WIIT 7752

Lesson D Guided Notes

Null and Alternative Hypotheses

A study published in the *Journal of the Air and Waste Management Association* reported that the mean amount of particulate matter (PM) produced by cars and light trucks in an urban setting is 35 milligrams of PM per mile of travel. Suppose that a new engine design is proposed that is intended to reduce the amount of PM in the air. There are two possible outcomes that could happen with the new engine design: either the new design will reduce the level of PM on average, or it will not.

These possibilities are called **hypotheses**. One of the hypotheses is called the **null hypothesis** and the other is called the **alternate hypothesis**.

Definition and Notation

The ___________________ about a parameter states that the parameter is equal to a specific value, μ_0. The null hypothesis is denoted H_0.

The ___________________ about a parameter states that the parameter differs from the value specified by the null hypothesis, μ_0. The alternate hypothesis is denoted H_1. There are three possible alternate hypotheses:

1. __________: States that the parameter is less than the value specified by the null hypothesis, for example, H_1: $\mu < \mu_0$
2. __________: States that the parameter is greater than value specified by the null hypothesis, for example, H_1: $\mu > \mu_0$
3. __________: States that the parameter is not equal to the value specified by the null hypothesis, for example, H_1: $\mu \neq \mu_0$

Left-tailed and right-tailed hypotheses are called __________ hypotheses.

Example 1 Boxes of a certain kind of cereal are labeled as containing 20 ounces. An inspector thinks that the mean weight may be less than this. State the appropriate null and alternate hypotheses.

Example 2 Last year, the mean monthly rent for an apartment in a certain city was \$800. A real estate agent believes that the mean rent is higher this year. State the appropriate null and alternate hypotheses.

Example 3 Scores on a standardized test have a mean of 70. Some modifications are made to the test, and an educator believes that the mean may have changed. State the appropriate null and alternate hypotheses.

More on Hypothesis Testing

The idea behind a hypothesis test is similar to a criminal trial. At the beginning of a trial, the defendant is assumed to be innocent. Then the evidence is presented. If the evidence strongly indicates that the defendant is guilty, we abandon the assumption of innocence and conclude the defendant is guilty. In a hypothesis test, the null hypothesis is like the defendant in a criminal trial.

At the start of a hypothesis test, we assume that the null hypothesis is true.

Then we look at the evidence, which comes from data that have been collected.

If the data strongly indicate that the null hypothesis is false, we abandon our assumption that it is true and believe the alternate hypothesis instead. This is referred to as **rejecting** the null hypothesis.

Stating Conclusions

We may either ___________________ the null hypothesis or _________________________ the null hypothesis.

If the null hypothesis is rejected, the conclusion is straightforward: We conclude that the alternate hypothesis, H_1, is true.

If the null hypothesis is not rejected, we are saying that there is not enough evidence to conclude that the alternate hypothesis, H_1, is true. We are **not** saying the null hypothesis is true. What we are saying is that the null hypothesis **might** be true.

Example 4 Boxes of a certain kind of cereal are labeled as containing 20 ounces. An inspector thinks that the mean weight may be less than this, so he performs a test of H_0: $\mu = 20$ versus H_0: $\mu < 20$. He rejects the null hypothesis. State an appropriate conclusion.

Example 5 Boxes of a certain kind of cereal are labeled as containing 20 ounces. An inspector thinks that the mean weight may be less than this, so he performs a test of H_0: $\mu = 20$ versus H_1: $\mu < 20$. He does not reject the null hypothesis. State an appropriate conclusion.

Errors

When a hypothesis test is conducted and a decision is made there is a possibility that it is the wrong decision.

There are two ways in which a wrong decision may occur with hypothesis testing.

1. If H_0 is true, we might mistakenly reject it. A **type I error** occurs when we reject H_0 when it is actually true.

2. If H_0 is false, we might mistakenly decide not to reject it. A **type II error** occurs when we do not reject H_0 when it is actaully false.

Decision	H_0 is True	H_0 is False
Reject H_0		
Fail to Reject H_0		

Example 6 The dean of a business school wants to determine whether the mean starting salary of graduates of her school is greater than \$50,000. She will perform a hypothesis test with the following null and alternate hypotheses:

$$H_0\colon \mu = \$50{,}000 \qquad H_1\colon \mu > \$50{,}000$$

Suppose that the true mean is $\mu = \$50{,}000$, and the dean rejects H_0. Is this a Type I error, Type II error, or a correct decision?

Example 7 The dean of a business school wants to determine whether the mean starting salary of graduates of her school is greater than $50,000. She will perform a hypothesis test with the following null and alternate hypotheses:

$$H_0\text{: } \mu = \$50{,}000 \qquad H_1\text{: } \mu > \$50{,}000$$

Suppose that the true mean is $\mu = \$55{,}000$, and the dean rejects H_0. Is this a Type I error, Type II error, or a correct decision?

Example 8 The dean of a business school wants to determine whether the mean starting salary of graduates of her school is greater than $50,000. She will perform a hypothesis test with the following null and alternate hypotheses:

$$H_0\text{: } \mu = \$50{,}000 \qquad H_1\text{: } \mu > \$50{,}000$$

Suppose that the true mean is $\mu = \$55{,}000$, and the dean does not reject H_0. Is this a Type I error, Type II error, or a correct decision?

T-test

When testing a hypothesis for a population mean, μ, we will use a T-test. This occurs when the population standard deviation, σ, is unknown, a case which is more realistic than the alternative.

In practice, since we will rarely know the population standard deviation, we replace σ with the sample standard deviation s and use the t statistic

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

When the null hypothesis is true, the t statistic has a Student's t distribution with $n - 1$ degrees of freedom. When we perform a test using the t statistic, we call this a ***t*-test**.

P-Values

The P-value is the probability, calculated assuming that the null hypothesis is true, of obtaining a value of the test statistic at least as contradictory to H_0 as the value calculated from the available sample.

This is not the same thing as the probability that the null hypothesis is true!

Select a significance level α. Then,

Assumptions

The **assumptions** for performing a hypothesis test about the mean when σ is not known are:

- We have a simple random sample.
- The sample size is large ($n > 30$), or the population is approximately normal.

When the assumptions are met, a hypothesis test may be performed using either the *P*-value method or critical value method.

T-Test Procedure

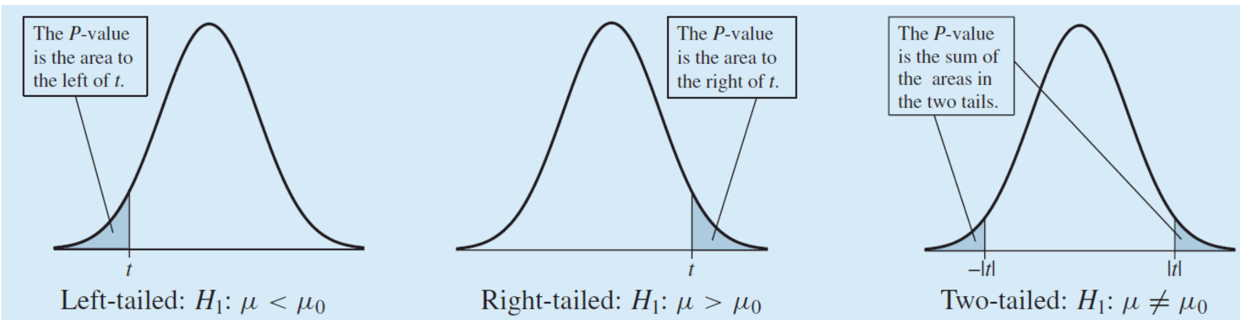
Check to be sure the assumptions are satisfied.

Step 1. State the null and alternate hypothesis.

Step 2. If making a decision, choose a significance level α.

Step 3. Calculate the test statistic.

Step 4. Calculate the P-value.



Step 5. Interpret the *P*-value. If making a decision, reject H_0 if the *P*-value is less than or equal to the significance level α.

Step 6. State a conclusion.

Example 9 In a recent medical study, 76 subjects were placed on a low-fat diet. After 12 months, their sample mean weight loss was $\bar{x}$ = 2.2 kilograms, with a sample standard deviation of s = 6.1 kilograms. Can we conclude that the mean weight loss is greater than 0? Use the α = 0.05 level of significance.

Example 10 A computer software vendor claims that a new version of their operating system will crash less than six times per year on average. A system administrator installs the operating system on a random sample of 41 computers. At the end of a year, the sample mean number of crashes is 7.1, with a standard deviation of 3.6. Can you conclude that the vendor's claim is false? Use the α = 0.05 significance level.

Example 11 Let *X* be a random variable that represents hemoglobin count (HC) in grams per 100 milliliters of whole blood. Then *x* has a distribution that is approximately normal with a population mean of 14 for healthy adult women (based on information from *Diagnostic Tests with Nursing Implications*, Springhouse Corporation). Suppose a female has taken 12 lab blood tests during the past year. The HC data sent to the patient's doctor were:

19	23	15	21	18	16
14	20	19	16	18	21

Is there evidence that this patient's mean hemoglobin count is greater than that of other healthy adult women? Use a 0.01 significance level.

Example 12 According to *The Wall Street Journal*, a professional employee working in a large company receives a mean of μ= 31.8 calls per day. Most of the calls are from other employees in the company. Because of the large number of calls, employees find themselves distracted and are unable to concentrate when they return to their tasks. In an effort to reduce distraction caused by such interruptions, one company has established a “priority list” that all employees were to use before making phone calls. One month after the new priorities were put into use, a random sample of 63 employees showed they were receiving a sample mean of $\bar{x}$ = 29.5 calls per day and the sample deviation is 10.7 calls per day. Can we conclude that the “priority list” reduces mean number of calls per day received per employee at the α= 0.01 level of significance?

Statistical Significance

When a result has a small *P*-value, we say that it is “statistically significant.” It is therefore tempting to think that statistically significant results must always be important. Sometimes statistically significant results do not have any practical importance or practical significance.

For example, a new study program may raise students’ scores by two points on a 100 point scale. This improvement may have statistical significance, but is the improvement important enough to offset the cost of training teachers and the time investment on behalf of the students.

Hypothesis Tests for Proportions

In the 2009–2010 GenX2Z American College Student Survey, 90% of female college students rated the social network site Facebook as “cool.” The other 10% rated it as “lame.” Assume that the survey was based on a sample of 500 students. A marketing executive at Facebook wants to advertise the site with the slogan “More than 85% of female college students think Facebook is cool.” Before launching the ad campaign, he wants to be confident that the slogan is true. Can he conclude that the proportion of female college students who think Facebook is cool is greater than 0.85?

This is an example of a problem that calls for a **hypothesis test about a population proportion**.

Notation

We use the following notation:

- *p* is the **population proportion** of individuals who are in a specified category.
- *x* is the **number of individuals** in the sample who are in the specified category.
- *n* is the **sample size**.
- $\hat{p}$ is the **sample proportion** of individuals who are in the specified category.
- p_0 is the value of *p* under the null hypothesis.

The sample proportion is given by $\hat{p} = \frac{x}{n}$.

Assumptions

We can perform a hypothesis for the population proportion whenever the sample proportion $\hat{p}$ is approximately normally distributed. This will occur when the following assumptions are met.

1. We have a simple random sample.

2. The population is at least 20 times as large as the sample.

3. The individuals in the population are divided into two categories.

4. The values np_0 and $n(1 - p_0)$ are both at least 10.

Once the assumptions are met, either the *P*-value method or critical value method may be used to perform the hypothesis test.

P-value Method

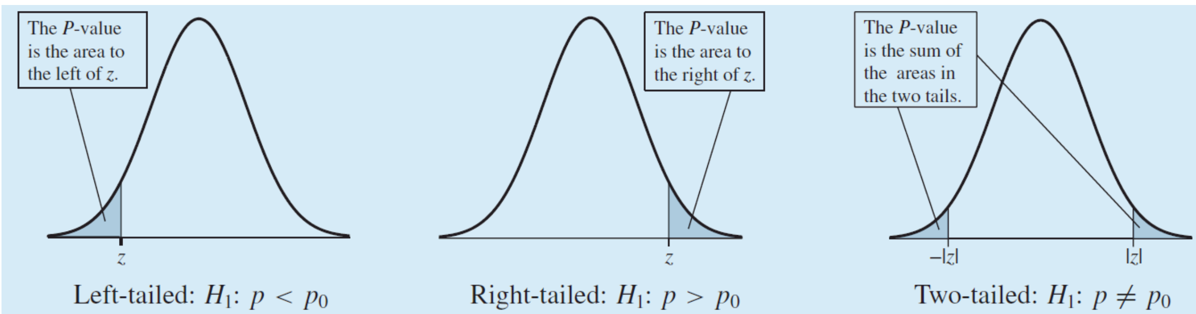
Check to be sure the assumptions are satisfied.

Step 1. State the null and alternate hypothesis.

Step 2. If making a decision, choose a significance level α.

Step 3. Calculate the test statistic.

Step 4. Calculate the P-value.



Step 5. Interpret the *P*-value. If making a decision, reject H_0 if the *P*-value is less than or equal to the significance level α.

Step 6. State a conclusion.

Example 13 In the 2009 – 2010 GenX2Z American College Student Survey, 90% of female college students rated the social network site Facebook as "cool." Assume that the survey was based on a random sample of 500 students. A marketing executive at Facebook wants to advertise the site with the slogan "More than 85% of female college students think Facebook is cool." Can you conclude that the proportion of female college students who think Facebook is cool is greater than 0.85? Use the α = 0.05 level of significance.

Example 14 A nationwide survey of working adults indicates that only 50% of them are satisfied with their jobs. The president of a large company believes that more than 50% of employees at his company are satisfied with their jobs. To test his belief, he surveys a random sample of 100 employees, and 54 of them report that they are satisfied with their jobs. Can he conclude that more than 50% of employees at the company are satisfied with their jobs? Use the α = 0.05 level of significance.

Example 15 The CEO of a large electric utility claims that 80 percent of his 1,000,000 customers are very satisfied with the service they receive. To test this claim, the local newspaper surveyed 100 customers, using simple random sampling. Among the sampled customers, 73 percent say they are very satisfied. Based on these findings, can we reject the CEO's hypothesis that 80% of the customers are very satisfied? Use a 0.05 level of significance.