HYPOTHESIS TESTING

Social science research, and by extension business research, uses a number of different approaches to study a variety of issues. This research may be a very informal, simple process or it may be a formal, somewhat sophisticated process. Regardless of the type of research, all research begins with a generalized idea in the form of a research question or a hypothesis. A research question usually is posed in the beginning of a research effort or in a specific area of study that had little formal research. A research question may take the form of a basic question about some issue or phenomena or a question about the relationship between two or more variables. For example, a research question might be: "Do flexible work hours improve employee productivity?" Another question might be: "How do flexible hours influence employees' work?"

A hypothesis differs from a research question; it is more specific and makes a prediction. It is a tentative statement about the relationship between two or more variables. The major difference between a research question and a hypothesis is that a hypothesis predicts an experimental outcome. For example, a hypothesis might state: "There is a positive relationship between the availability of flexible work hours and employee productivity."

Hypotheses provide the following benefits:

1. They determine the focus and direction for a research effort.
2. Their development forces the researcher to clearly state the purpose of the research activity.
3. They determine what variables will not be considered in a study, as well as those that will be considered.
4. They require the researcher to have an operational definition of the variables of interest.

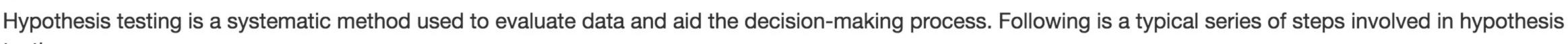The worth of a hypothesis often depends on the researcher's skills. Since the hypothesis is the basis of a research study, it is necessary for the hypothesis to be developed with a great deal of thought and contemplation. There are basic criteria to consider when developing a hypothesis, in order to ensure that it meets the study and the research. A good hypothesis should:

1. Have logical consistency. Based on the current research literature and knowledge base, does this hypothesis make sense?
2. Be in step with the current literature and/or provide a good basis for any differences. Though it does not have to support the current body of literature, it is necessary to provide a good rationale for stepping away from the mainstream.
3. Be testable. If one cannot design the means to conduct the research, the hypothesis means nothing.
4. Be stated in clear and simple terms in order to reduce confusion.

[Photo by: oily]

HYPOTHESIS TESTING PROCESS

Hypothesis testing is a systematic method used to evaluate data and aid the decision-making process. Following is a typical series of steps involved in hypothesis testing:

1. State the hypotheses of interest
2. Determine the appropriate test statistic
3. Specify the level of statistical significance
4. Determine the decision rule for rejecting or not rejecting the null hypothesis
5. Collect the data and perform the needed calculations
6. Decide to reject or not reject the null hypothesis

Each step in the process will be discussed in detail, and an example will follow the discussion of the steps.

STATING THE HYPOTHESES.

A research study includes at least two hypotheses—the null hypothesis and the alternative hypothesis. The hypothesis being tested is referred to as the null hypothesis and it is designated as H It also is referred to as the hypothesis of no difference and should include a statement of equality (=, ≥, or ≤). The alternative hypothesis presents the alternative to the null and includes a statement of inequality (≠). The null hypothesis and the alternative hypothesis are complementary.

The null hypothesis is the statement that is believed to be correct throughout the analysis, and it is the null hypothesis upon which the analysis is based. For example, the null hypothesis might state that the average age of entering college freshmen is 21 years.

H_0 The average age of entering college freshman = 21 years

If the data one collects and analyzes indicates that the average age of entering college freshmen is greater than or less than 21 years, the null hypothesis is rejected. In this case the alternative hypothesis could be stated in the following three ways: (1) the average age of entering college freshman is not 21 years (the average age of entering college freshman is ≠ 21); (2) the average age of entering college freshman is less than 21 years (the average age of entering college freshmen < 21); or (3) the average age of entering college freshman is greater than 21 years (the average age of entering college freshmen > 21 years).

The choice of which alternative hypothesis to use is generally determined by the study's objective. The preceding second and third examples of alternative hypotheses involve the use of a "one-tailed" statistical test. This is referred to as "one-tailed" because a direction (greater than [>] or less than [<]) is implied in the statement. The first example represents a "two-tailed" test. There is inequality expressed (age ≠ 21 years), but the inequality does not imply direction. One-tailed tests are used more often in management and marketing research because there usually is a need to imply a specific direction in the outcome. For example, it is more likely that a researcher would want to know if Product A performed better than Product B (Product A performance > Product B performance), or vice versa (Product A performance < Product B performance), than whether Product A performed differently than Product B (Product A performance ≠ Product B performance). Additionally, more useful information is gained by knowing that employees who work from 7:00 a.m. to 4:00 p.m. are more productive than those who work from 3:00 p.m. to 12:00 a.m. (early shift employee production > late shift employee production), rather than simply knowing that these employees have different levels of productivity (early shift employee production ≠ late shift employee production).

Both the alternative and the null hypotheses must be determined and stated prior to the collection of data. Before the alternative and null hypotheses can be formulated it is necessary to decide on the desired or expected conclusion of the research. Generally, the desired conclusion of the study is stated in the alternative hypothesis. This is true as long as the null hypothesis can include a statement of equality. For example, suppose that a researcher is interested in exploring the effects of amount of study time on tests scores. The researcher believes that students who study longer perform better on tests. Specifically, the research suggests that students who spend four hours studying for an exam will get a better score than those who study two hours. In this case the hypotheses might be:

H_0 The average test scores of students who study 4 hours for the test = the average test scores of those who study 2 hours.

H_1 The average test score of students who study 4 hours for the test < the average test scores of those who study 2 hours.

As a result of the statistical analysis, the null hypothesis can be rejected or not rejected. As a principle of rigorous scientific method, this subtle but important point means that the null hypothesis cannot be accepted. If the null is rejected, the alternative hypothesis can be accepted; however, if the null is not rejected, we can't conclude that the null hypothesis is true. The rationale is that evidence that supports a hypothesis is not conclusive, but evidence that negates a hypothesis is ample to discredit a hypothesis. The analysis of study time and test scores provides an example. If the results of one study indicate that the test scores of students who study 4 hours are significantly better than the test scores of students who study two hours, the null hypothesis can be rejected because the researcher has found one case when the null is not true. However, if the results of the study indicate that the test scores of those who study 4 hours are not significantly better than those who study 2 hours, the null hypothesis cannot be rejected. One also cannot conclude that the null hypothesis is accepted because these results are only one set of score comparisons. Just because the null hypothesis is true in one situation does not mean it is always true.

DETERMINING THE APPROPRIATE TEST STATISTIC.

The appropriate test statistic (the statistic to be used in statistical hypothesis testing) is based on various characteristics of the sample population of interest, including sample size and distribution. The test statistic can assume many numerical values. Since the value of the test statistic has a significant effect on the decision, one must use the appropriate statistic in order to obtain meaningful results. Most test statistics follow this general pattern:

For example, the appropriate statistic to use when testing a hypothesis about a population means is:

In this formula Z = test statistic, $\bar{X}$ = mean of the sample, μ = mean of the population, σ = standard deviation of the sample, and η = number in the sample.

SPECIFYING THE STATISTICAL SIGNIFICANCE SEVEL.

As previously noted, one can reject a null hypothesis or fail to reject a null hypothesis. A null hypothesis that is rejected may, in reality, be true or false. Additionally, a null hypothesis that fails to be rejected may, in reality, be true or false. The outcome that a researcher desires is to reject a false null hypothesis or to fail to reject a true null hypothesis. However, there always is the possibility of rejecting a true hypothesis or failing to reject a false hypothesis.

Rejecting a null hypothesis that is true is called a Type I error and failing to reject a false null hypothesis is called a Type II error. The probability of committing a Type I error is termed α and the probability of committing a Type II error is termed β. As the value of α increases, the probability of committing a Type I error increases. As the value of β increases, the probability of committing a Type II error increases. While one would like to decrease the probability of committing of both types of errors, the reduction of α results in the increase of β and vice versa. The best way to reduce the probability of decreasing both types of error is to increase sample size.

The probability of committing a Type I error, α, is called the level of significance. Before data is collected one must specify a level of significance, or the probability of committing a Type I error (rejecting a true null hypothesis). There is an inverse relationship between a researcher's desire to avoid making a Type I error and the selected value of α; if not making the error is particularly important, a low probability of making the error is sought. The greater the desire is to not reject a true null hypothesis, the lower the selected value of α. In theory, the value of α can be any value between 0 and 1. However, the most common values used in social science research are .05, .01, and .001, which respectively correspond to the levels of 95 percent, 99 percent, and 99.9 percent likelihood that a Type I error is not being made. The tradeoff for choosing a higher level of certainty (significance) is that it will take much stronger statistical evidence to ever reject the null hypothesis.

DETERMINING THE DECISION RULE.

Before data are collected and it is necessary to determine under what circumstances the null hypothesis will be rejected or fail to be rejected. The decision rule can be stated in terms of the computed test statistic, or in probabilistic terms. The same decision will be reached regardless of which method is chosen.

COLLECTING THE DATA AND PERFORMING THE CALCULATIONS.

The method of data collection is determined early in the research process. Once a research question is determined, one must make decisions regarding what type of data is needed and how the data will be collected. This decision establishes the bases for how the data will be analyzed. One should use only approved research methods for collecting and analyzing data.

DECIDING WHETHER TO REJECT THE NULL HYPOTHESIS.

This step involves the application of the decision rule. The decision rule allows one to reject or fail to reject the null hypothesis. If one rejects the null hypothesis, the alternative hypothesis can be accepted. However, as discussed earlier, if one fails to reject the null he or she can only suggest that the null may be true.

EXAMPLE.

XYZ Corporation is a company that is focused on a stable workforce that has very little turnover. XYZ has been in business for 50 years and has more than 10,000 employees. The company has always promoted the idea that its employees stay with them for a very long time, and it has used the following line in its recruitment brochures: "The average tenure of our employees is 20 years." Since XYZ isn't quite sure if that statement is still true, a random sample of 100 employees is taken and the average tenure turns out to be 19 years with a standard deviation of 2 years. Can XYZ continue to make its claim, or does it need to make a change?

1. State the hypotheses.
   H_0 = 20 years
   $H_1 \neq$ 20 years
2. Determine the test statistic. Since we are testing a population mean that is normally distributed, the appropriate test statistic is:
3. Specify the significance level. Since the firm would like to keep its present message to new recruits, it selects a fairly weak significance level (α = .05). Since this is a two-tailed test, half of the alpha will be assigned to each tail of the distribution. In this situation the critical value of Z = +1.96 and −1.96.
4. State the decision rule. If the computed value of Z is greater than or equal to +1.96 or less than or equal to −1.96, the null hypothesis is rejected.
5. Calculations.
6. Reject or fail to reject the null. Since 2.5 is greater than 1.96, the null is rejected. The mean tenure is not 20 years, therefore XYZ needs to change its statement.

SEE ALSO: Research Methods and Processes ; Statistics

Donna T. Mayo

Revised by Marcia Simmering

FURTHER READING:

Anderson, David R., Dennis J. Sweeney, and Thomas A. Williams. *Statistics for Business and Economics.* 9th ed. Mason, OH: South-Western College Publishing, 2004.

Kerlinger, Fred N., and Howard B. Lee. *Foundations of Behavioral Research.* 4th ed. Fort Worth, TX: Harcourt College Publishers, 2000.

Pedhazur, Elazar J., and Liora Pedhazur Schmelkin. *Measurement, Design, and Analysis: An Integrated Approach.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1991.

Schwab, Donald P. *Research Methods for Organizational Studies.* Mahwah, NJ: Lawrence Erlbaum Associates, 1999.

Also read article about Hypothesis Testing from Wikipedia

User Contributions:

Comment about this article, ask questions, or add new information about this topic:

Name:

E-mail:
Show my email publicly

Type the code shown:
[DMCU]

Public Comment: (50-4000 characters)

📨 Send comment

💬 Hypothesis Testing forum