Section 10.1

Steps in Hypothesis Testing

1. A claim is made. (More than 64% of ECC students vote, in our case.)
2. Evidence is collected to test the claim. (We found that 326 of 460 voted.)
3. The data are analyzed to assess the plausibility of the claim. (We determined that the proportion is most likely higher than 64% for ECC students.)

Null and Alternative Hypotheses

In statistics, we call these claims *hypotheses*. We have two types of hypotheses, a **null hypothesis** and an **alternative hypothesis**.

The **null hypothesis**, denoted H_0 ("H-naught"), is a statement to be tested. It is usually the status quo, and is assumed true until evidence is found otherwise.

The **alternative hypothesis**, denoted H_1 ("H-one"), is a claim to be tested. We will try to find evidence to support the alternative hypothesis.

There are three general ways in this chapter that we'll set up the null and alternative hypothesis.

1. **two tailed**
   H_0: parameter = some value
   H_1: parameter ≠ the value
2. **left-tailed**
   H_0: parameter = some value
   H_1: parameter < the value
3. **right-tailed**
   H_0: parameter = some value
   H_1: parameter > the value

Let's look an example of each.

Example 1

In the introduction of this section, we were considering the proportion of ECC students who voted in the 2008 presidential election. We assumed that it was the same as the national proportion in 2004 (64%), and tried to find evidence that it was higher than that. In that case, our null and alternative hypotheses would be:

H_0: $p = 0.64$
H_1: $p > 0.64$

Example 2

According to the Elgin Community College website, the average age of ECC students is 28.2 years. We might claim that the average is less for online Mth120 students. In that case, our null and alternative hypotheses would be:

H_0: $\mu = 28$
H_1: $\mu < 28$

Example 3

It's fairly standard knowledge that IQ tests are designed to be normally distributed, with an average of 100. We wonder whether the IQ of ECC students is different from this average. Our hypotheses would then be:

H_0: $\mu = 100$
H_1: $\mu \neq 100$

There are four possible outcomes from a hypothesis test when we compare our decision with what is true in reality - which we will never know!

- We could decide to not reject the null hypothesis when in reality the null hypothesis was true. This would be a correct decision.
- We could reject the null hypothesis when in reality the alternative hypothesis is true. This would also be a correct decision.
- We could reject the null hypothesis when it really is true. We call this error a **Type I error**.
- We could decide to not reject the null hypothesis, when in reality we should have, because the alternative was true. We call this error a **Type II error**.

		reality	
		H_0 true	H_1 true
decision	do not reject H_0	correct decision	Type II error
	reject H_0	Type I error	correct decision

To help illustrate the idea, let's look at an example.

Example 4

Let's consider a pregnancy test. The tests work by looking for the presence of the hormone human chorionic gonadotropin (hCG), which is secreted by the placenta after the fertilized egg implants in a woman's uterus.

If we consider this in the language of a hypothesis test, the null hypothesis here is that the woman is not pregnant - this is what we assume is true until proven otherwise.

The corresponding chart for this test would look something like this:

		reality	
		not pregnant	pregnant
pregnancy test	not pregnant	correct decision	Type II error
	pregnant	Type I error	correct decision

In this case, we would call a Type I error a "false positive" - the test was positive for pregnancy, when in reality the mother was not pregnant.

The Type II error in this context would be a "false negative" - the test did not reveal the pregnancy, when the woman really was pregnant.

When a test claims that it is "99% Accurate at Detecting Pregnancies", it is referring to Type II errors. The tests claim to detect 99% of pregnancies, so it will make a Type II error (not detecting the pregnancy) only 1% of the time.

Note: The "99% Accurate" claim is not entirely correct. Many tests do not have this accuracy until a few days after a missed period.