Lecture Notes

# Chapter 8: Testing Hypotheses

## Learning Objectives

8.1 Describe the assumptions of statistical hypothesis testing.

8.2 Define and apply the components in hypothesis testing.

8.3 Explain what it means to reject or fail to reject a null hypothesis.

8.4 Calculate and interpret a test for two sample cases with means or proportions.

8.5 Determine the significance of *t*-test and *Z*-test statistics.

## Chapter Outline

1. Assumptions of Statistical Hypothesis Testing
   1. **Statistical hypothesis testing** (a procedure that allows us to evaluate hypotheses about population parameters based on sample statistics) requires several assumptions.
   2. These assumptions include considerations of the level of measurement of the variable, the method of sampling, the shape of the population distribution, and the sample size.
   3. All assume random sampling.
   4. Tests of hypotheses about means also assume interval-ratio level of measurement and require that the population under consideration be normally distributed or that the sample size be larger than 50*.*
2. Stating the Research and Null Hypotheses
   1. The Research Hypothesis ()
      1. The substantive hypothesis is called the **research hypothesis** (a statement reflecting the substantive hypothesis; it is always expressed in terms of population parameters, but its specific form varies from test to test)and is symbolized as .
      2. **One-tailed test** is a type of hypothesis test that involves a directional research hypothesis.
      3. When a one-tailed test specifies that the population mean is greater than some specified value, it is called a **right-tailed test** because we will evaluate the outcome at the right tail of the sampling distribution.
      4. If the research hypothesis specifies that the population mean is less than some specified value, it is called a **left-tailed test** because the outcome will be evaluated at the left tail of the sampling distribution.
      5. When we have no theoretical reason for specifying a direction in the research hypothesis, we conduct a **two-tailed test***.*
   2. The Null Hypothesis ()
      1. The **null hypothesis**, symbolized as , contradicts the research hypothesis and states that there is no difference between the population mean and some specified value.
      2. In hypothesis testing, rejecting the null hypothesis provides indirect support for the research hypothesis.
3. Probability Values and Alpha
   1. Standard deviation (standard error) of .
   2. *Z* score, modified formula: .
   3. Converting the sample mean to a *Z*-score equivalent is called computing the test statistic. The *Z* value obtained is called the ***Z* statistic (obtained).**
   4. ***p* value** can be defined as the probability associated with the obtained value of *Z*.
      1. It is a measure of how unusual or rare our obtained statistic is compared with what is stated in our null hypothesis.
   5. Researchers usually define in advance what a sufficiently improbable *Z* value is by specifying a cutoff point below which *p* must fall to reject the null hypothesis.
      1. This cutoff point, called **alpha**and denoted by the Greek letter α, is customarily set at the *.*05, .01, or .001 level.
      2. The difference between *p* and alpha is that *p* is the actual probability associated with the obtained value of *Z*, whereas alpha is the level of probability determined in advance at which the null hypothesis is rejected.
      3. The null hypothesis is rejected when, *p* ≤ α.
4. The Five Steps in Hypothesis Testing: A Summary
   1. Making Assumptions
      1. **Statistical hypothesis testing** involves making several assumptions regarding the level of measurement of the variable, the method of sampling, the shape of the population distribution, and the sample size.
   2. Stating the Research and Null Hypotheses and Selecting Alpha
      1. The null hypothesis, symbolized as contradicts the research hypothesis in a statement of no difference between the population mean and our hypothesized value.
   3. Selecting the Sampling Distribution and Specifying the Test Statistic
      1. The normal distribution and the *Z* statistic are used to test the null hypothesis.
   4. Computing the Test Statistic
   5. Making a Decision and Interpreting the Results
5. Errors in Hypothesis Testing
   1. If the null hypothesis is true and we reject it nonetheless, we are making an incorrect decision. This type of error is called a **Type I error**.
   2. If the null hypothesis is false but we fail to reject it, this incorrect decision is a **Type II error***.*
   3. The *t* Statistic and Estimating the Standard Error
      1. ***t* statistic (obtained): ***.*
      2. The obtained *t* represents the number of standard deviation units (or standard error units) that the sample mean is from the hypothesized value of μ, assuming that the null hypothesis is true.
   4. The *t* Distribution and Degrees of Freedom
      1. The ***t* distribution** is actually a family of curves, each determined by its degrees of freedom.
      2. The **degrees of freedom (*df*)** represent the number of scores that are free to vary in calculating each statistic.
         1. In the case of a single-sample mean, the *df* is calculated as follows: *df* = *N* – 1.
   5. Comparing the *t* and *Z* Statistics
      1. The only apparent difference is in the denominator: The denominator of *Z* is the standard error based on the population standard deviation *σ*.
      2. Because it is estimated from sample data, the denominator of the *t* statistic is subject to sampling error.
6. Hypothesis Testing With One Sample and Population Variance Unknown
   1. For example, let’s apply the five-step model to test the hypothesis that the average earnings of White women differed from the average earnings of all women working full time in the United States in 2014.
      1. Making Assumptions: Our assumptions are as follows:
         1. A random sample is selected.
         2. Because *N* > 50, the assumption of normal population is not required.
         3. The level of measurement of the variable income is interval ratio.
      2. Stating the Research and the Null Hypotheses and Selecting Alpha:
7. The research hypothesis is *H*1:μ > $41,977 and the null hypothesis is *H*0:μ > $41,977. We will set alpha at .05, meaning that we will reject the null hypothesis if the probability of our obtained statistic is less than or equal to .05.
   * 1. Selecting the Sampling Distribution and Specifying the Test Statistic:
8. We use the *t* distribution and the *t* statistic to test the null hypothesis.
   * 1. Computing the Test Statistic:
9. We first calculate the *df* associated with our test: *df* = (*N* –1) = (280 – 1) =279..
10. To evaluate the probability of obtaining a sample mean of $45,785, assuming the average earnings of White women were equal to the national average of $41,977, we need to calculate the obtained *t* statistic, 
    * 1. Making a Decision and Interpreting the Results:
11. Given our research hypothesis, we will conduct a two-tailed test. We can see that 279 *df* is not listed, so we will have to use the last row, *df* = ∞, to assess the significance of our obtained *t* statistic.
12. Our obtained *t* statistic of 2.49 is not listed in the last row. It is greater than 2.326 (*t* critical for .01 one-tailed test) but less than 2.576 (*t* critical for .005 one-tailed test). The probability of 2.49 can be estimated as *p* < .05, leading to the conclusion that we reject the null hypothesis.
13. Hypothesis Testing with Two Sample Means
    1. Hypothesis testing with two samples follows the same structure as for one-sample tests.
    2. The Assumption of Independent Samples
       1. One important difference between one- and two-sample hypothesis testing involves sampling procedures.
       2. With a two-sample case, we assume that the samples are independent of each other.
    3. Stating the Research and Null Hypotheses
       1. The second difference between one- and two-sample tests is in the form taken by the research and the null hypotheses.
       2. In one-sample tests, both the null and the research hypotheses are statements about a single population parameter, μ; with two-sample tests, we compare two population parameters.
14. The Sampling Distribution of the Difference Between Means
    1. The **sampling distribution of the difference between two sample means** is a theoretical probability distribution that would be obtained by calculating all the possible mean differences by drawing all possible independent random samples of size *N*1 and *N*2 from two populations.
    2. The central limit theorem assumes that our samples are independently drawn from normal populations.
    3. Estimating the Standard Error
       1. The only data we have are based on sample data, and we do not know the true value of the population variances,  and .
       2. Thus, we need to estimate the standard error from the sample variances,  and .
    4. Calculating the Estimated Standard Error
       1. When we can assume that the two population variances are equal, we combine information from the two sample variances to calculate the estimated standard error:
    5. The t Statistic
       1. As with single sample means, we use the *t* distribution and the *t* statistic whenever we estimate the standard error for a difference between means test.
       2. The formula for computing the *t* statistic for a difference between means test is t = 

where  is the estimated standard error.

* 1. Calculating the Degrees of Freedom for a Difference Between Means Test
     1. To use the t distribution for testing the difference between two sample means, we need to calculate the degrees of freedom.
     2. When calculating the *t* statistic for the two-sample test, we lose 2 *df*, one for every population variance we estimate.
     3. When population variances are assumed to be equal or if the size of both samples is greater than 50, the *df* is calculated as follows: *df* = (*N*1 + *N*2) –2.

1. The Five Steps in Hypothesis Testing About Difference Between Means: A Summary
   1. Statistical hypothesis testing involving two sample means can be organized into five steps:
      1. Making assumptions
      2. Stating the research and null hypotheses and selecting alpha.
      3. Selecting the sampling distribution and specifying the test statistic:The *t* distribution and the *t* statistic are used to test the significance of the difference between the two sample means.
      4. Computing the test statistic.
      5. Making a decision and interpreting the results.
2. Statistics in Practice: Vape Use Among Teens
   1. Data collected from the Monitoring the Future (MTF) surveys have revealed decreases or stability in drug use among youths, particularly for cigarettes, alcohol, marijuana, cocaine, and methamphetamine, although vaping (use of e-cigarettes) has been on the rise.
3. Hypothesis Testing With Two Sample Proportions
   1. Numerous variables in the social sciences are measured at a nominal or an ordinal level.
   2. Hypothesis testing with two sample proportions follows the same structure as the statistical tests presented earlier.
   3. Apply five-step model.
4. Reading the Research Literature: Reporting the Results of Hypothesis Testing
   1. It is not uncommon for a single research article to include the results of multiple statistical tests.
      1. Results have to be presented succinctly and in summary form.
      2. An author’s findings are usually presented in a summary table that may include the sample statistics (e.g., the sample means), the obtained test statistics (*t* or *Z*), the *p* level, and an indication of whether or not the results are statistically significant.
   2. Robert Emmet Jones and Shirley A. Rainey studied how people of color and the poor are more likely than Whites and more affluent groups to live in areas with poor environmental quality and protection, exposing them to greater health risks.
      1. They studied residents from the Red River community in Tennessee from 2001 to 2003.