Test of Hypothesis

- **Null Hypothesis:** The statement being stated in a test of significance is called the *null hypothesis*. The test of significance is designed to assess the strength of the evidence against the null hypothesis. Usually the null hypothesis is a statement of "no effect" or "no difference". The null hypothesis is usually denoted by H_0.

- **Alternative Hypothesis:** The statement that we suspect to be true. Or the statement that we wish to conclude. This *alternative hypothesis* is usually denoted by H_a or H_1.

  Common Types:

  1. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$ (One-sided alternative)

  2. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$ (One-sided alternative)

  3. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$ (Two-sided alternative)

- **Test Statistic:** A *test statistic* measures compatibility between the null hypothesis and the observed data.

- **Z Test for a Population Mean (σ is known):** To test the hypothesis $H_0 : \mu = \mu_0$ based on a SRS of size n from a population with unknown mean μ and known standard deviation σ, compute the test statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1). \tag{1}$$

- **Steps to do Hypothesis Testing:**

  1. Formulate the Null hypothesis and Alternative hypothesis.

  2. Specify the level of significance (Commonly used: $\alpha = 0.05$ or 0.01).

  3. Determine the appropriate *test statistic* to use.

  4. Define your rejection rule (Not needed if you decide to use the p-value).

  5. Compute the observed value of the test statistic (Compute the p-value).

  6. Write your conclusion.

- **Examples:**

  1. In a discussion of SAT scores, someone comments: "Because only a minority of high school students take the test, the scores overestimate the ability of typical high school seniors. The mean SAT mathematics score is about 519, but I think that if all seniors took the test, the mean score would be no more than 450." To verify his claim, you gave the test to a SRS of 500 seniors from California. These students had an average score of $\bar{x} = 461$. Is this enough evidence against the claim that the mean for all California seniors is no more than 450? Use a level of significance equal to $\alpha = 0.01$. (Assume that $\sigma = 100$).

2. Do middle-aged male executives have different average blood pressure than the general population? The national Center for Health Statistics reports that the mean systolic blood pressure for males 35 to 44 years of age is 128 and the standard deviation in this population is 15. The medical director of a company looks at the medical records of 72 company executives in this age group and finds that the mean systolic blood pressure is this sample is $\bar{x} = 126.07$. Is this enough evidence that executive blood pressures differ from the national average? Use $\alpha = 0.05$.

- **Practice:**

1. The Federal Trade Commission (FTC) periodically conducts statistical studies designed to test the claims that manufacturers make about their products. For example, the label on a large can of Hilltop Coffee states that the can contains 3 pounds of coffee. To protect the consumers, FTC wants to make sure that the population mean amount of coffee in each can is at least 3 pounds. If a sample of 36 Hilltop coffee cans provides a sample mean of $\bar{x} = 2.92$ pounds and σ is known to be 0.18, is there enough evidence to conclude that the population mean is statistically lower than 3 pounds per can? Use a level of significance equal to $\alpha = 0.05$.

2. Reis, Inc., a New York real state research firm, tracks the cost of apartment rentals in the United States. In mid-2002, the nationwide mean apartment rental rate was \$895 per month (*The Wall Street Journal*, July 8, 2002). Assume that, based on the historical quarterly surveys, a population standard deviation $\sigma = \$225$ is reasonable. In a current study of apartment rental, a sample of 180 apartments nationwide provided a sample mean of \$915 per month. Do the sample data enable Reis to conclude that the population mean apartment rental rate now exceeds the level reported in 2002? $\alpha = 0.01$.

3. The Florida Department of Labor and Employment Security reported the state mean annual wage was \$26,133 (*The Naples Daily News*, Feb. 13, 1999). A hypothesis test of wages by county can be conducted to see whether the mean annual wage for a particular county differs from the state mean. A sample of 550 individuals from Collier County showed a sample mean annual wage of \$25,457. Assuming that $\sigma = \$7600$, is there enough evidence to conclude that mean annual wage of people from this county is different than the state mean? Use $\alpha = 0.05$.

4. The national mean sales price for new one-family homes is \$181,900 (*The New York Times Almanac 2000*). A sample of 40 one-family home sales in the south showed a sample mean of \$166,400. If $\sigma = \$33,500$, is there enough evidence to say that the population mean sales price for new one-family homes in the south is less than the national mean? Use $\alpha = 0.01$.

5. Individuals filing federal income tax returns prior to March 31 received an average refund of \$1056. Consider the population of "last minute" filers who mail their tax return during the last five days of the income tax period (typically April 10 to April 15).

a. A researcher suggests that a reason individuals wait until the last five days is that on average these individuals receive lower refunds than do early filers. Formulate appropriate hypotheses such that the rejection of H_0 will support the researcher's contention. Clearly define the parameter you used.

b. For a sample of 400 individuals who filed a tax return between April 10 to April 15, the sample mean refund was \$910. Based on prior experience a population standard deviation of $\sigma = \$1600$ may be assumed. Calculate the value of the appropriate test statistic.

c. At $\alpha = 0.05$, what is your conclusion?

d. Compute the $p-$value? What is your conclusion based on the value of the $p-$value?

Type I and Type II Errors:

1. *Type I error* is rejecting H_0 when it is true.
   Note: Pr(Type I error)=α.
2. *Type II error* is not rejecting H_0 when it is false.
   Note: Pr(Type II error)=β, and the *Power* of the test is $(1 - \beta)$.

Homework:

1. Due Wednesday, April 2.
   Sec 8.1: (pp. 375-376) # 8, 9, 10, 11, 12, 13.
   Sec 8.2: (pp. 381-383) # 21(a,d,f), 22, 23, 24, 26, 27, 33.

2. Due Thursday, April 3.
   Sec 8.1: (pp. 375-376) # 14, 15, 16.
   Sec 8.3: (pp. 388-390) # 38, 39, 41, 42, 43, 44, 46, 47.