Hypothesis

In research, a hypothesis is a suggested explanation of a phenomenon

A null hypothesis is a hypothesis which a researcher tries to disprove. Normally, the null hypothesis represents the current view/explanation of an aspect of the world that the researcher wants to challenge. Research methodology involves the researcher providing an alternative hypothesis, a research hypothesis, as an alternate way to explain the phenomenon.

The researcher tests the hypothesis to disprove the null hypothesis, not because he/she loves the research hypothesis, but because it would mean coming closer to finding an answer to a specific problem. The research hypothesis is often based on observations that evoke suspicion that the null hypothesis is not always correct.

A hypothesis is an assumption about relations between variables. Hypothesis can be defined as a logically conjectured relationship between two or more variables expressed in the form of a testable statement. Relationships are conjectured on the basis of the network of associations established in the theoretical framework formulated for the research study. Research Hypothesis is a predictive statement that relates an independent variable to a dependant variable. Hypothesis must contain at least one independent variable and one dependant variable.

1. Hypothesis is tentative, intelligent guesses as to the solution of the problem.
2. Hypothesis is a specific statement of prediction. It describes in concrete terms what you expect to happen in the study.
3. Hypothesis is an assumption about the population of the study.
4. It delimits the area of research and keeps the researcher on the right track

Problem (Vs) Hypothesis

1. Hypothesis is an assumption that can be tested and can be proved to be right or wrong.
2. A problem is a broad question which cannot be directly tested. A problem can be scientifically investigated after converting it into a form of hypothesis.

Characteristics of Hypothesis

1. **Conceptual Clarity** - It should be clear and precise.
2. **Specificity** - It should be specific and limited in scope.
3. **Consistency** - It should be consistent with the objectives of research.
4. **Testability** - It should be capable of being tested.
5. **Expectancy** - It should state the expected relationships between variables.
6. **Simplicity** - It should be stated as far as possible in simple terms.
7. **Objectivity** - It should not include value judgments, relative terms or any moral preaching.
8. **Theoretical Relevance** - It should be consistent with a substantial body of established or known facts or existing theory.
9. **Availability of Techniques** - Statistical methods should be available for testing the proposed hypothesis.

Sources of Hypothesis

1. **Discussions** with colleagues and experts about the problem, its origin and objectives in seeking a solution.
2. Examination of data and records for possible trends, peculiarities.
3. Review of similar studies.
4. Exploratory personal investigation / Observation.
5. Logical deduction from the existing theory.
6. Continuity of research.
7. Intuition and personal experience.

Types of Hypothesis

Descriptive Hypothesis

These are assumptions that describe the characteristics (such as size, form or distribution) of a variable. The variable may be an object, person, organisation, situation or event.

Examples: "Public enterprises are more amenable for centralized planning".

Relational Hypothesis [Explanatory Hypothesis]

These are assumptions that describe the relationship between two variables. The relationship suggested may be positive, negative or causal relationship. ***Examples: "Families with higher incomes spend more for recreation".***

Causal Hypothesis state that the existence of or change in one variable causes or leads to an effect on another variable. The first variable is called the independent variable and the latter is the dependant variable.

Null Hypothesis

When a hypothesis is stated negatively, it is called null hypothesis. It is a 'no difference', 'no relationship' hypothesis. ie., It states that, no difference exists between the parameter and statistic being compared to or no relationship exists between the variables being compared.It is

usually represented as H_O or H_0 .***Example: H_0: There is no relationship between a family's income and expenditure on recreation*.**

Alternate Hypothesis

It is the hypothesis that describes the researcher's prediction that, there exist a relationship between two variables or it is the opposite of null hypothesis. It is represented as H_A or H_1. Example:

H_A: There is a definite relationship between family's income and expenditure on recreation.

Functions or Role of Hypothesis

1. It gives a definite point to the investigation and provides direction to the study.
2. It determines the data needs.
3. It specifies the sources of data.
4. It suggests which type of research is likely to be more appropriate.
5. It determines the most appropriate technique of analysis.
6. It contributes to the development of theory.

Variable

A variable is anything that can vary can be considered as a variable. A variable is anything that can take on differing or varying values For example, Age, Production units, Absenteeism, Sex, Motivation, Income, Height, and Weight etc. **Note**: The values can differ at various times for the same object or person (or) at the same time for different objects or persons.

Variable / Attribute

A variable is a characteristic that takes on two or more values whereas, an attribute is a specific value on a variable (qualitative).For example; the variable SEX/GENDER **has 2 attributes** - Male and Female. The variable AGREEMENT has 5 attributes - Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree.

Types of Variables

Explanatory vs Extraneous Variable

The variables selected for analysis are called explanatory variables and all other variables that are not related to the purpose of the study but may affect the dependant variable are extraneous.

Dependant vs Independent Variable

The variable that changes in relationship to changes in another variable(s) is called dependant variable. The variable whose change results in the change in another variable is called an independent variable.

OR

An independent variable is the one that influences the dependant variable in either a positive or negative way.

Variables as they are measured

Nominal - measurement scale in which numbers are used as names of categories; i.e., categorizes without order. (Frequency data)

Ordinal - measurement scale that categorizes AND indicates relative amount or rank-order of a characteristic. (Ordered data)

Interval - measurement scale that categorizes, indicates relative amount or rank-order of a characteristic, and has units of equal distance between consecutive points on the scale. (Score data)

Ratio - measurement scale that categorizes, indicates relative amount or rank-order of a characteristic, has units of equal distance between consecutive points on the scale, and compares terms as ratios of one to another (i.e. has a true zero point). Rarely used in social science research.

NOTE - measurement scales are hierarchical, with each higher level of measurement provided greater information and permitting more stringent statistical analyses to be performed.

OPERATIONAL DEFINITION - a definition expressed in terms of the processes or operations and conditions that are being used to measure the characteristic under study.

Diagram100

<u>The Research Question and Hypothesis</u>

What is a research question?	This is the question that you are trying to answer when you do research on a topic or write a research report.
Should a research question be general or specific?	It should be as specific as possible. In some cases, you may make two or more research questions to cover a complex topic.
What is an example of a research question?	For example, if you are studying the effects of sleep on reflexes, you might formulate the following research question: *What are the effects of sleep on reflexes?* A similar question might be: *Does sleep have an effect on reflexes?* Or: *Is maximum reflex efficiency achieved after eight hours of sleep?* The goal of your research is to find the answer to the research question.
What is a hypothesis?	A hypothesis is a statement that can be proved or disproved. A research question can be made into a hypothesis by changing it into a statement. For

	example, the third research question above can be made into the hypothesis: *Maximum reflex efficiency is achieved after eight hours of sleep.*
What is a null hypothesis?	A null hypothesis (abbreviated H_0) is a hypothesis to be disproved. The hypothesis above can be turned into a working null hypothesis simply by adding "not". *Maximum reflex efficiency is not achieved after eight hours of sleep.* Another null hypothesis is: *Sleep does not have an effect on reflexes.* Null hypotheses are used in the sciences. In the scientific method, a null hypothesis is formulated, and then a scientific investigation is conducted to try to disprove the null hypothesis. If it can be disproved, another null hypothesis is constructed and the process is repeated. As an example, we might begin with the null hypothesis: *Sleep does not affect reflexes.* If we can disprove this, we find that sleep does have an effect. We might then go to the next null hypothesis: *Different amounts of sleep have the same effect on reflexes.* If we can disprove this, we can go to: *Maximum reflex efficiency is not achieved after eight hours of sleep.* And so on. At each stage in the investigation, we conduct experiments designed to try to disprove the hull hypothesis.
What is the relationship between the null hypothesis and the thesis statement of a research report?	A generalized form of the final hypothesis (not the null hypothesis) can be used as a thesis statement. For example, if our final proved hypothesis is: *Maximum reflex efficiency is achieved after eight hours of sleep* we might generalize this to a thesis statement such as: *This investigation demonstrated that sleep has an effect on reflex efficiency and that, in fact, maximum reflex efficiency is achieved after a specific period of slee*

Test of hypothesis

Nonparametric tests

Many nonparametric procedures are based on ranked data. Data are ranked by ordering them from lowest to highest and assigning them, in order, the integer values from 1 to the sample size. Ties are resolved by assigning tied values the mean of the ranks they would have received if there were no ties, e.g., 117, 119, 119, 125, 128 becomes 1, 2.5, 2.5, 4, 5. (If the two 119s were not tied, they would have been assigned the ranks 2 and 3. The mean of 2 and 3 is 2.5.)

Advantages of nonparametric procedures

(1) Nonparametric test make less stringent demands of the data. For standard parametric procedures to be valid, certain underlying conditions or assumptions must be met, particularly for smaller sample sizes. The one-sample t test, for example, requires that the observations be drawn from a normally distributed population. For two independent samples, the t test has the additional requirement that the population standard deviations be equal. If these assumptions/conditions are violated, the resulting P-values and confidence intervals may not be trustworthy[3]. However, normality is not required for the Wilcoxon signed rank or rank sum tests to produce valid inferences about whether the median of a symmetric population is 0 or whether two samples are drawn from the same population.

(2) Nonparametric procedures can sometimes be used to get a quick answer with little calculation.

Two of the simplest nonparametric procedures are the sign test and median test. The *sign test* can be used with paired data to test the hypothesis that differences are equally likely to be positive or negative, (or, equivalently, that the median difference is 0). For small samples, an exact test of whether the proportion of positives is 0.5 can be obtained by using a binomial distribution. For large samples, the test statistic is

$(\text{plus} - \text{minus})^2 / (\text{plus} + \text{minus})$, where *plus* is the number of positive values and *minus* is the number of negative values. Under the null hypothesis that the positive and negative values are equally likely, the test statistic follows the chi-square distribution with 1 degree of freedom. Whether the sample size is small or large, the sign test provides a quick test of whether two paired treatments are equally effective simply by counting the number of times each treatment is better than the other.

Example: 15 patients given both treatments A and B to test the hypothesis that they perform equally well. If 13 patients prefer A to B and 2 patients prefer B to A, the test statistic is $(13 - 2)^2 / (13 + 2)$ [= 8.07] with a corresponding P-value of 0.0045. The null hypothesis is therefore rejected.

The *median test* is used to test whether two samples are drawn from populations with the same median. The median of the combined data set is calculated and each original observation is classified according to its original sample (A or B) and whether it is less than or greater than the overall median. The chi-square test for homogeneity of proportions in the resulting 2-by-2 table tests whether the population medians are equal.

(3) **Nonparametric methods** provide an air of objectivity when there is no reliable (universally recognized) underlying scale for the original data and there is some concern that the results of standard parametric techniques would be criticized for their dependence on an artificial metric. For example, patients might be asked whether they feel *extremely uncomfortable* / *uncomfortable* / *neutral* / *comfortable* / *very comfortable*. What scores should be assigned to the comfort categories and how do we know whether the outcome would change dramatically with a slight change in scoring? Some of these concerns are blunted when the data are converted to ranks[4].

(4) A historical appeal of rank tests is that it was easy to construct tables of exact critical values, provided there were no ties in the data. The same critical value could be used for all data sets with the same number of observations because every data set is reduced to the ranks 1... *n*. However, this advantage has been eliminated by the ready availability of personal computers[5].

(5) Sometimes the data do not constitute a random sample from a larger population. The data in hand are all there are. Standard parametric techniques based on sampling from larger populations are no longer appropriate. Because there are no larger populations, there are no population

parameters to estimate. Nevertheless, certain kinds of nonparametric procedures can be applied to such data by using ***randomization models*.**

Consider, for example, a situation in which a company's workers are assigned in haphazard fashion to work in one of two buildings. After yearly physicals are administered, it appears that workers in one building have higher lead levels in their blood. Standard sampling theory techniques are inappropriate because the workers do not represent samples from a large population--there is no large population. The randomization model, however, provides a means for carrying out statistical tests in such circumstances. The model states that if there were no influence exerted by the buildings, the lead levels of the workers in each building should be no different from what one would observe after combining all of the lead values into a single data set and dividing it in two, at random, according to the number of workers in each building. The stochastic component of the model, then, exists only in the analyst's head; it is not the result of some physical process, except insofar as the haphazard assignment of workers to buildings is truly random.

Of course, randomization tests cannot be applied blindly any more than normality can automatically be assumed when performing a t test. (Perhaps, in the lead levels example, one building's workers tend to live in urban settings while the other building's workers live in rural settings. Then the randomization model would be inappropriate.) Nevertheless, there will be many situations where the less stringent requirements of the randomization test will make it the test of choice. In the context of randomization models, randomization tests are the ONLY legitimate tests; standard parametric test are valid only as approximations to randomization tests.[6]

Disadvantages of nonparametric procedures

Such a strong case has been made for the benefits of nonparametric procedures that some might ask why parametric procedures aren't abandoned entirely in favor of nonparametric methods!

The major disadvantage of nonparametric techniques is contained in its name. Because the procedures are *nonparametric*, there are no parameters to describe and it becomes more difficult to make quantitative statements about the actual difference between populations. (For example, when the sign test says two treatments are different, there's no confidence interval and the test doesn't say by how much the treatments differ.) However, it is sometimes possible with the right software to compute estimates (and even confidence intervals!) for medians, differences between medians. However, the calculations are often too tedious for pencil-and-paper. A computer is required. As statistical software goes though its various iterations, such confidence intervals may become readily available, but I'm still waiting!

The second disadvantage is that nonparametric procedures throw away information! The sign test, for example, uses only the signs of the observations. Ranks preserve information about the order of the data but discard the actual values. Because information is discarded, nonparametric

procedures can never be as powerful (able to detect existing differences) as their parametric counterparts when parametric tests can be used.

Conceptual Framework vs theoretical framework

Conceptual framework

After formulating the theoretical framework, the researcher has to develop the conceptual framework of the study. A **concept** is an image or symbolic representation of an abstract idea. Chinn and Kramer (1999) define a concept as a "complex mental formulation of experience".

While the theoretical framework is the theory on which the study is based, the conceptual framework is the operationalization of the theory. Conceptual framework is the researcher's own position on the problem and gives direction to the study. It may be an adaptation of a model used in a previous study, with modifications to suit the inquiry. Aside from showing the direction of the study, through the conceptual framework, the researcher can be able to show the relationships of the different constructs that he wants to investigate.

1. The conceptual framework is the schematic diagram which shows the variables included in the study.
2. Arrows or line should be properly placed and connected between boxes to show the relationship between the independent and dependent variables.
3. All the independent and dependent variables should be clearly discussed and explained how these would influence the results of the study.

Theoretical Framework

The theoretical framework of the study is a structure that can hold or support a theory of a research work. It presents the theory which explains <u>why the problem under study exists</u>. Thus, the theoretical framework is but a theory that serves as a basis for conducting research.

1. The theoretical framework consists of theories, principles, generalizations and research findings which are closely related to the present study under investigation. It is in this framework where the present research problem understudy evolved.
2. Authors of these theories and principles should be cited. As much as possible research findings and theories should be correct.

Formulating the Theoretical Framework

Purpose:

1. It helps the researcher see clearly the variables of the study;
2. It can provide him with a general framework for data analysis;
3. It is essential in preparing a research proposal using descriptive and experimental methods.

Based on the foregoing example, how should the theoretical framework formulated?

1. Specifies the theory used as basis for the study
2. Mentions the proponents of the theory
3. Cites the main points emphasized in the theory
4. supports his exposition of the theory by ideas from other experts;
5. Illustrates his theoretical framework by means of a diagram; and,
6. Reiterates his theoretical proposition in the study.