

Hypothesis Development and Testing

Sendil Mourougan¹, Dr. K. Sethuraman²

^{1,2} (Department of Business Administration, Annamalai university, India)

Abstract : Hypothesis testing is an important activity of evidence-based research. A well worked up hypothesis is half the answer to the research question. For this, both knowledge of the subject derived from extensive review of the literature and working knowledge of basic statistical concepts are desirable. This paper discusses the methods of working up a good hypothesis and statistical concepts of hypothesis testing.

Keywords: Effect size, Hypothesis testing, Type I error, Type II error

I. Introduction

A research hypothesis is the statement created by researchers when they speculate upon the outcome of a research or experiment. Every true experimental design must have this statement at the core of its structure, as the ultimate aim of any experiment. The hypothesis is generated via a number of means, but is usually the result of a process of inductive reasoning where observations lead to the formation of a theory. Scientists then use a large battery of deductive methods to arrive at a hypothesis that is testable, falsifiable and realistic. The first step in the scientific process is not observation but the generation of a hypothesis which may then be tested critically by observations and experiments. The goal of the scientist's efforts is not the verification but the falsification of the initial hypothesis.

II. Hypothesis

Research usually starts with a problem. Questions, objectives and hypotheses provide a specific restatement and clarification of the problem statement/research question. Hypothesis is a tentative explanation that accounts for a set of facts and can be tested by further investigation. Hypothesis should be statements expressing the relation between two or more measurable variables. It should carry clear implications for testing the stated relations.

- A hypothesis may be precisely defined as a tentative proposition suggested as a solution to a problem or as an explanation of some phenomenon. (Ary, Jacobs and Razavieh, 1984)
- A hypothesis is a conjectural statement of the relation between two or more variables. (Kerlinger, 1956)
- Hypothesis is a formal statement that presents the expected relationship between an independent and dependent variable. (Creswell, 1994)
- Hypothesis relates theory to observation and observation to theory. (Ary, Jacobs and Razavieh, 1984)
- Hypotheses are relational propositions. (Kerlinger, 1956)

2.1 Purpose and Importance of Hypotheses in a Scientific Research

- It provides a tentative explanation of phenomena and facilitates the extension of knowledge in an area.
- It provides the investigator with a relational statement that is directly testable in a research study.
- It provides direction to the research.
- It provides a framework for reporting conclusions of the study.
- It could be considered as the working instrument of theory. Hypotheses can be deduced from theory and from other hypotheses.
- It could be tested and shown to be probably supported or not supported, apart from man's own values and opinions.

2.2 Characteristics of a Good Hypothesis

A good hypothesis must be based on a good research question.

- Hypothesis should be simple, specific and stated in advance.
- It must have explanatory power.
- It must state the expected relationship between variables.
- It must be testable.
- It should be consistent with the existing body of knowledge.
- It should be stated as simply and concisely as possible.

2.3 Types of Hypotheses

For the purpose of testing statistical significance, hypotheses are classified into two types:

(a) Null Hypotheses

A null hypothesis is a statement that there is no actual relationship between variables. (H_0 or H_N). A null hypothesis may read, "There is no difference between.....". H_0 states the opposite of what the experimenter would expect or predict. The final conclusion of the investigator will either retain a null hypothesis or reject a null hypothesis in favor of an alternative hypothesis. Not rejecting H_0 does not really mean that H_0 is true. There might not be enough evidence against H_0 . Once the null hypothesis has been stated, it is easy to construct the alternative hypothesis. It is essentially the statement that the null hypothesis is false. Example can be "There is no significant difference in the anxiety level of children of High IQ and those of low IQ."

(b) Alternate Hypotheses

An alternative hypothesis is a statement that suggests a potential outcome that the researcher may expect. (H_1 or H_A). It is established only when a null hypothesis is rejected. Often an alternative Hypothesis is the desired conclusion of the investigator. The two types of alternative hypothesis are: (a) Directional Hypothesis and (b) Non-directional Hypothesis.

- **Directional Hypothesis:** It is a type of alternative hypothesis that specifies the direction of expected findings. Sometimes directional hypothesis are created to examine the relationship among variables rather than to compare groups. Directional hypothesis may read, "...is more than..", "...will be lesser.." Example can be "Children with high IQ will exhibit more anxiety than children with low IQ"
- **Non-directional Hypothesis:** It is a type of alternative hypothesis in which no definite direction of the expected findings is specified. The researcher may not know what can be predicted from the past literature. It may read, "...there is a difference between.." Example can be "There is a difference in the anxiety level of the children of high IQ and those of low IQ."

III. Hypothesis Formulation

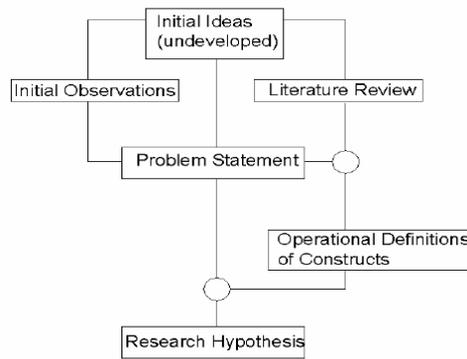
Formulation of Hypothesis differs with the method of research conducted such as (a) Qualitative method and (b) Quantitative method.

- **Qualitative Method:** The researcher uses questions, not objectives or hypothesis. He poses a central question, which is being examined in the study in the most general form-"the broadest question that can be asked in a study." Several sub-questions are raised related to the central question to narrow the focus of study but to leave the questioning at an open end. They are under continual review and restructuring along the course of research. Questions begin with "what", "how". They focus on a single concept. The questions use non-directional terminology.
- **Quantitative Method:** The researcher uses questions and hypothesis to compose and focus the purpose of study. Hypothesis or research questions used to compare, relate and describe variables. The research follows from a test of theory and the question and the hypothesis or the research questions are included in the theory. Independent and dependant variables are measured separately. It is not a combination of both either a research questions or a hypothesis.

Hypothesis needs to be structured before the data-gathering and interpretation phase of the research. A well-grounded hypothesis indicates that the researcher has sufficient knowledge in the area to undertake the investigation. The hypothesis gives direction to the collection and interpretation of data.

Consider the example of a simple association between two variables, Y and X.

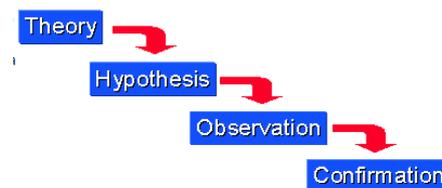
1. Y and X are associated (or, there is an association between Y and X).
 2. Y is related to X (or, Y is dependent on X).
 3. As X increases, Y decreases (or, increases in values of X appear to effect reduction in values of Y).
- The first hypothesis provides a simple statement of association between Y and X. Nothing is indicated about the association that would allow the researcher to determine which variable, Y or X, would tend to cause the other variable to change in value.
 - The second hypothesis is also a simple statement of association between Y and X, but this time it may be inferred that values of Y are in some way contingent upon the condition of the X variable.
 - The third hypothesis is the most specific of the three. Not only does it say that Y and X are related and that Y is dependent on X for its value, but it also reveals something more about the nature of the association between the two variables.



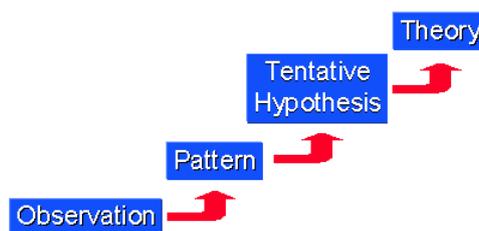
Deductive and Inductive Approaches

There are two broad methods of reasoning such as deductive and inductive approaches. The main difference between inductive and deductive approaches to research is that whilst a deductive approach is aimed and testing theory, an inductive approach is concerned with the generation of new theory emerging from the data.

- Deductive Approach:** Deductive reasoning works from the more general to the more specific. Sometimes this is informally called a "top-down" approach. Researcher might begin with thinking up a theory about our topic of interest. He then narrows that down into more specific hypotheses that we can test. Researcher narrows down even further when he collects observations to address the hypotheses. This ultimately leads us to be able to test the hypotheses with specific data -- a confirmation (or not) of our original theories.



- Inductive Approach:** Inductive reasoning works the other way, moving from specific observations to broader generalizations and theories. Informally, sometimes this is called as a "bottom up" approach. In inductive reasoning, researcher begin with specific observations and measures, begin to detect patterns and regularities, formulate some tentative hypotheses that he can explore, and finally end up developing some general conclusions or theories.



These two methods of reasoning have a very different "feel" while conducting a research. Inductive reasoning, by its very nature, is more open-ended and exploratory, especially at the beginning. Deductive reasoning is narrower in nature and is concerned with testing or confirming hypotheses. Even though a particular study may look like it's purely deductive, most social research involves both inductive and deductive reasoning processes at some time in the project.

IV. Hypothesis Testing

Hypothesis testing is a statistical technique that is used in a variety of situations. Testing a hypothesis involves

- Deducing the consequences that should be observable if the hypothesis is correct.
- Selecting the research methods that will permit the observation, experimentation, or other procedures necessary to show whether or not these do occur.
- Applying this method and gathering the data that can be analyzed to indicate whether or not the hypothesis is supported.

There are two possibilities

- Nothing Happened - Null Hypothesis (H_0)
- Something Happened - Alternative Hypothesis (H_1)

Though the technical details differ from situation to situation, all hypothesis tests use the same core set of terms and concepts. The following descriptions of common terms and concepts refer to a hypothesis test in which the means of two populations are being compared.

In the field of statistics, a hypothesis is a claim about some aspect of a population. A hypothesis test allows us to test the claim about the population and find out how likely it is to be true. The hypothesis test consists of several components; two statements, the null hypothesis and the alternative hypothesis, the test statistic and the critical value, which in turn gives us the p-value and the rejection region (α), respectively.

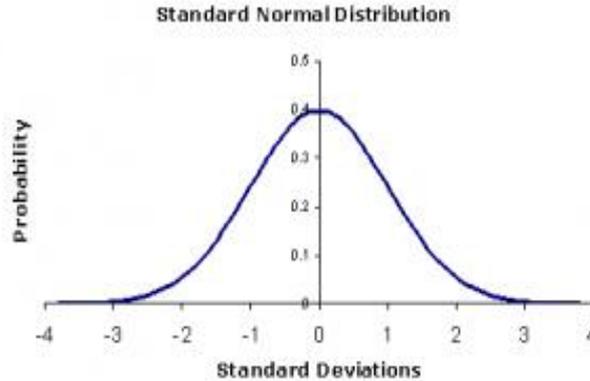
- **Null Hypothesis:** The null hypothesis, denoted as H_0 is the statement that the value of the parameter is, in fact, equal to the claimed value. It is assumed that the null hypothesis is true until the researcher prove that it is not.
- **Alternate Hypothesis:** The alternative hypothesis, denoted as H_1 is the statement that the value of the parameter differs in some way from the null hypothesis. The alternative hypothesis can use the symbols $<$, $>$, or \neq .
- **Test Statistic:** The test statistic is the tool researcher use to decide whether or not to reject the null hypothesis. It is obtained by taking the observed value (the sample statistic) and converting it into a standard score under the assumption that the null hypothesis is true. The test statistic depends fundamentally on the number of observations that are being evaluated. It differs from situation to situation. The whole notion of hypothesis rests on the ability to specify (exactly or approximately) the distribution that the test statistic follows.
- **Significance:** The significance level is a measure of the statistical strength of the hypothesis test. It is often characterized as the probability of incorrectly concluding that the null hypothesis is false. The significance level should be specified up front. The significance level is typically one of three values: 10%, 5%, or 1%. A 1% significance level represents the strongest test of the three. For this reason, 1% is a higher significance level than 10%.
- **Power:** Related to significance, the power of a test measures the probability of correctly concluding that the null hypothesis is true. Power is not something that researcher can choose. It is determined by several factors, including the significance level selected and the size of the difference between the things researcher is trying to compare. Unfortunately, significance and power are inversely related. Increasing significance decreases power. This makes it difficult to design experiments that have both very high significance and power.
- **Critical Value:** The critical value is the standard score that separates the rejection region (α) from the rest of a given curve. The critical value in a hypothesis test is based on two things: the distribution of the test statistic and the significance level. The critical value(s) refer to the point in the test statistic distribution that give the tails of the distribution an area (meaning probability) exactly equal to the significance level that was chosen.
- **Decision:** Your decision to reject or accept the null hypothesis is based on comparing the test statistic to the critical value. If the test statistic exceeds the critical value, you should reject the null hypothesis. In this case, you would say that the difference between the two population means is significant. Otherwise, you accept the null hypothesis.
- **p-Value:** It is the area to the left or right of the test statistic. The p-value of a hypothesis test gives another way to evaluate the null hypothesis. The p-value represents the highest significance level at which particular test statistic would justify rejecting the null hypothesis. For example, if the significance level of 5% is chosen, and the p-value turns out to be .03 (or 3%), it would be justified in rejecting the null hypothesis.
- **Type I (also known as ' α ') Errors:** A Type I Error occur when researcher is rejecting a null hypothesis and accepting alternate hypothesis.
- **Type II (also known as ' β ') Errors:** A Type II Error occur when researcher is accepting null hypothesis and rejecting alternate hypothesis.
- **Z-Value:** Z value is a measure of standard deviation i.e. how many standard deviation away from mean is the observed value. For example, the value of z value = +1.8 can be interpreted as the observed value is +1.8 standard deviations away from the mean. p-values are probabilities. Both these statistics terms are associated with the standard normal distribution. p-values associated with each z-value can be looked up in Z-table. Below is the formula to calculate z value:

$$z = \frac{x - \mu}{\sigma}$$

μ = Mean

σ = Standard Deviation

Here X is the point on the curve, μ is mean of the population and σ is standard deviation of population.



One-Tailed and Two-Tailed Test

The two ways of carrying out statistical significance test of a characteristic, drawn from the population, with respect to the test statistic, are a one-tailed test and two-tailed test. The former refers to a test of null hypothesis, in which the alternative hypothesis is articulated directionally. However, if the alternative hypothesis is not exhibited directionally, then it is known as two-tailed test of the null hypothesis. In the field of research and experiments, it pays to know the difference between one-tailed and two-tailed test, as they are quite commonly used in the process.

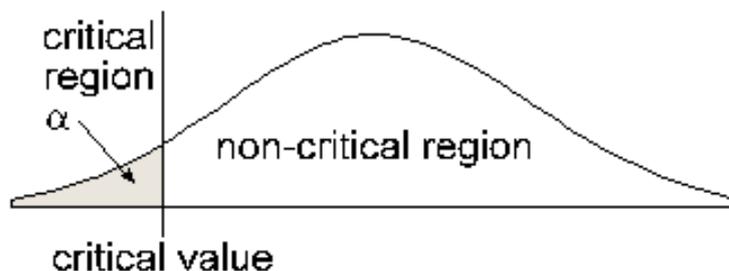
Definition of One-Tailed Test

One-tailed test alludes to the significance test in which the region of rejection appears on one end of the sampling distribution. It represents that the estimated test parameter is greater or less than the critical value. When the sample tested falls in the region of rejection, i.e. either left or right side, as the case may be, it leads to the acceptance of alternative hypothesis rather than the null hypothesis. It is primarily applied in chi-square distribution; that ascertains the goodness of fit.

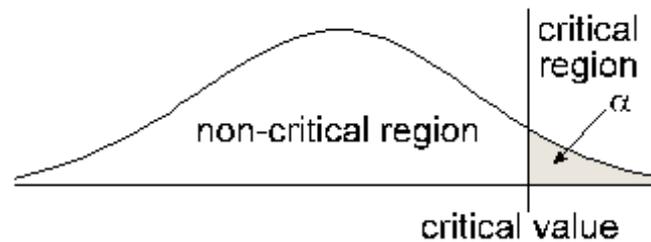
In this statistical hypothesis test, all the critical region, related to α , is placed in any one of the two tails. One-tailed test can be:

- Left-tailed test: When the population parameter is believed to be lower than the assumed one, the hypothesis test carried out is the left-tailed test.
- Right-tailed test: When the population parameter is supposed to be greater than the assumed one, the statistical test conducted is a right-tailed test.

Left Tailed Test



H_1 : parameter < value. Notice the inequality points to the left.
Decision Rule: Reject H_0 if test parameter. < critical value.

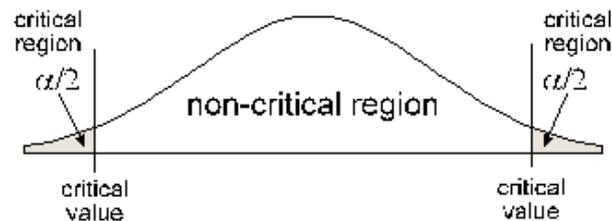
Right Tailed Test

H_1 : parameter $>$ value. Notice the inequality points to the right
 Decision Rule: Reject H_0 if test parameter. $>$ Critical value.

Definition of Two-tailed Test

The two-tailed test is described as a hypothesis test, in which the region of rejection or say the critical area is on both the ends of the normal distribution. It determines whether the sample tested falls within or outside a certain range of values. Therefore, an alternative hypothesis is accepted in place of the null hypothesis, if the calculated value falls in either of the two tails of the probability distribution.

In this test, α is bifurcated into two equal parts, placing half on each side, i.e. it considers the possibility of both positive and negative effects. It is performed to see, whether the estimated parameter is either above or below the assumed parameter, so the extreme values, work as evidence against the null hypothesis.

Two Tailed Test

H_1 : parameter not equal value. Another way to write not equal is $<$ or $>$.

Notice the inequality points to both sides

Decision Rule: Reject H_0 if test parameter $<$ critical value. (Left) or test parameter. $>$ critical value. (right)

Key Differences between One-Tailed and Two-Tailed Test

The fundamental differences between one-tailed and two-tailed test, is explained below in points:

- One-tailed test, as the name suggest is the statistical hypothesis test, in which the alternative hypothesis has a single end. On the other hand, two-tailed test implies the hypothesis test; wherein the alternative hypothesis has dual ends.
- In the one-tailed test, the alternative hypothesis is represented directionally. Conversely, the two-tailed test is a non-directional hypothesis test.
- In a one-tailed test, the region of rejection is either on the left or right of the sampling distribution. On the contrary, the region of rejection is on both the sides of the sampling distribution.
- A one-tailed test is used to ascertain if there is any relationship between variables in a single direction, i.e. left or right. As against this, the two-tailed test is used to identify whether or not there is any relationship between variables in either direction.
- In a one-tailed test, the test parameter calculated is more or less than the critical value. Unlike, two-tailed test, the result obtained is within or outside critical value.
- When an alternative hypothesis has ' \neq ' sign, then a two-tailed test is performed. In contrast, when an alternative hypothesis has ' $>$ ' or ' $<$ ' sign, then one-tailed test is carried out.

To sum up, we can say that the basic difference between one-tailed and two-tailed test lies in the direction, i.e. in case the research hypothesis entails the direction of interrelation or difference, then one-tailed test is applied, but if the research hypothesis does not signify the direction of interaction or difference, we use two-tailed test. This explains how to determine if the test is a left tail, right tail, or two-tail test. The type of test is determined by the Alternative Hypothesis (H_1).

Steps to Perform Hypothesis Testing

There are 7 steps for a hypothesis test.

1. The first step in hypothesis testing is to specify the null hypothesis (H_0) and the alternative hypothesis (H_1). If the research concerns whether one method of presenting pictorial stimuli leads to better recognition than another, the null hypothesis would be that there is no difference between methods ($H_0: \mu_1 - \mu_2 = 0$). The alternative hypothesis would be $H_1: \mu_1 \neq \mu_2$.
2. The next step is to select a significance level. Typically the 0.05 or the 0.01 level is used.
3. The third step is to calculate a statistic analogous to the parameter specified in the null hypothesis. If the null hypothesis were defined by the parameter $\mu_1 - \mu_2$, then the statistic $M_1 - M_2$ would be computed.
4. The fourth step is to calculate the probability value (often called the p value). The p value is the probability of obtaining a statistic as different or more different from the parameter specified in the null hypothesis as the statistic computed from the data. The calculations are made assuming that the null hypothesis is true.
5. The probability value computed in Step 4 is compared with the significance level chosen in Step 2. If the probability is less than or equal to the significance level, then the null hypothesis is rejected; if the probability is greater than the significance level then the null hypothesis is not rejected. When the null hypothesis is rejected, the outcome is said to be "statistically significant" when the null hypothesis is not rejected then the outcome is said to be "not statistically significant."
6. If the outcome is statistically significant, then the null hypothesis is rejected in favor of the alternative hypothesis. If the rejected null hypothesis were that $\mu_1 - \mu_2 = 0$, then the alternative hypothesis would be that $\mu_1 \neq \mu_2$. If M_1 were greater than M_2 then the researcher would naturally conclude that $\mu_1 \geq \mu_2$.
7. The final step is to describe the result and the statistical conclusion in an understandable way. Be sure to present the descriptive statistics as well as whether the effect was significant or not.

V. Conclusion

The empirical approach to research cannot eliminate uncertainty completely. At the best, it can quantify uncertainty. This uncertainty can be of 2 types: Type I error (falsely rejecting a null hypothesis) and type II error (falsely accepting a null hypothesis). The acceptable magnitudes of type I and type II errors are set in advance and are important for sample size calculations. Another important point to remember is that we cannot 'prove' or 'disprove' anything by hypothesis testing and statistical tests. We can only knock down or reject the null hypothesis and by default accept the alternative hypothesis. If we fail to reject the null hypothesis, we accept it by default. In this article, we have looked at the complete process of undertaking hypothesis testing. Initially, we looked at the concept of hypothesis followed by the types of hypothesis and way to validate hypothesis to make an informed decision.

References

- [1]. Daniel W. W. In: Biostatistics. 7th ed. New York: John Wiley and Sons, Inc; 2002. Hypothesis testing; pp. 204–294.
- [2]. Hulley S. B, Cummings S. R, Browner W. S, Grady D, Hearst N, Newman T. B. 2nd ed. Philadelphia: Lippincott Williams and Wilkins; 2001. Getting ready to estimate sample size: Hypothesis and underlying principles In: Designing Clinical Research-An epidemiologic approach; pp. 51–63.
- [3]. Medawar P. B. Philadelphia: American Philosophical Society; 1969. Induction and intuition in scientific thought. Popper K. Unended Quest. An Intellectual Autobiography. Fontana Collins; p. 42.
- [4]. Wulff H. R, Pedersen S. A, Rosenberg R. Oxford: Blackwell Scientific Publications; Empirism and Realism: A philosophical problem. In: Philosophy of Medicine.