Key Ideas
Hypothesis (Null and Alternative), Hypothesis Test, Test Statistic, P-value
Type I Error, Type II Error, Significance Level, Power

<u>**Section 8-1: Overview**</u>
Confidence Intervals (Chapter 7) are great for estimating the value of a parameter, but in some situations estimation is not the goal. For example, suppose that researchers somehow know that the average number of fast food meals eaten per week by families in the 1990s was 3.5. Furthermore, they are interested in seeing if families today eat *more* fast food than in the 1990s. In this case, researchers don't need to estimate how much fast food families eat today. All they need to know is whether that average is larger than 3.5. For the purposes of this example, let μ denote the average number of fast food meals eaten per week by families today. The researchers' question could then be formulated as follows:

Is the average number of meals eaten today the same as in the 1990s ($\mu = 3.5$) or is it more ($\mu > 3.5$)?

There are two different possibilities, and the correct answer is impossible to know, since μ is unknown. However, the researchers *could* observe the sample mean and determine which possibility is the most *likely*. This process is called **hypothesis testing**.

<u>Definition</u>
Hypothesis – A claim or statement about a property of the population (e.g. "$\mu = 3.5$" from above)
Hypothesis Test – A standard procedure for testing a claim about a property of the population.

To guide this discussion of hypothesis testing, remember the "Rare Event Rule":
If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.

<u>Example</u>
Your friend claims to have a fair coin, and you want to test this claim. If $p = P(\text{Head})$, then we are testing the claim that $p = 0.5$. Suppose you flip the coin 100 times.

a. If you get 54 heads and 46 tails, would you conclude $p = 0.5$?

    Here, you probably would. We should expect around 50 heads and 50 tails, and this is not far off from the mark. This is the kind of outcome one would expect from a fair coin.

b. If you get 89 heads and 11 tails, would you conclude $p = 0.5$?

    You probably would not say $p = 0.5$ in this case. Why does the coin seem unfair? If it really was fair, then this kind of outcome (while possible) would have a very low chance of occurring. It is much more likely that the coin is actually unfair. This is the rare event rule mentioned above. Here, the "given assumption" is that the coin is fair. Since the probability of 89 heads and 11 tails on a fair coin is "exceptionally small", we conclude that the given assumption (fair coin) is probably not correct.

c. What would you say if you got 40 heads and 60 tails? 65 heads and 35 tails?

    This is getting into a kind of gray area where it is hard to decide whether a fair coin would have a good chance of getting these outcomes. When observations are on the borderline like this, it is harder to choose between the two possible scenarios. Making these decisions is where hypothesis testing comes into play.

Section 8-2: Basics of Hypothesis Testing

Definitions

Null Hypothesis – The null hypothesis, denoted H_0, is a statement that the value of a population parameter (e.g. population mean, proportion, or standard deviation) is *equal* to a particular value. For the purposes of the test, we assume that the null hypothesis is true, and then decide whether there is enough evidence to reject that assumption. For example, here are some null hypotheses:

$$H_0: \mu = 3.5 \qquad\qquad H_0: p = 0.5 \qquad\qquad H_0: \sigma = 12$$
$$\text{(first example)} \qquad\qquad \text{(second example)}$$

Alternative Hypothesis – The alternative hypothesis, denoted H_1, is a statement that the parameter has some value that is different than the one in the null hypothesis. These statements are all inequalities and come in 3 forms: $>$, $<$, and $\neq$. Here are some examples of alternative hypotheses:

$$H_0: \mu > 3.5 \qquad\qquad H_0: p \neq 0.5 \qquad\qquad H_0: \sigma < 12$$
$$\text{(first example)} \qquad\qquad \text{(second example)}$$

Test Statistic – The test statistic is a value that is used to decide whether to reject the null hypothesis. It is a quantity based on the sample data and has a known distribution when the null hypothesis is true. This process will be discussed in more detail soon.

The General Idea of a Hypothesis Test

To run a hypothesis test, there are a few general steps, which will be elaborated on later.
1. Based on the question you want to answer, formulate the test as a choice between two hypotheses (the null and alternative).
2. Find a test statistic whose distribution is known when the null hypothesis H_0 is true.
3. Figure out if the value of the test statistic computed for your sample is an "unlikely" value from that distribution.
4. If it is unlikely enough, reject H_0 and conclude that H_1 is more likely. If it is not unlikely, then conclude that there is not enough evidence to reject H_0. Note that this is not saying H_0 is true – rather it is just saying that there isn't enough evidence to conclude that it is *false*.

Example: In the fast food example from the start of this chapter, researchers want to test whether the average number of fast food meals eaten per week by modern families is larger than it was in the 1990s, when it was 3.5. Suppose that they know the population standard deviation is $\sigma = 0.6$, the sample size is 100, and the sample mean was $\bar{x} = 3.7$.

1. There are two possibilities: the new mean is still 3.5 or the new mean is larger than 3.5. Thus, the hypotheses are:

$$H_0: \mu = 3.5$$
$$H_1: \mu > 3.5$$

2. If H_0 were true, is there some expression (a test statistic) whose distribution is known? Notice that here we also know that $\sigma = 0.6$, $n = 100$, and $\bar{x} = 3.7$.

   Recall:
   $$Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \text{ has a known distribution.}$$
   This distribution is the *standard normal distribution*.

3. What is the value of Z for this sample?

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{3.7 - 3.5}{0.6 / \sqrt{100}} = 3.333$$

4. Is this value of Z unlikely for a standard normal distribution? Yes, it is. Remember, the chance of Z-scores being above 3 is very low. In this case, it is much more likely that the new mean is larger than 3.5. Therefore, we would reject H_0.

Determining Which Values are "Unlikely"

In the previous example, we know that 3.333 was an unlikely value for the standard normal distribution. However, the definition of "unlikely" is subjective, and it is important to find a way to make a clear cut-off for "too unlikely" versus "likely enough". To do this, we introduce some new terminology:

Definitions

Critical Region – The set of all values of the test statistic that lead to rejection of the null hypothesis (i.e. the area where values would be considered too extreme).

Significance Level – The probability that the test statistic will fall in the critical region when H_0 is actually true. In other words, this is the chance that we would mistakenly reject H_0 even though it is actually true. This probability is denoted α, and it is typically a small value, like $\alpha = 0.05$ or $\alpha = 0.01$.

Critical Value – The cut-off value between the critical region and the range of acceptable values.

Example: Fast food example from the beginning of the chapter.
We saw that Z had a standard normal distribution when H_0 was true. Since the alternative hypothesis is that $\mu > 3.5$, the "unlikely" values of Z will be those that are too *large*. Therefore, the critical region, critical value, and significance level can be found on the normal distribution below:
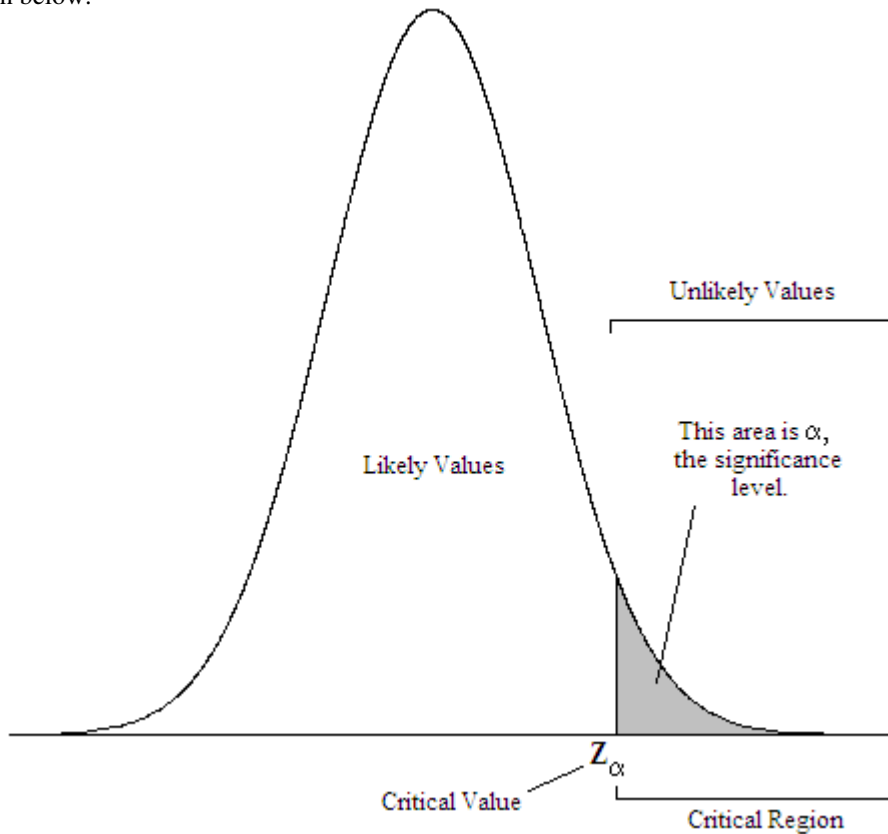
Decision-Making in Hypothesis Testing

There are 3 general approaches to hypothesis testing, and the differences between the methods are in the decision-making process at the end of the test.

These approaches are:

1. The *Traditional Method*
2. The *P-Value Method*
3. The *Confidence Interval Method*

The traditional method uses critical regions and critical values (as mentioned above) to make decisions. Here are the steps to using the traditional method:

1. Formulate the question into a null hypothesis H_0 and an alternative hypothesis H_1.
2. Identify a test statistic which has a known distribution when H_0 is true.
3. Define a critical region on the distribution and a critical value that marks the edge of the critical region as follows:

If H_1 has a > sign, the critical region is the upper tail (values are "too large")
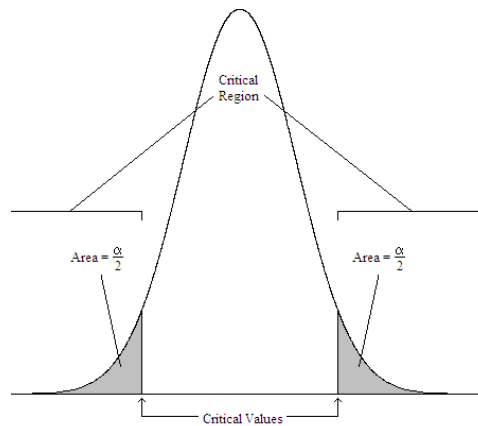
Critical Region

Area = α
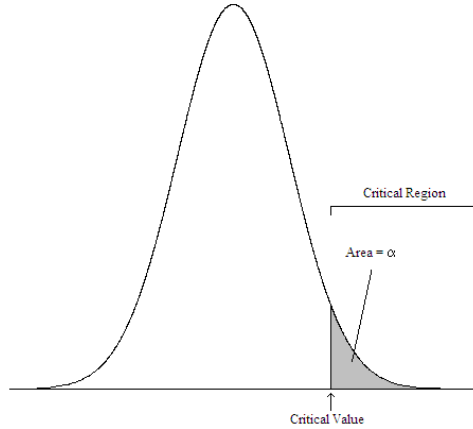
Critical Value

If H_1 has a < sign, the critical region is the lower tail (values are "too small")

Critical Region

Area = α
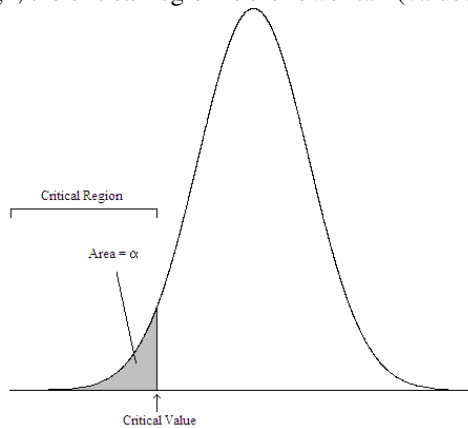
Critical Value

If H_1 has a ≠ sign, the critical region is *both* tails (values are either "too small" or "too large")
Note: the overall area is still α.

Critical
Region

Area = $\frac{\alpha}{2}$ Area = $\frac{\alpha}{2}$

Critical Values

4. Compute the test statistic from the sample data. If the test statistic falls in the critical region, reject H_0. Otherwise, there is not enough evidence to reject H_0.

The P-Value Method
One drawback to the traditional method is that a new critical value must be computed for each value of the significance level α. Thus, while one person thinks an area of α = 0.05 is small enough, someone else might think α must be smaller, like 0.01. The second person would have to re-compute the critical value for α = 0.01 to run the test at their significance level. To solve this problem, we can use the p-value method.

1. Formulate the question into a null hypothesis H_0 and an alternative hypothesis H_1.
2. Identify a test statistic which has a known distribution when H_0 is true.
3. Compute the test statistic from the sample data.
4. Find the **p-value** for the test statistic, which is the probability of having a value *at least as extreme* as the value of the statistic:

If H_1 has a > sign, the p-value is the area above the test statistic ("at least as extreme" means "larger")

P-Value = This Area
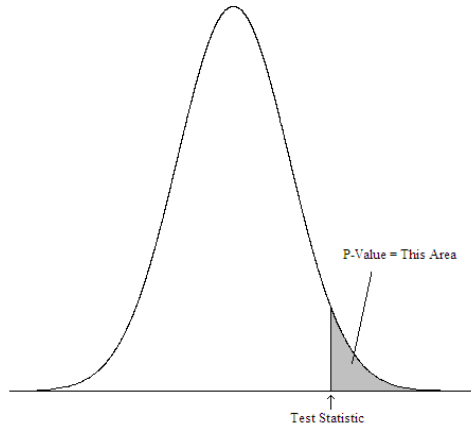
Test Statistic

If H_1 has a < sign, the p-value is the area below the test statistic ("at least as extreme" means "smaller")

P-Value = This Area
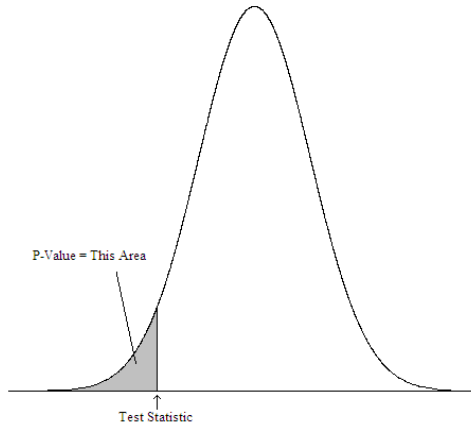
Test Statistic

If H_1 has a ≠ sign, the p-value is the area above and below the test statistic and its opposite. So if the statistic is positive, it is the area above the test statistic and below its negative. If the statistic is negative, it is the area below the statistic and above its absolute value.
("at least as extreme" means "larger in *absolute value*")

P-Value = This Area

One of these is the test
statistic... the other is
its opposite.
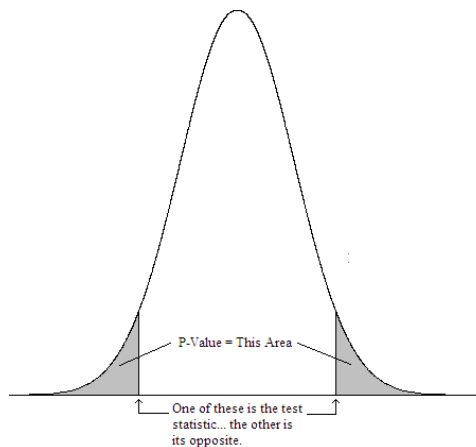
5. If the p-value is less than α, reject H_0. Otherwise, there is not enough evidence to reject H_0.

<u>The Confidence Interval Method</u>
This method is not commonly used, but it still works.

1. Compute a confidence interval for the population parameter as in Chapter 7.
2. Since the confidence interval contains all likely values of the parameter, reject H_0 if the quantity in the null hypothesis does not fall in the confidence interval. If it does fall in the interval, there is not enough evidence to reject H_0.

<u>Types of Error</u>
There are two different ways that the hypothesis test could give the wrong conclusion.
1. H_0 is in reality true, but the test rejects H_0.
   This is called **Type I Error**, and is denoted α (it is also called the **significance level**).
2. H_0 is in reality false, but the test does not reject H_0.
   This is called **Type II Error**, and is denoted β.

		Reality	
		H_0 True	H_0 False
Test Decision	Do not reject H_0	Correct Answer	Type II Error (β)
	Reject H_0	Type I Error (α)	Correct Answer

In testing situations, α is chosen (most scientists select $\alpha = 0.05$ or $\alpha = 0.01$).
β cannot be selected, but there are techniques to reduce it:
- For fixed α, increasing the sample size n will reduce β.
- For fixed sample size n, increasing α will reduce β.
- To decrease both α and β, increase the sample size n.

<u>Remark</u>: The quantity $1 - \beta$ is the probability of rejecting H_0 when it is actually false (see table above). It has a special name, which is the *power* of a test. Power is something we will not worry about in this class, but to increase it, one can use the techniques above (reducing β is the same as increasing power). See p.400 – 403 for more information.

Section 8-3: Testing a Claim About a Proportion

In order to test a claim about a proportion, the hypothesis test requires that a few conditions be met. These conditions satisfy some of the theoretical assumptions made in using this test. In particular, testing claims about a proportion requires that the Central Limit Theorem be used in order to *approximate* a Binomial distribution with a Normal distribution (we did not cover that section).

Conditions
1. The sample must be a simple random sample.
2. The conditions for a Binomial distribution must be met (i.e. n independent trials with 2 outcomes, P(success) = p and is the same for each trial)
3. $np \geq 5$ and $nq \geq 5$

(Condition 2 makes sure it is binomial, and Condition 3 is for the approximation with a normal distribution)

Notation
n = sample size
p = the true population proportion of successes
$q = 1 - p$
$\hat{p}$ = the sample proportion of successes

Using the Hypothesis Test for a Proportion
To run the hypothesis test, we use the following test statistic (notice that the components of it are similar to the ones used in the confidence intervals in the previous chapter):

Test Statistic: $Z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$ Critical Values and P-Values come from the Standard Normal Distribution

(Here, p is the value used in the null hypothesis)

To test a hypothesis about the population proportion, follow these steps:
1. Write down the null and alternative hypotheses, as given in the statement of the problem.
2. Identify the values of importance: p, q, $\hat{p}$, n, α
3. Calculate the test statistic Z.
4. Note which sign is used in the alternative hypothesis H_1.
5. Using either the traditional or p-value method, determine whether the test statistic is "unlikely".
   With a > sign, you look at the area at the high end of the distribution.
   With a < sign, you look at the area at the low end of the distribution.
   With a ≠ sign, you divide the area between the high and low ends of the distribution.
6. If it is unlikely, reject H_0. Otherwise, there is not enough evidence to reject H_0.

Example
Poll workers want to determine if the national percentage of Americans who identify themselves as supporting an "Independent" political party is *larger* than 8%. To do this, they sample 64 people and find that 9 of them are Independent. Determine whether the poll workers' suspicions are correct using significance level $\alpha = 0.05$.

Solution
From the information given, we see that:

$n = 64$, $\alpha = 0.05$, $p = 0.08$, $q = 1 - 0.08 = 0.92$, $\hat{p} = \dfrac{9}{64} = 0.140625$

H_0: $p = 0.08$
H_1: $p > 0.08$

Test Statistic: $Z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} = \dfrac{0.140625 - 0.08}{\sqrt{\dfrac{(0.08)(0.92)}{64}}} = \dfrac{0.060625}{0.033912} = 1.788$

Traditional Method

- Since H_1 has a ">" sign, we want to find the critical value that has an area of α above it on the standard normal distribution (i.e. we want the value Z_α).

- From the table, this cut-off value with an area of 0.05 above is 1.645.

- Now we compare the test statistic to 1.645 and find that $Z = 1.788 > 1.645$.

- This means that Z is in the critical region (shaded area above 1.645).

- Therefore, we reject H_0 and conclude that the nationwide percentage of Independents is larger than 8%.
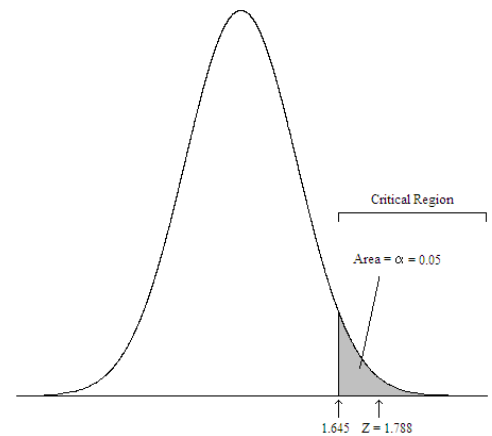
P-Value Method

- Since H_1 has a ">" sign, we want to find area above $Z = 1.788$ for the standard normal distribution.

- From the Z-Table, this area is $1 - 0.9633 = 0.0367$.

- Now we compare this area to α and see that $0.0367 < 0.05$.

- Therefore, we reject H_0 and conclude that the nationwide percentage of Independents is larger than 8%.
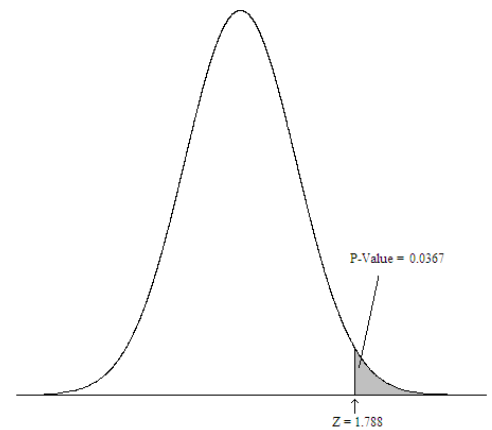
Example
The newspaper claims that 50% of Americans own a pet, but Fred thinks that it is different than that. To test this claim, he takes a simple random sample of 100 Americans and finds that 57 of them own pets. Test his claim with significance level α = 0.05.

Solution
From the information given, we see that:

$n = 100$, $\alpha = 0.05$, $p = 0.50$, $q = 1 - 0.50 = 0.50$, $\hat{p} = \frac{57}{100} = 0.57$

H_0: $p = 0.50$
H_1: $p \neq 0.50$

Test Statistic: $Z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} = \dfrac{0.57 - 0.50}{\sqrt{\dfrac{(0.50)(0.50)}{100}}} = \dfrac{0.07}{0.05} = 1.40$

Traditional Method

- Since H_1 has a "$\neq$" sign, we want to find the critical value that has an area of α/2 above it and α/2 below it on the standard normal distribution (i.e. we want the value $Z_{\alpha/2}$).

- From the table, this cut-off value with an area of 0.025 above is 1.96.

- Now we compare the test statistic to 1.96 and find that Z = 1.40 < 1.96.

- This means that Z is not in the critical region (shaded area).

- Therefore, we do not reject H_0. In other words, we conclude that there is not enough evidence to claim that the percentage of Americans owning pets is not 50%.
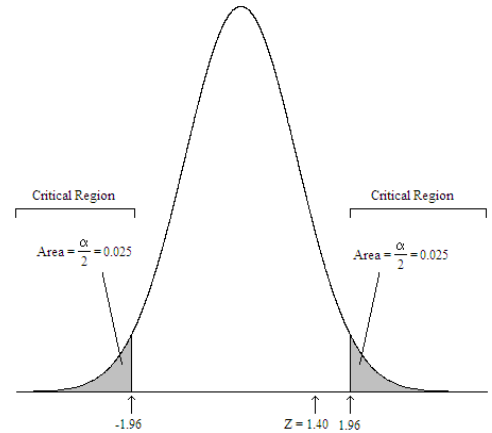
P-Value Method

- Since H_1 has a "$\neq$" sign, we want to find area above |Z| = 1.40 and below -|Z| = -1.40 for the standard normal distribution.

- From the Z-Table, this area is 2(0.0808) = 0.1616.

- Now we compare this area to α and see that 0.1616 > 0.05.

- Therefore, we do not reject H_0. In other words, we conclude that there is not enough evidence to claim that the percentage of Americans owning pets is not 50%.
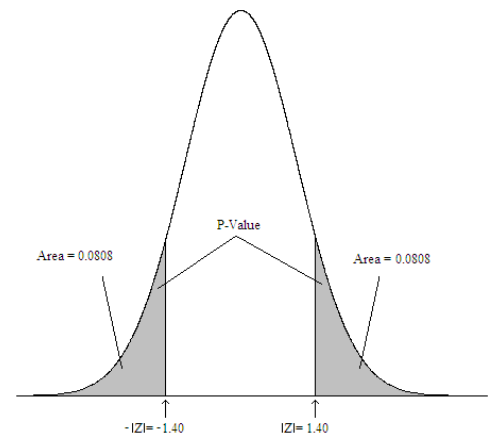
Section 8-4: Testing a Claim About a Mean: σ Known

To test a claim about a mean, we must again rely on the Central Limit Theorem. As a result, there are a few requirements that must be met in order for these techniques to work well.

Conditions

1. The sample must be a simple random sample.
2. The population standard deviation σ is known.
3. Either the population is normally distributed or $n > 30$.

(Condition 3 is for the approximation with a normal distribution using the CLT)

Notation

n = sample size
μ = the true population mean
σ = the population standard deviation
$\bar{x}$ = the sample mean

Using the Hypothesis Test for a Mean

To run the hypothesis test, we use the following test statistic (notice that the components of it are similar to the ones used in the confidence intervals in the previous chapter):

Test Statistic: $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ Critical Values and P-Values come from the Standard Normal Distribution

(Here, μ is the value used in the null hypothesis)

To test a hypothesis about the population mean, follow these steps:

1. Write down the null and alternative hypotheses, as given in the statement of the problem.
2. Identify the values of importance: μ, σ, $\bar{x}$, n, α
3. Calculate the test statistic Z.
4. Note which sign is used in the alternative hypothesis H_1.
5. Using either the traditional or p-value method, determine whether the test statistic is "unlikely".
   With a > sign, you look at the area at the high end of the distribution.
   With a < sign, you look at the area at the low end of the distribution.
   With a ≠ sign, you divide the area between the high and low ends of the distribution.
6. If it is unlikely, reject H_0. Otherwise, there is not enough evidence to reject H_0.

Example
The newspaper claims that the average number of pets owned by Americans is 1.5, but Fred thinks that it is different than that. To test this claim, he takes a simple random sample of 100 Americans and finds a sample average of 1.4. Test his claim with significance level $\alpha = 0.05$, assuming that the population standard deviation is 0.4.

Solution
From the information given, we see that:
$n = 100$, $\alpha = 0.05$, $\mu = 1.5$, $\sigma = 0.4$, $\bar{x} = 1.4$

H_0: $\mu = 1.5$
H_1: $\mu \neq 1.5$

Test Statistic: $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \dfrac{1.4 - 1.5}{0.4 / \sqrt{100}} = \dfrac{-0.1}{0.04} = -2.5$

Traditional Method

- Since H_1 has a "$\neq$" sign, we want to find the critical value that has an area of $\alpha/2$ above it and $\alpha/2$ below it on the standard normal distribution (i.e. we want the value $Z_{\alpha/2}$).

- From the table, this cut-off value with an area of 0.025 above is 1.96.

- Now we compare the test statistic to -1.96 and find that $Z = -2.5 < -1.96$.

- This means that Z is in the critical region (shaded area).

- Therefore, we reject H_0. The evidence suggests that the average number of pets owned by Americans is less than 1.5.
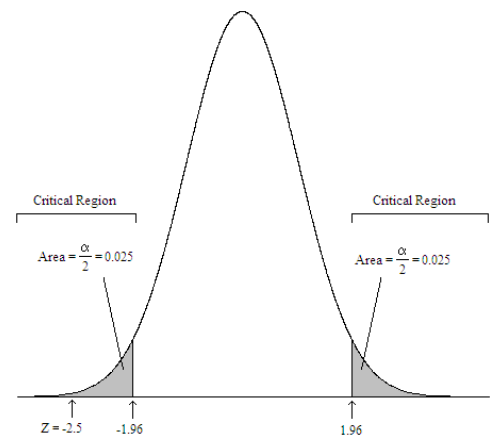
P-Value Method

- Since H_1 has a "$\neq$" sign, we want to find area above $|Z| = 2.5$ and below $-|Z| = -2.5$ for the standard normal distribution.

- From the Z-Table, this area is $2(0.0062) = 0.0124$.

- Now we compare this area to α and see that $0.0124 < 0.05$.

- Therefore, we reject H_0. The evidence suggests that the average number of pets owned by Americans is less than 1.5.
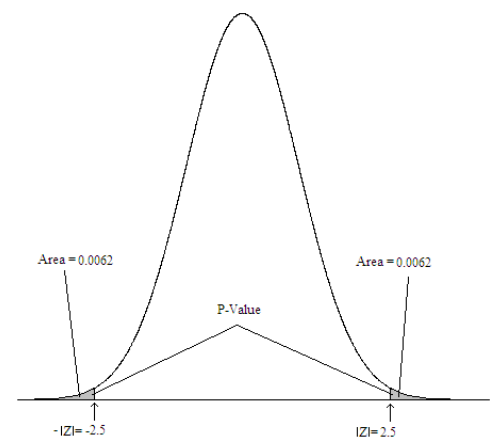
Example

Engineers in a bottle manufacturing facility want to see if the average width of the bottle wall is less than 0.1 cm. To test this, they randomly sample 500 bottles and measure the wall width. The sample average width is 0.999 cm. They also know from previous experience that the standard deviation in wall widths for the population is 0.01 cm. Run a hypothesis test with $\alpha = 0.05$ to see if the population average wall width is less than 0.1 cm.

Solution

From the information given, we see that:
$n = 500$, $\alpha = 0.05$, $\mu = 0.1$, $\sigma = 0.01$, $\bar{x} = 0.999$

H_0: $\mu = 0.1$
H_1: $\mu < 0.1$

Test Statistic: $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \dfrac{0.999 - 0.1}{0.01 / \sqrt{500}} = \dfrac{-0.001}{0.0004472} = -2.236$

Traditional Method

- Since H_1 has a "<" sign, we want to find the critical value that has an area of α below it on the standard normal distribution (i.e. we want the value $-Z_\alpha$).

- From the table, this cut-off value with an area of 0.05 below is -1.645.

- Now we compare the test statistic to -1.645 and find that $Z = -2.236 < -1.645$.

- This means that Z is in the critical region (shaded area).

- Therefore, we reject H_0. The evidence suggests that the average wall width is less than 0.1 cm.
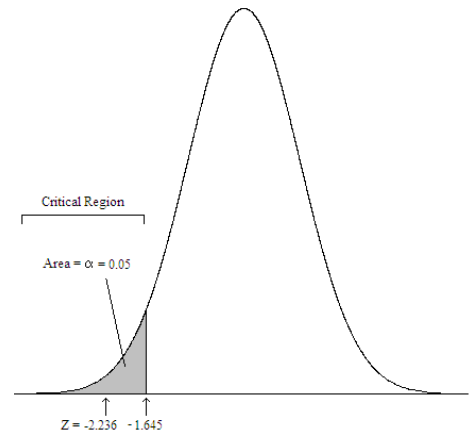
P-Value Method

- Since H_1 has a "<" sign, we want to find area below $Z = -2.236$ for the standard normal distribution.

- From the Z-Table, this area is 0.0125.

- Now we compare this area to α and see that $0.0125 < 0.05$.

- Therefore, we reject H_0. The evidence suggests that the average wall width is less than 0.1 cm.
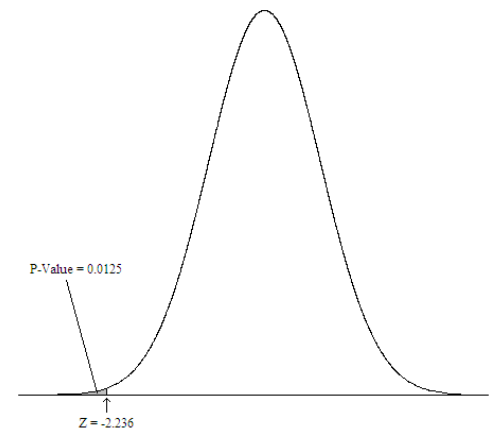
Section 8-5: Testing a Claim About a Mean: σ Unknown

When σ is unknown, we must rely on a different technique for the hypothesis test. However, the Central Limit Theorem is still used, and these conditions verify that the theoretical assumptions are met.

Conditions
1. The sample must be a simple random sample.
2. The population standard deviation σ is unknown.
3. Either the population is normally distributed or $n > 30$.

(Condition 3 is for the CLT)

Notation
n = sample size
μ = the true population mean
$\bar{x}$ = the sample mean
s = the sample standard deviation
$df = n - 1$ = degrees of freedom

Using the Hypothesis Test for a Mean
To run the hypothesis test, we use the following test statistic (notice that the components of it are similar to the ones used in the confidence intervals in the previous chapter):

Test Statistic: $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ Critical Values and P-Values come from the Student-t Distribution with $n - 1$ degrees of freedom

(Here, μ is the value used in the null hypothesis)

To test a hypothesis about the population mean, follow these steps:
1. Write down the null and alternative hypotheses, as given in the statement of the problem.
2. Identify the values of importance: μ, $\bar{x}$, s, n, α
3. Calculate the test statistic t.
4. Note which sign is used in the alternative hypothesis H_1.
5. Using either the traditional or p-value method, determine whether the test statistic is "unlikely".
    With a > sign, you look at the area at the high end of the distribution.
    With a < sign, you look at the area at the low end of the distribution.
    With a ≠ sign, you divide the area between the high and low ends of the distribution.
6. If it is unlikely, reject H_0. Otherwise, there is not enough evidence to reject H_0.

Example
A power company wants to see if the average amount of current passing through a series of connections is larger than 30 milliamperes (mA). They randomly select 35 of the connections and find a sample average current of 32 mA, with a sample standard deviation of 5 mA. Run a hypothesis test with $\alpha = 0.05$ to see if the population average current is more than 30 mA.

Solution
From the information given, we see that:
$n = 35$, $\alpha = 0.05$, $\mu = 30$, $\bar{x} = 32$, $s = 5$, $df = n - 1 = 34$

H_0: $\mu = 30$
H_1: $\mu > 30$

Test Statistic: $t = \dfrac{\bar{x} - \mu}{s / \sqrt{n}} = \dfrac{32 - 30}{5 / \sqrt{35}} = \dfrac{2}{0.845} = 2.366$

Traditional Method

- Since H_1 has a ">" sign, we want to find the critical value that has an area of α above it on the t distribution (i.e. we want the value t_α).

- From the table, this cut-off value with an area of 0.05 above is 1.691.

- Now we compare the test statistic to 1.691 and find that $t = 2.366 > 1.691$.

- This means that t is in the critical region (shaded area).

- Therefore, we reject H_0. The evidence suggests that the average current is more than 30 mA.
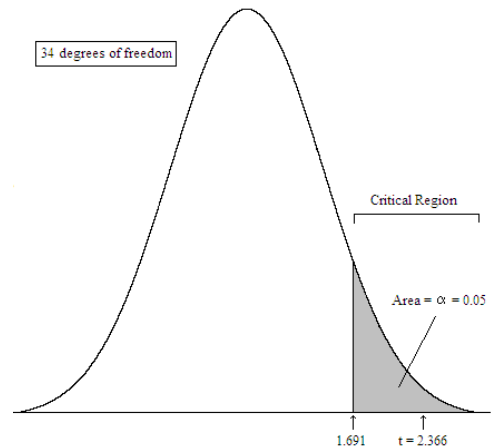
P-Value Method

- Since H_1 has a ">" sign, we want to find area above $t = 2.366$ for the t distribution.

- Now, notice that the t-Table does not allow one to directly find this area. However, for $df = 34$ we see that 2.441 has an area above of 0.01 and 2.032 has an area above of 0.025.

- Since $2.032 < t < 2.441$, the p-value will fall between 0.01 and 0.025. (see picture)

- Now we compare this area to α and see that $0.01 < p < 0.025 < 0.05 = \alpha$.

- Since $p < 0.05$, we reject H_0. The evidence suggests that the average current is more than 30 mA.
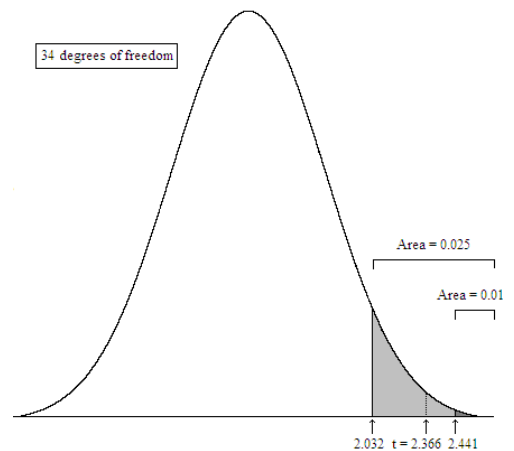
Example

10 years ago, the census bureau estimated the average number of U.S. citizens residing in rural counties to be 75,000. Now, they wish to see if the average has changed. They take a sample of 31 counties across the U.S. From the counties sampled, the average number of residents was 77,000, with a standard deviation of 5700. Using a significance level of $\alpha = 0.05$, run a test to see if the average has changed over the past decade.

Solution

From the information given, we see that:

$n = 31$, $\alpha = 0.05$, $\mu = 75,000$, $\bar{x} = 77,000$, $s = 5700$, $df = n - 1 = 30$

H_0: $\mu = 75,000$
H_1: $\mu \neq 75,000$

Test Statistic: $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{77,000 - 75,000}{5700/\sqrt{31}} = \dfrac{2000}{1023.75} = 1.954$

Traditional Method

- Since H_1 has a "$\neq$" sign, we want to find the critical value that has an area of $\alpha/2$ above it and $\alpha/2$ below it on the t distribution (i.e. we want the value $t_{\alpha/2}$).

- From the table, this cut-off value with a one-tailed area of 0.025 is 2.042.

- Now we compare the test statistic to 2.042 and find that $t = 1.954 < 2.042$.

- This means that t is not in the critical region (shaded area).

- Therefore, we do not reject H_0. There is not sufficient evidence to conclude that the average number of residents has changed.
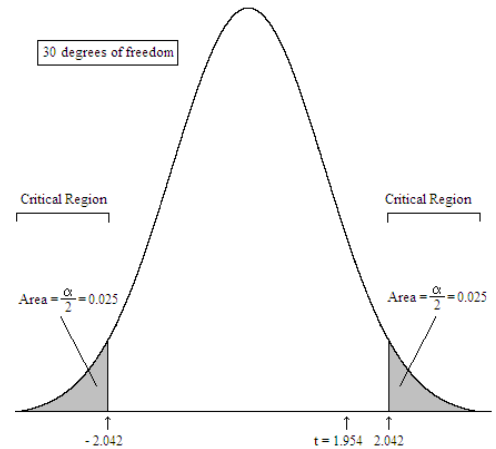
P-Value Method

- Since H_1 has a "$\neq$" sign, the p-value will be the area above $|t| = 1.954$ and below $-|t| = -1.954$ for the t distribution.

- Again, notice that the t-Table does not allow one to directly find this area. However, for $df = 30$ we see that 1.697 has a two-tailed area of 0.10 and 2.042 has a two-tailed area of 0.05.

- Since $1.697 < t < 2.042$, the p-value will fall between 0.05 and 0.10. (see picture)

- Now we compare this area to α and see that $p > 0.05 = \alpha$.

- Since $p > 0.05$, we do not reject H_0. There is not sufficient evidence to conclude that the average number of residents has changed.
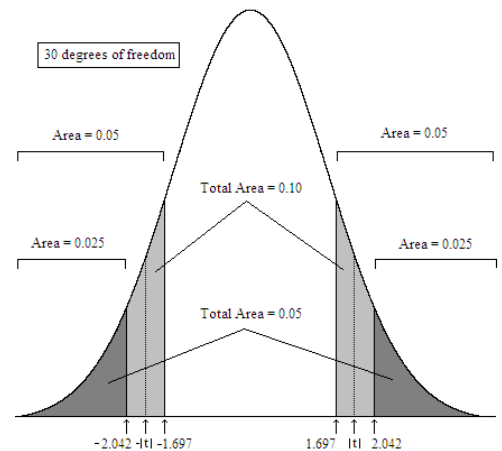
Section 8-6: Testing a Claim About a Standard Deviation or Variance

Whether you want to test a claim about standard deviation or one about variance, the procedure is the same. However, as was the case in the discussion of confidence intervals for the standard deviation or variance, the conditions under which the technique works are much more stringent.

Conditions
1. The sample must be a simple random sample.
2. The population is normally distributed.

(Note that by Condition 2, a Central Limit Theorem argument is not sufficient)

Notation
n = sample size
σ = the true population standard deviation
s = the sample standard deviation
$df = n - 1$ = degrees of freedom

Using the Hypothesis Test for a Standard Deviation or Variance

To run the hypothesis test, we use the following test statistic (notice that the components of it are similar to the ones used in the confidence intervals in the previous chapter):

Test Statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$ Critical Values and P-Values come from the Chi-Square Distribution with $n - 1$ d.f.

(Here, σ – or σ^2 – is the value used in the null hypothesis)

To test a hypothesis about the population standard deviation or variance, follow these steps:
1. Write down the null and alternative hypotheses, as given in the statement of the problem.
2. Identify the values of importance: σ, s, n, α
3. Calculate the test statistic χ^2.
4. Note which sign is used in the alternative hypothesis H_1.
5. Using either the traditional or p-value method, determine whether the test statistic is "unlikely".
     With a > sign, you look at the area at the high end of the distribution.
     With a < sign, you look at the area at the low end of the distribution.
     With a ≠ sign, you divide the area between the high and low ends of the distribution.
6. If it is unlikely, reject H_0. Otherwise, there is not enough evidence to reject H_0.

Example

In the previous census bureau question from Section 8-5, the bureau sampled 31 counties and found that the sample standard deviation in the number of residents in rural counties was 5700. Now use a hypothesis test ($\alpha = 0.05$) to see if the true population standard deviation is larger than 4800.

Solution

From the information given, we see that:

$n = 31$, $\alpha = 0.05$, $s = 5700$, $\sigma = 4800$, $df = n - 1 = 30$

H_0: $\sigma = 4800$
H_1: $\sigma > 4800$

Test Statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{30(5700^2)}{4800^2} = 42.30$

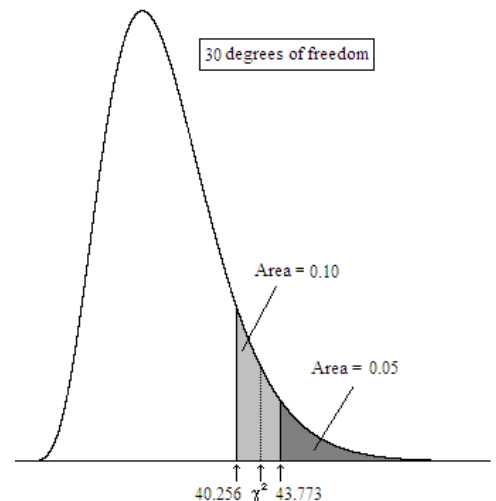*Traditional Method*

- Since H_1 has a ">" sign, we want to find the critical value that has an area of α above it on the Chi-Square distribution.

- From the table, this cut-off value with an area of 0.05 above is 43.773. (df = 30)

- Now we compare the test statistic to 43.773 and find that $\chi^2 = 42.30 < 43.773$.

- This means that χ^2 is not in the critical region (shaded area).

- Therefore, we do not reject H_0. There is not sufficient evidence to conclude that the standard deviation of residents is larger than 4800.

30 degrees of freedom

Area = α = 0.05

42.30 43.773
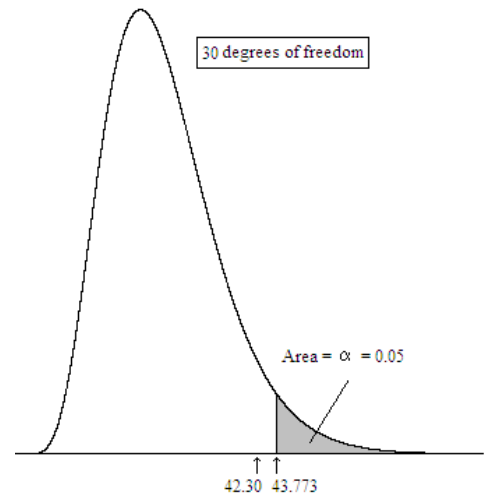
P-Value Method

- Since H_1 has a ">" sign, the p-value will be the area above $\chi^2 = 42.30$ for the Chi-Square distribution.

- As with the t distribution, notice that the table does not allow one to directly find this area. However, for $df = 30$ we see that 40.256 has an area of 0.10 above and 43.773 has an area of 0.05 above.

- Since $40.256 < \chi^2 < 43.773$, the p-value will fall between 0.05 and 0.10. (see picture)

- Now we compare this area to α and see that $p > 0.05 = \alpha$.

- Therefore, we do not reject H_0. There is not sufficient evidence to conclude that the standard deviation of residents is larger than 4800.

30 degrees of freedom

Area = 0.10

Area = 0.05

40.256 χ^2 43.773

Example
A farmer wants to see if the variance in the yield of her apple crop (in bushels) is different than 0.8. She samples 61 trees and finds that the sample variance is 0.47 bushels. Run a hypothesis test with $\alpha = 0.05$ to answer this question.

Solution
From the information given, we see that:
$n = 61$, $\alpha = 0.05$, $s^2 = 0.47$, $\sigma^2 = 0.8$, $df = n - 1 = 60$

H_0: $\sigma^2 = 0.8$
H_1: $\sigma^2 \neq 0.8$

Test Statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{60(0.47)}{0.8} = 35.25$

Traditional Method

- Since H_1 has a "$\neq$" sign, we want to find the critical value that has an area of $\alpha/2$ above it and the one with $\alpha/2$ below it on the Chi-Square distribution.

- From the table, these values are 40.482 and 83.298. (df = 60)

- Now we compare the test statistic to 40.482 and find that $\chi^2 = 35.25 < 40.482$.

- This means that χ^2 is in the critical region (shaded area).

- Therefore, we reject H_0. There is strong evidence suggesting that the true population variance in yield is different than 0.8.

60 degrees of freedom
Area $= \frac{\alpha}{2} = 0.025$
Area $= \frac{\alpha}{2} = 0.025$
35.25 40.482
83.298
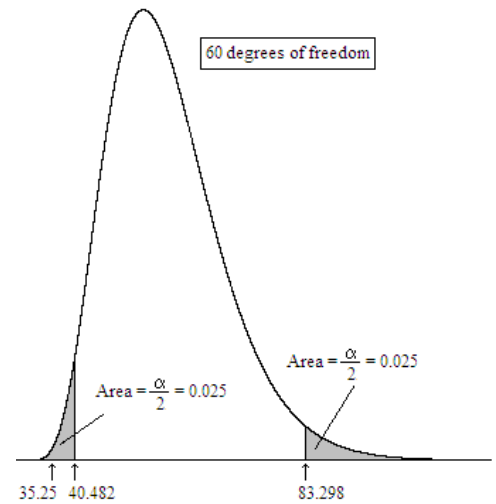
P-Value Method

- Since H_1 has a "$\neq$" sign, the p-value will twice the tail area of $\chi^2 = 42.30$ for the Chi-Square distribution.

- For $df = 30$ we see that 35.534 has an area of 0.995 above (or 0.005 below).

- Since $\chi^2 < 35.534$, the p-value will be less than 2*0.005 = 0.01.
  (see picture)

- Now we compare this area to α and see that $p < 0.01 < 0.05 = \alpha$.

- Therefore, we reject H_0. There is strong evidence suggesting that the true population variance in yield is different than 0.8.

60 degrees of freedom
P-Value
Area = 0.005
χ^2 35.534