Chapter 7: One-Sample Inference

Now that you have all this information about descriptive statistics and probabilities, it is time to start inferential statistics. There are two branches of inferential statistics: hypothesis testing and confidence intervals.

Hypothesis Testing: making a decision about a parameter(s) based on a statistic(s).
Confidence Interval: estimating a parameter(s) based on a statistic(s).

Section 7.1: Basics of Hypothesis Testing

To understand the process of a hypothesis tests, you need to first have an understanding of what a hypothesis is, which is an educated guess about a parameter. Once you have the hypothesis, you collect data and use the data to make a determination to see if there is enough evidence to show that the hypothesis is true. However, in hypothesis testing you actually assume something else is true, and then you look at your data to see how likely it is to get an event that your data demonstrates with that assumption. If the event is very unusual, then you might think that your assumption is actually false. If you are able to say this assumption is false, then your hypothesis must be true. This is known as a proof by contradiction. You assume the opposite of your hypothesis is true and show that it can't be true. If this happens, then your hypothesis must be true. All hypothesis tests go through the same process. Once you have the process down, then the concept is much easier. It is easier to see the process by looking at an example. Concepts that are needed will be detailed in this example.

Example #7.1.1: Basics of Hypothesis Testing

Suppose a manufacturer of the XJ35 battery claims the mean life of the battery is 500 days with a standard deviation of 25 days. You are the buyer of this battery and you think this claim is inflated. You would like to test your belief because without a good reason you can't get out of your contract.

What do you do?

Well first, you should know what you are trying to measure. Define the random variable.

Let x = life of a XJ35 battery

Now you are not just trying to find different x values. You are trying to find what the true mean is. Since you are trying to find it, it must be unknown. You don't think it is 500 days. If you did, you wouldn't be doing any testing. The true mean, μ, is unknown. That means you should define that too.

Let μ = mean life of a XJ35 battery

Now what?

You may want to collect a sample. What kind of sample?

You could ask the manufacturers to give you batteries, but there is a chance that there could be some bias in the batteries they pick. To reduce the chance of bias, it is best to take a random sample.

How big should the sample be?

A sample of size 30 or more means that you can use the central limit theorem. Pick a sample of size 30.

Table #7.1.1 contains the data for the sample you collected:

Table #7.1.1: Data on Battery Life

491	485	503	492	482	490
489	495	497	487	493	480
482	504	501	486	478	492
482	502	485	503	497	500
488	475	478	490	487	486

Now what should you do? Looking at the data set, you see some of the times are above 500 and some are below. But looking at all of the numbers is too difficult. It might be helpful to calculate the mean for this sample.

The sample mean is $\bar{x} = 490 \text{ days}$. Looking at the sample mean, one might think that you are right. However, the standard deviation and the sample size also plays a role, so maybe you are wrong.

Before going any farther, it is time to formalize a few definitions.

You have a guess that the mean life of a battery is less than 500 days. This is opposed to what the manufacturer claims. There really are two hypotheses, which are just guesses here – the one that the manufacturer claims and the one that you believe. It is helpful to have names for them.

Null Hypothesis: historical value, claim, or product specification. The symbol used is H_o.

Alternate Hypothesis: what you want to prove. This is what you want to accept as true when you reject the null hypothesis. There are two symbols that are commonly used for the alternative hypothesis: H_A or H_1. The symbol H_A will be used in this book.

In general, the hypotheses look something like this:

$H_o : \mu = \mu_o$

$H_A : \mu < \mu_o$

where μ_o just represents the value that the claim says the population mean is actually equal to.
Also, H_A can be less than, greater than, or not equal to.

For this problem:

$H_o: \mu = 500 \text{ days}$, since the manufacturer says the mean life of a battery is 500 days.

$H_A: \mu < 500 \text{ days}$, since you believe that the mean life of the battery is less than 500 days.

Now back to the mean. You have a sample mean of 490 days. Is this small enough to believe that you are right and the manufacturer is wrong? How small does it have to be?

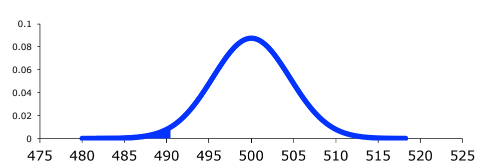
If you calculated a sample mean of 235, you would definitely believe the population mean is less than 500. But even if you had a sample mean of 435 you would probably believe that the true mean was less than 500. What about 475? Or 483? There is some point where you would stop being so sure that the population mean is less than 500. That point separates the values of where you are sure or pretty sure that the mean is less than 500 from the area where you are not so sure. How do you find that point?

Well it depends on how much error you want to make. Of course you don't want to make any errors, but unfortunately that is unavoidable in statistics. You need to figure out how much error you made with your sample. Take the sample mean, and find the probability of getting another sample mean less than it, assuming for the moment that the manufacturer is right. The idea behind this is that you want to know what is the chance that you could have come up with your sample mean even if the population mean really is 500 days.

You want to find $P\left(\bar{x} < 490 \middle| H_o \text{ is true}\right) = P\left(\bar{x} < 490 \middle| \mu = 500\right)$

To compute this probability, you need to know how the sample mean is distributed. Since the sample size is at least 30, then you know the sample mean is approximately normally distributed. Remember $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

A picture is always useful.



Before calculating the probability, it is useful to see how many standard deviations away from the mean the sample mean is. Using the formula for the *z*-score from chapter 6, you find

$$z = \frac{\overline{x} - \mu_o}{\sigma/\sqrt{n}} = \frac{490 - 500}{25/\sqrt{30}} = -2.19$$

This sample mean is more than two standard deviations away from the mean. That seems pretty far, but you should look at the probability too.

$$P\left(\overline{x} < 490 | \mu = 500\right) = \text{normalcdf}\left(-1\text{E}99, 490, 500, 25 \div \sqrt{(30)}\right) \approx 0.0142$$

There is a 1.42% chance that you could find a sample mean less than 490 when the population mean is 500 days. This is really small, so the chances are that the assumption that the population mean is 500 days is wrong, and you can reject the manufacturer's claim. But how do you quantify really small? Is 5% or 10% or 15% really small? How do you decide?

Before you answer that question, a couple more definitions are needed.

Test statistic: $z = \frac{\overline{x} - \mu_o}{\sigma/\sqrt{n}}$ since it is calculated as part of the testing of the hypothesis

p – value: probability that the test statistic will take on more extreme values than the observed test statistic, given that the null hypothesis is true. It is the probability that was calculated above.

Now, how small is small enough? To answer that, you really want to know the types of errors you can make.

There are actually only two errors that can be made. The first error is if you say that H_o is false, when in fact it is true. This means you reject H_o when H_o was true. The second error is if you say that H_o is true, when in fact it is false. This means you fail to reject H_o when H_o is false. The following table organizes this for you:

Type of errors:

	H_o true	H_o false
Reject H_o	Type I error	No error
Fail to reject H_o	No error	Type II error

Thus

Type I Error is rejecting H_o when H_o is true, and
Type II Error is accepting H_o when H_o is false.

Since these are the errors, then one can define the probabilities attached to each error.

α = P(type I error) = P(rejecting H_o / H_o is true)
β = P(type II error) = P(failing to reject H_o / H_o is false)

α is also called the **level of significance**.

Another common concept that is used is Power = $1-\beta$.

Now there is a relationship between α and β. They are not complements of each other. How are they related?

If α increases that means the chances of making a type I error will increase. It is more likely that a type I error will occur. It makes sense that you are less likely to make type II errors, only because you will be rejecting H_o more often. You will be failing to reject H_o less, and therefore, the chance of making a type II error will decrease. Thus, as α increases, β will decrease, and vice versa. That makes them seem like complements, but they aren't complements. What gives? Consider one more factor – sample size.

Consider if you have a larger sample that is representative of the population, then it makes sense that you have more accuracy then with a smaller sample. Think of it this way, which would you trust more, a sample mean of 490 if you had a sample size of 35 or sample size of 350 (assuming a representative sample)? Of course the 350 because there are more data points and so more accuracy. If you are more accurate, then there is less chance that you will make any error. By increasing the sample size of a representative sample, you decrease both α and β.

Summary of all of this:
1. For a certain sample size, *n*, if α increases, β decreases.
2. For a certain level of significance, α, if *n* increases, β decreases.

Now how do you find α and β? Well α is actually chosen. There are only three values that are usually picked for α: 0.01, 0.05, and 0.10. β is very difficult to find, so usually it isn't found. If you want to make sure it is small you take as large of a sample as you can afford provided it is a representative sample. This is one use of the Power. You want β to be small and the Power of the test is large. The Power word sounds good.

Which pick of α do you pick? Well that depends on what you are working on. Remember in this example you are the buyer who is trying to get out of a contract to buy these batteries. If you create a type I error, you said that the batteries are bad when they aren't, most likely the manufacturer will sue you. You want to avoid this. You might pick α to be 0.01. This way you have a small chance of making a type I error. Of course this means you have more of a chance of making a type II error. No big deal right? What if the batteries are used in pacemakers and you tell the person that their pacemaker's batteries are good for 500 days when they actually last less, that might be bad. If you make a type II error, you say that the batteries do last 500 days when they last less, then you have the possibility of killing someone. You certainly do not want to do this. In this case you might want to pick α as 0.10. If both errors are equally bad, then pick α as 0.05.

The above discussion is why the choice of α depends on what you are researching. As the researcher, you are the one that needs to decide what α level to use based on your analysis of the consequences of making each error is.

If a type I error is really bad, then pick $\alpha = 0.01$. If a type II error is really bad, then pick $\alpha = 0.10$ If neither error is bad, or both are equally bad, then pick $\alpha = 0.05$

The main thing is to always pick the α before you collect the data and start the test.

The above discussion was long, but it is really important information. If you don't know what the errors of the test are about, then there really is no point in making conclusions with the tests. Make sure you understand what the two errors are and what the probabilities are for them.

Now it is time to go back to the example and put this all together. This is the basic structure of testing a hypothesis, usually called a hypothesis test. Since this one has a test statistic involving *z*, it is also called a *z*-test. And since there is only one sample, it is usually called a one-sample *z*-test.

Example #7.1.2: Battery Example Revisited.

1. State the random variable and the parameter in words

    x = life of battery
    μ = mean life of a XJ35 battery

2. State the null and alternative hypothesis and the level of significance

    $H_o: \mu = 500$ days
    $H_A: \mu < 500$ days
    $\alpha = 0.10$ (from above discussion about consequences)

3. State and check the assumptions for a hypothesis test

    Every hypothesis has some assumptions that be met to make sure that the results of the hypothesis are valid. The assumptions are different for each test. This test has the following assumptions.

    a. A random sample of size *n* is taken.
    This occurred in this example, since it was stated that a random sample of 30 battery lives were taken.
    b. The population standard deviation is known.
    This is true, since it was given in the problem.
    c. The sample size is at least 30 or the population of the random variable is normally distributed.

The sample size was 30, so this condition is met.

4. Find the sample statistic, test statistic, and p-value

   The test statistic depends on how many samples there are, what parameter you are testing, and assumptions that need to be checked. In this case, there is one sample and you are testing the mean. The assumptions were checked above.

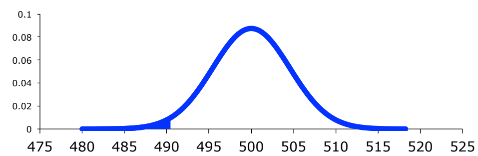
   Sample statistic:

   $\bar{x} = 490$

   Test statistic:

   $$z = \frac{\bar{x} - \mu_o}{\sigma/\sqrt{n}} = \frac{490 - 500}{25/\sqrt{30}} = -2.19$$

   p-value:



   $$P\left(\bar{x} < 490 \mid \mu = 500\right) = \text{normalcdf}\left(-1E99, 490, 500, 25/\sqrt{(30)}\right) \approx 0.0142$$

5. Conclusion:

   Now what? Well, this p-value is 0.0142. This is a lot smaller than the amount of error you would accept in the problem - $\alpha = 0.10$. That means that finding a sample mean less than 490 days is unusual to happen if H_o is true. This should make you think that H_o is not true. You should reject H_o.

In fact, in general:
Reject H_o if the p-value < α and
Fail to reject H_o if the p-value $\geq \alpha$.

6. Interpretation:

   Since you rejected H_o, what does this mean in the real world? That is what goes in the interpretation. Since you rejected the claim by the manufacturer that the mean life of the batteries is 500 days, then you now can believe that your hypothesis was correct. In other words, there is enough evidence to show that the mean life of the battery is less than 500 days.

Now that you know that the batteries last less than 500 days, should you cancel the contract? Statistically, there is evidence that the batteries do not last as long as the manufacturer says they should. However, based on this sample there are only ten days less on average that the batteries last. There may not be practical significance in this case. Ten days do not seem like a large difference. In reality, if the batteries are used in pacemakers, then you would probably tell the patient to have the batteries replaced every year. You have a large buffer whether the batteries last 490 days or 500 days. It seems that it might not be worth it to break the contract over ten days. What if the 10 days was practically significant? Are there any other things you should consider? You might look at the business relationship with the manufacturer. You might also look at how much it would cost to find a new manufacturer. These are also questions to consider before making any changes. What this discussion should show you is that just because a hypothesis has statistical significance does not mean it has practical significance. The hypothesis test is just one part of a research process. There are other pieces that you need to consider.

That's it. That is what a hypothesis test looks like. All hypothesis tests are done with the same six steps. Those general six steps are outlined below.

1. State the random variable and the parameter in words. This is where you are defining what the unknowns are in this problem.
   x = random variable
   μ = mean of random variable, if the parameter of interest is the mean. There are other parameters you can test, and you would use the appropriate symbol for that parameter.

2. State the null and alternative hypotheses and the level of significance
   $H_o : \mu = \mu_o$, where μ_o is the known mean
   $H_A : \mu < \mu_o$
   $H_A : \mu > \mu_o$, use the appropriate one for your problem
   $H_A : \mu \neq \mu_o$
   Also, state your α level here.

3. State and check the assumptions for a hypothesis test
   Each hypothesis test has its own assumptions. They will be stated when the different hypothesis tests are discussed.

4. Find the sample statistic, test statistic, and p-value
   This depends on what parameter you are working with, how many samples, and the assumptions of the test. The p-value depends on your H_A. If you are doing the H_A with the less than, then it is a left-tailed test, and you find the probability of being in that left tail. If you are doing the H_A with the greater than, then it is a right-tailed test, and you find the probability of being in the right tail. If you are doing the H_A with the not equal to, then you are doing a two-tail test, and you find the probability of being in both tails. Because of symmetry, you could find the probability in one tail and double this value to find the probability in both tails.

5. Conclusion
   This is where you write reject H_o or fail to reject H_o. The rule is: if the p-value $< \alpha$, then reject H_o. If the p-value $\geq \alpha$, then fail to reject H_o

6. Interpretation
   This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show H_A is true, or you do not have enough evidence to show H_A is true.

Sorry, one more concept about the conclusion and interpretation. First, the conclusion is that you reject H_o or you fail to reject H_o. Why was it said like this? It is because you never **accept** the null hypothesis. If you wanted to accept the null hypothesis, then why do the test in the first place? In the interpretation, you either have enough evidence to show H_A is true, or you do not have enough evidence to show H_A is true. You wouldn't want to go to all this work and then find out you wanted to accept the claim. Why go through the trouble? You always want to show that the alternative hypothesis is true. Sometimes you can do that and sometimes you can't. It doesn't mean you proved the null hypothesis; it just means you can't prove the alternative hypothesis. Here is an example to demonstrate this.

Example #7.1.3: Conclusions in Hypothesis Tests

In the U.S. court system a jury trial could be set up as a hypothesis test. To really help you see how this works, let's use OJ Simpson as an example. In the court system, a person is presumed innocent until he/she is proven guilty, and this is your null hypothesis. OJ Simpson was a football player in the 1970s. In 1994 his ex-wife and her friend were killed. OJ Simpson was accused of the crime, and in 1995 the case was tried. The prosecutors wanted to prove OJ was guilty of killing his wife and her friend, and that is the alternative hypothesis

$$H_0: \text{OJ is innocent of killing his wife and her friend}$$
$$H_A: \text{OJ is guilty of killing his wife and her friend}$$

In this case, a verdict of not guilty was given. That does not mean that he is innocent of this crime. It means there was not enough evidence to prove he was guilty. Many people believe that OJ was guilty of this crime, but the jury did not feel that the evidence presented was enough to show there was guilt. The verdict in a jury trial is always guilty or not guilty!

The same is true in a hypothesis test. There is either enough evidence or not enough evidence to show that alternative hypothesis. It is not that you proved the null hypothesis true.

Example #7.1.4: Stating Hypotheses

Identify the hypotheses necessary to test the following statements:

a.) The average salary of a teacher is more than $30,000.

Solution:

x = salary of teacher

μ = mean salary of teacher

The guess is that $\mu > \$30,000$ and that is the alternative hypothesis.

The null hypothesis has the same parameter and number with an equal sign.

$$H_0 : \mu = \$30,000$$

$$H_A : \mu > \$30,000$$

b.) The proportion of students who like math is less than 10%.

Solution:

x = number of students who like math

p = proportion of students who like math

The guess is that $p < 0.10$ and that is the alternative hypothesis.

$$H_0 : p = 0.10$$

$$H_A : p < 0.10$$

c.) The average age of students in this class differs from 21.

Solution:

x = age of students in this class

μ = mean age of students in this class

The guess is that $\mu \neq 21$ and that is the alternative hypothesis.

$$H_0 : \mu = 21$$

$$H_A : \mu \neq 21$$

Example #7.1.5: Stating Type I and II Errors and Picking Level of Significance

a.) The plant-breeding department at a major university developed a new hybrid raspberry plant called YumYum Berry. Based on research data, the claim is made that from the time shoots are planted 90 days on average are required to obtain the first berry with a standard deviation of 9.2 days. A corporation that is interested in marketing the product tests 60 shoots by planting them and recording the number of days before each plant produces its first berry. The sample mean is 92.3 days. The corporation wants to know if the mean number of days is more than the 90 days claimed. State the type I and type II errors in terms of this problem, consequences of each error, and state which level of significance to use.

Solution:

x = time to first berry for YumYum Berry plant

μ = mean time to first berry for YumYum Berry plant

$$H_0 : \mu = 90$$

$$H_A : \mu > 90$$

Type I Error: If the corporation does a type I error, then they will say that the plants take longer to produce than 90 days when they don't. They probably will not want to market the plants if they think they will take longer. They will not market them even though in reality the plants do produce in 90 days. They may have loss of future earnings, but that is all.

Type II error: The corporation do not say that the plants take longer then 90 days to produce when they do take longer. Most likely they will market the plants. The plants will take longer, and so customers might get upset and then the company would get a bad reputation. This would be really bad for the company.

Level of significance: It appears that the corporation would not want to make a type II error. Pick a 10% level of significance, $\alpha = 0.10$.

b.) A concern was raised in Australia that the percentage of deaths of Aboriginal prisoners was higher than the percent of deaths of non-indigenous prisoners, which is 0.27%. State the type I and type II errors in terms of this problem, consequences of each error, and state which level of significance to use.

Solution:

x = number of Aboriginal prisoners who have died

p = proportion of Aboriginal prisoners who have died

$$H_o : p = 0.27\%$$

$$H_A : p > 0.27\%$$

Type I error: Rejecting that the proportion of Aboriginal prisoners who died was 0.27%, when in fact it was 0.27%. This would mean you would say there is a problem when there isn't one. You could anger the Aboriginal community, and spend time and energy researching something that isn't a problem.

Type II error: Accepting that the proportion of Aboriginal prisoners who died was 0.27%, when in fact it is higher than 0.27%. This would mean that you wouldn't think there was a problem with Aboriginal prisoners dying when there really is a problem. You risk causing deaths when there could be a way to avoid them.

Level of significance: It appears that both errors may be issues in this case. You wouldn't want to anger the Aboriginal community when there isn't an issue, and you wouldn't want people to die when there may be a way to stop it. It may be best to pick a 5% level of significance, $\alpha = 0.05$.

Hint – hypothesis testing is really easy if you follow the same recipe every time. The only differences in the various problems are the assumptions of the test and the test statistic you calculate so you can find the p-value. Do the same steps, in the same order, with the same words, every time and these problems become very easy.

Section 7.1: Homework

For the problems in this section, a question is being asked. This is to help you understand what the hypotheses are. You are not to run any hypothesis tests in this section.

1.) Eyeglassomatic manufactures eyeglasses for different retailers. They test to see how many defective lenses they made in a given time period and found that 11% of all lenses had defects of some type. Looking at the type of defects, they found in a three-month time period that out of 34,641 defective lenses, 5865 were due to scratches. Are there more defects from scratches than from all other causes? State the random variable, population parameter, and hypotheses.

2.) According to the February 2008 Federal Trade Commission report on consumer fraud and identity theft, 23% of all complaints in 2007 were for identity theft. In that year, Alaska had 321 complaints of identity theft out of 1,432 consumer complaints ("Consumer fraud and," 2008). Does this data provide enough evidence to show that Alaska had a lower proportion of identity theft than 23%? State the random variable, population parameter, and hypotheses.

3.) The Kyoto Protocol was signed in 1997, and required countries to start reducing their carbon emissions. The protocol became enforceable in February 2005. In 2004, the mean CO_2 emission was 4.87 metric tons per capita. Is there enough evidence to show that the mean CO_2 emission is lower in 2010 than in 2004? State the random variable, population parameter, and hypotheses.

4.) Stephen Stigler determined in 1977 that the speed of light is 299,710.5 km/sec. In 1882, Albert Michelson had collected measurements on the speed of light ("Student t-distribution," 2013). Is there evidence to show that Michelson's data is different from Stigler's value of the speed of light? State the random variable, population parameter, and hypotheses.

5.) Eyeglassomatic manufactures eyeglasses for different retailers. They test to see how many defective lenses they made in a given time period and found that 11% of all lenses had defects of some type. Looking at the type of defects, they found in a three-month time period that out of 34,641 defective lenses, 5865 were due to scratches. Are there more defects from scratches than from all other causes? State the type I and type II errors in this case, consequences of each error type for this situation, and the appropriate alpha level to use.

6.) According to the February 2008 Federal Trade Commission report on consumer fraud and identity theft, 23% of all complaints in 2007 were for identity theft. In that year, Alaska had 321 complaints of identity theft out of 1,432 consumer complaints ("Consumer fraud and," 2008). Does this data provide enough evidence to show that Alaska had a lower proportion of identity theft than 23%? State the type I and type II errors in this case, consequences of each error type for this situation, and the appropriate alpha level to use.

7.) The Kyoto Protocol was signed in 1997, and required countries to start reducing their carbon emissions. The protocol became enforceable in February 2005. In 2004, the mean CO_2 emission was 4.87 metric tons per capita. Is there enough evidence to show that the mean CO_2 emission is lower in 2010 than in 2004? State the type I and type II errors in this case, consequences of each error type for this situation, and the appropriate alpha level to use.

8.) Stephen Stigler determined in 1977 that the speed of light is 299,710.5 km/sec. In 1882, Albert Michelson had collected measurements on the speed of light ("Student t-distribution," 2013). Is there evidence to show that Michelson's data is different from Stigler's value of the speed of light? State the type I and type II errors in this case, consequences of each error type for this situation, and the appropriate alpha level to use.

Section 7.2: One-Sample Proportion Test

There are many different parameters that you can test. There is a test for the mean, such as was introduced with the z-test. There is also a test for the population proportion, p. This is where you might be curious if the proportion of students who smoke at your school is lower than the proportion in your area. Or you could question if the proportion of accidents caused by teenage drivers who do not have a drivers' education class is more than the national proportion.

To test a population proportion, there are a few things that need to be defined first. Usually, Greek letters are used for parameters and Latin letters for statistics. When talking about proportions, it makes sense to use p for proportion. The Greek letter for p is π, but that is too confusing to use. Instead, it is best to use p for the population proportion. That means that a different symbol is needed for the sample proportion. The convention is to use, $\hat{p}$, known as p-hat. This way you know that p is the population proportion, and that $\hat{p}$ is the sample proportion related to it.

Now proportion tests are about looking for the percentage of individuals who have a particular attribute. You are really looking for the number of successes that happen. Thus, a proportion test involves a binomial distribution.

Hypothesis Test for One Population Proportion (1-Prop Test)

1. State the random variable and the parameter in words.
   x = number of successes
   p = proportion of successes

2. State the null and alternative hypotheses and the level of significance
   $H_o : p = p_o$, where p_o is the known proportion
   $H_A : p < p_o$
   $H_A : p > p_o$, use the appropriate one for your problem
   $H_A : p \neq p_o$
   Also, state your α level here.

3. State and check the assumptions for a hypothesis test
   a. A simple random sample of size n is taken.
   b. The conditions for the binomial distribution are satisfied
   c. To determine the sampling distribution of $\hat{p}$, you need to show that $np \geq 5$ and $nq \geq 5$, where $q = 1 - p$. If this requirement is true, then the sampling distribution of $\hat{p}$ is well approximated by a normal curve.

4. Find the sample statistic, test statistic, and p-value
   Sample Proportion:
   $$\hat{p} = \frac{x}{n} = \frac{\text{\# of successes}}{\text{\# of trials}}$$
   Test Statistic:

$$z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$$

p-value:
Use normalcdf(lower limit, upper limit, 0, 1)
(Note: if $H_A: p < p_o$, then lower limit is $-1E99$ and upper limit is your test statistic. If $H_A: p > p_o$, then lower limit is your test statistic and the upper limit is $1E99$. If $H_A: p \neq p_o$, then find the p-value for $H_A: p < p_o$, and multiply by 2.)

5. Conclusion
   This is where you write reject H_o or fail to reject H_o. The rule is: if the p-value $< \alpha$, then reject H_o. If the p-value $\geq \alpha$, then fail to reject H_o

6. Interpretation
   This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show H_A is true, or you do not have enough evidence to show H_A is true.

Example #7.2.1: Hypothesis Test for One Proportion Using Formula

A concern was raised in Australia that the percentage of deaths of Aboriginal prisoners was higher than the percent of deaths of non-Aboriginal prisoners, which is 0.27%. A sample of six years (1990-1995) of data was collected, and it was found that out of 14,495 Aboriginal prisoners, 51 died ("Indigenous deaths in," 1996). Do the data provide enough evidence to show that the proportion of deaths of Aboriginal prisoners is more than 0.27%?

Solution:

1. State the random variable and the parameter in words.
   x = number of Aboriginal prisoners who die
   p = proportion of Aboriginal prisoners who die

2. State the null and alternative hypotheses and the level of significance
   $H_o: p = 0.0027$
   $H_A: p > 0.0027$
   Example #7.1.5b argued that the $\alpha = 0.05$.

3. State and check the assumptions for a hypothesis test

a. A simple random sample of 14,495 Aboriginal prisoners was taken. However, the sample was not a random sample, since it was data from six years. It is the numbers for all prisoners in these six years, but the six years were not picked at random. Unless there was something special about the six years that were chosen, the sample is probably a representative sample. This assumption is probably met.

b. There are 14,495 prisoners in this case. The prisoners are all Aboriginals, so you are not mixing Aboriginal with non-Aboriginal prisoners. There are only two outcomes, either the prisoner dies or doesn't. The chance that one prisoner dies over another may not be constant, but if you consider all prisoners the same, then it may be close to the same probability. Thus the conditions for the binomial distribution are satisfied

c. In this case p = 0.0027 and n = 14,495. $np = 14495 * 0.0027 \approx 39 \geq 5$ and $nq = 14495 * (1 - 0.0027) \approx 14456 \geq 5$. So, the sampling distribution for $\hat{p}$ is a normal distribution.

4. Find the sample statistic, test statistic, and p-value

   Sample Proportion:

$$x = 51$$
$$n = 14495$$
$$\hat{p} = \frac{x}{n} = \frac{51}{14495} \approx 0.003518$$

   Test Statistic:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.003518 - 0.0027}{\sqrt{\frac{0.0027(1 - 0.0027)}{14495}}} \approx 1.8979$$

   p-value:

   p-value = $P(z > 1.8979) = \text{normalcdf}(1.8979, 1E99, 0, 1) \approx 0.029$

5. Conclusion

   Since the p-value < 0.05, then reject H_o .

6. Interpretation

   There is enough evidence to show that the proportion of deaths of Aboriginal prisoners is more than for non-Aboriginal prisoners.

Example #7.2.2: Hypothesis Test for One Proportion Using TI-83/84 Calculator

A researcher who is studying the effects of income levels on breastfeeding of infants hypothesizes that countries where the income level is lower have a higher rate of infant breastfeeding than higher income countries. It is known that in Germany, considered a high-income country by the World Bank, 22% of all babies are breastfeed. In Tajikistan, considered a low-income country by the World Bank, researchers found that in a random sample of 500 new mothers that 125 were breastfeeding their infant. At the 5% level of significance, does this show that low-income countries have a higher incident of breastfeeding?

Solution:

1. State you random variable and the parameter in words.

   x = number of woman who breastfeed in a low-income country

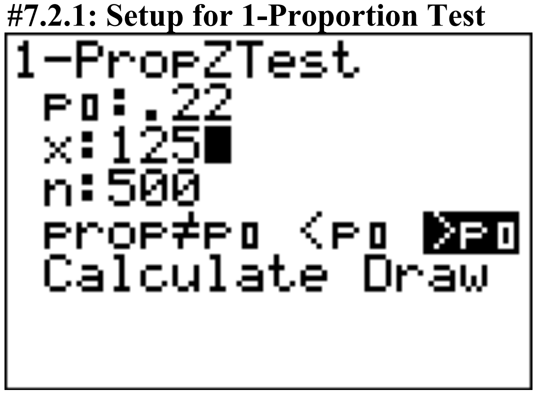
   p = proportion of woman who breastfeed in a low-income country

2. State the null and alternative hypotheses and the level of significance
   $H_o : p = 0.22$
   $H_A : p > 0.22$
   $\alpha = 0.05$

3. State and check the assumptions for a hypothesis test
   a. A simple random sample of 500 breastfeeding habits of woman in a low-income country was taken as was stated in the problem.
   b. There were 500 women in the study. The women are considered identical, though they probably have some differences. There are only two outcomes, either the woman breastfeeds or she doesn't. The probability of a woman breastfeeding is probably not the same for each woman, but it is probably not very different for each woman. The conditions for the binomial distribution are satisfied
   c. In this case, n = 500 and p = 0.22. $np = 500(0.22) = 110 \ge 5$ and $nq = 500(1-0.22) = 390 \ge 5$, so the sampling distribution of $\hat{p}$ is well approximated by a normal curve.

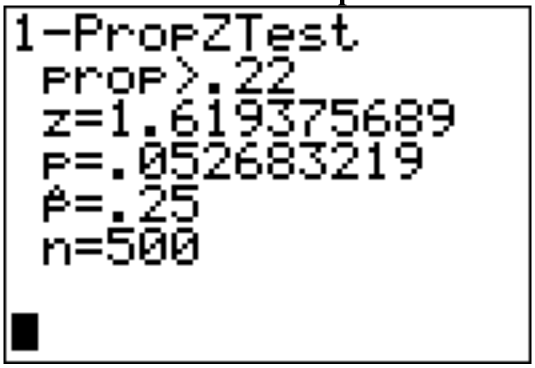
4. Find the sample statistic, test statistic, and p-value
   This time, all calculations will be done on the TI-83/84 calculator. Go into the STAT menu, then arrow over to TESTS. This test is a 1-propZTest. Then type in the information just as shown in figure #7.2.1.

Figure #7.2.1: Setup for 1-Proportion Test



Once you press Calculate, you will see the results as in figure #7.2.2.

Figure #7.2.2: Results for 1-Proportion Test



The z in the results is the test statistic. The p = 0.052683219 is the p-value, and the $\hat{p} = 0.25$ is the sample proportion.

The p-value is approximately 0.053

5. Conclusion
   Since the p-value is more than 0.05, you fail to reject H_o.

6. Interpretation
   There is not enough evidence to show that the proportion of women who breastfeed in low-income countries is more than in high-income countries.

Notice, the conclusion is that there wasn't enough evidence to show what H_1 said. The conclusion was not that you proved H_o true. There are many reasons why you can't say that H_o is true. It could be that the countries you chose were not very representative of what truly happens. If you instead looked at all high-income countries and compared them to low-income countries, you might have different results. It could also be that the sample you collected in the low-income country was not representative. It could also be that income level is not an indication of breastfeeding habits. There could be other factors involved. This is why you can't say that you have proven H_o is true. There are too many other factors that could be the reason that you failed to reject H_o.

Section 7.2: Homework

In each problem show all steps of the hypothesis test. If some of the assumptions are not met, note that the results of the test may not be correct and then continue the process of the hypothesis test.

1.) Eyeglassomatic manufactures eyeglasses for different retailers. They test to see how many defective lenses they made in a given time period and found that 11% of all lenses had defects of some type. Looking at the type of defects, they found in a three-month time period that out of 34,641 defective lenses, 5865 were due to scratches. Are there more defects from scratches than from all other causes? Use a 1% level of significance.

2.) In July of 1997, Australians were asked if they thought unemployment would increase, and 47% thought that it would increase. In November of 1997, they were asked again. At that time 284 out of 631 said that they thought unemployment would increase ("Morgan gallup poll," 2013). At the 5% level, is there enough evidence to show that the proportion of Australians in November 1997 who believe unemployment would increase is less than the proportion who felt it would increase in July 1997?

3.) According to the February 2008 Federal Trade Commission report on consumer fraud and identity theft, 23% of all complaints in 2007 were for identity theft. In that year, Arkansas had 1,601 complaints of identity theft out of 3,482 consumer complaints ("Consumer fraud and," 2008). Does this data provide enough evidence to show that Arkansas had a higher proportion of identity theft than 23%? Test at the 5% level.

4.) According to the February 2008 Federal Trade Commission report on consumer fraud and identity theft, 23% of all complaints in 2007 were for identity theft. In that year, Alaska had 321 complaints of identity theft out of 1,432 consumer complaints ("Consumer fraud and," 2008). Does this data provide enough evidence to show that Alaska had a lower proportion of identity theft than 23%? Test at the 5% level.

5.) In 2001, the Gallup poll found that 81% of American adults believed that there was a conspiracy in the death of President Kennedy. In 2013, the Gallup poll asked 1,039 American adults if they believe there was a conspiracy in the assassination, and found that 634 believe there was a conspiracy ("Gallup news service," 2013). Do the data show that the proportion of Americans who believe in this conspiracy has decreased? Test at the 1% level.

6.) In 2008, there were 507 children in Arizona out of 32,601 who were diagnosed with Autism Spectrum Disorder (ASD) ("Autism and developmental," 2008). Nationally 1 in 88 children are diagnosed with ASD ("CDC features -," 2013). Is there sufficient data to show that the incident of ASD is more in Arizona than nationally? Test at the 1% level.

Section 7.3 One-Sample Test for the Mean

It is time to go back to look at the test for the mean that was introduced in section 7.1 called the z-test. In the example, you knew what the population standard deviation, σ, was. What if you don't know σ?

You could just use the sample standard deviation, *s*, as an approximation of σ. That means the test statistic is now $\frac{\bar{x}-\mu}{s/\sqrt{n}}$. Great, now you can go and find the p-value using the normal curve. Or can you? Is this new test statistic normally distributed? Actually, it is not. How is it distributed? A man named W. S. Gossett figured out what this distribution is and called it the Student's t-distribution. There are some assumptions that must be made for this formula to be a Student's t-distribution. These are outlined in the following theorem. Note: the t-distribution is called the Student's t-distribution because that is the name he published under because he couldn't publish under his own name due to employer not wanting him to publish under his own name. His employer by the way was Guinness and they didn't want competitors knowing they had a chemist working for them. It is not called the Student's t-distribution because it is only used by students.

Theorem: If the following assumptions are met

a. A random sample of size *n* is taken.

b. The distribution of the random variable is normal or the sample size is 30 or more.

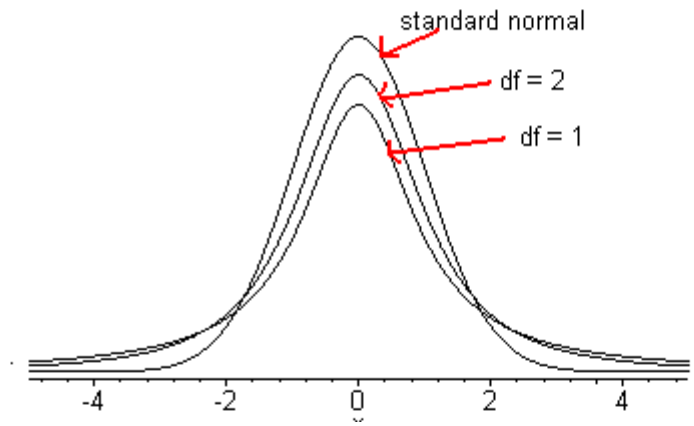
Then the distribution of $t=\frac{\bar{x}-\mu}{s/\sqrt{n}}$ is a Student's t-distribution with $n-1$ degrees of freedom.

Explanation of degrees of freedom: Recall the formula for sample standard deviation is $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$. Notice the denominator is $n-1$. This is the same as the degrees of freedom. This is no accident. The reason the denominator and the degrees of freedom are both $n-1$ comes from how the standard deviation is calculated. Remember, first you take each data value and subtract $\bar{x}$. If you add up all of these new values, you will get 0. This must happen. Since it must happen, the first $n-1$ data values you have "freedom of choice", but the nth data value, you have no freedom to choose. Hence, you have $n-1$ degrees of freedom. Another way to think about it is that if you five people and five chairs, the first four people have a choice of where they are sitting, but the last person does not. They have no freedom of where to sit. Only $5-1=4$ people have freedom of choice.

The Student's t-distribution is a bell-shape that is more spread out than the normal distribution. There are many t-distributions, one for each different degree of freedom.

Here is a graph of the normal distribution and the Student's t-distribution for df = 1 and df = 2.

Figure #7.3.1: Typical Student t-Distributions



As the degrees of freedom increases, the student's t-distribution looks more like the normal distribution.

To find probabilities for the t-distribution, again technology can do this for you. There are many technologies out there that you can use. On the TI-83/84, the command is in the DISTR menu and is tcdf(. The syntax for this command is

$$\text{tcdf}\left(\text{lower limit, upper limit, } df\right)$$

Hypothesis Test for One Population Mean (t-Test)

1. State the random variable and the parameter in words.
   x = random variable
   μ = mean of random variable

2. State the null and alternative hypotheses and the level of significance
   $H_o : \mu = \mu_o$, where μ_o is the known proportion
   $H_A : \mu < \mu_o$
   $H_A : \mu > \mu_o$, use the appropriate one for your problem
   $H_A : \mu \neq \mu_o$
   Also, state your α level here.

3. State and check the assumptions for a hypothesis test
   a. A random sample of size n is taken.
   b. The population of the random variable is normally distributed, though the t-test is fairly robust to the condition if the sample size is large. This means that if this condition isn't met, but your sample size is quite large (over 30), then the results of the t-test are valid.
   c. The population standard deviation, σ, is unknown.

4. Find the sample statistic, test statistic, and p-value

   Test Statistic:

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

   with degrees of freedom = $df = n - 1$

   p-value:

   Use $\text{tcdf}(\text{lower limit, upper limit, } df)$

   (Note: if $H_A : \mu < \mu_o$, then lower limit is $-1E99$ and upper limit is your test statistic. If $H_A : \mu > \mu_o$, then lower limit is your test statistic and the upper limit is $1E99$. If $H_A : \mu \neq \mu_o$, then find the p-value for $H_A : \mu < \mu_o$, and multiply by 2.)

5. Conclusion

   This is where you write reject H_o or fail to reject H_o. The rule is: if the p-value $< \alpha$, then reject H_o. If the p-value $\geq \alpha$, then fail to reject H_o

6. Interpretation

   This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show H_A is true, or you do not have enough evidence to show H_A is true.

How to check the assumptions of t-test:

In order for the t-test to be valid, the assumptions of the test must be true. Whenever you run a t-test, you must make sure the assumptions are true. You need to check them. Here is how you do this:

1. For the condition that the sample is a random sample, describe how you took the sample. Make sure your sampling technique is random.
2. For the condition that population of the random variable is normal, remember the process of assessing normality from chapter 6.

Note: if the assumptions behind this test are not valid, then the conclusions you make from the test are not valid. If you do not have a random sample, that is your fault. Make sure the sample you take is as random as you can make it following sampling techniques from chapter 1. If the population of the random variable is not normal, then take a sample larger than 30. If you cannot afford to do that, or if it is not logistically possible, then you do different tests called non-parametric tests. There is an entire course on non-parametric tests, and they will not be discussed in this book.

Example #7.3.1: Test of the Mean Using the Formula

A random sample of 20 IQ scores of famous people was taken from the website of IQ of Famous People ("IQ of famous," 2013) and a random number generator was used to pick 20 of them. The data are in table #7.3.1. Do the data provide evidence at the 5% level that the IQ of a famous person is higher than the average IQ of 100?

Table #7.3.1: IQ Scores of Famous People

158	180	150	137	109
225	122	138	145	180
118	118	126	140	165
150	170	105	154	118

Solution:

1. State the random variable and the parameter in words.
   x = IQ score of a famous person
   μ = mean IQ score of a famous person

2. State the null and alternative hypotheses and the level of significance
   $H_o : \mu = 100$
   $H_A : \mu > 100$
   $\alpha = 0.05$

3. State and check the assumptions for a hypothesis test
   a. A random sample of 20 IQ scores was taken, since that was said in the problem.
   b. The population of IQ score is normally distributed as was shown in example #6.4.2.

4. Find the sample statistic, test statistic, and p-value
   Sample Statistic:
   $\bar{x} = 145.4$
   $s \approx 29.27$
   Test Statistic:
   $$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{145.4 - 100}{\frac{29.27}{\sqrt{20}}} \approx 6.937$$
   p-value:
   $df = n - 1 = 20 - 1 = 19$
   p-value = $\text{tcdf}(6.937, 1E99, 19) = 6.5 \times 10^{-7}$

5. Conclusion
   Since the p-value is less than 5%, then reject H_o.

6. Interpretation
   There is enough evidence to show that famous people have a higher IQ than the average IQ of 100.

Example #7.3.2: Test of the Mean Using the TI-83/84 Calculator

In 2011, the average life expectancy for a woman in Europe was 79.8 years. The data in table #7.3.2 are the life expectancies for men in European countries in 2011 ("WHO life expectancy," 2013). Do the data indicate that men's life expectancy is less than women's? Test at the 1% level.

Table #7.3.2: Life Expectancies for Men in European Countries in 2011

73	79	67	78	69	66	78	74
71	74	79	75	77	71	78	78
68	78	78	71	81	79	80	80
62	65	69	68	79	79	79	73
79	79	72	77	67	70	63	82
72	72	77	79	80	80	67	73
73	60	65	79	66			

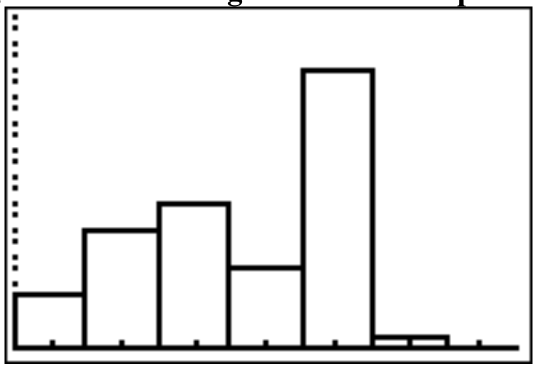
Solution:

1. State the random variable and the parameter in words.
   x = life expectancy for a European man in 2011
   μ = mean life expectancy for European men in 2011

2. State the null and alternative hypotheses and the level of significance
   $H_o : \mu = 79.8$ years
   $H_A : \mu < 79.8$ years
   $\alpha = 0.01$

3. State and check the assumptions for a hypothesis test
   a. A random sample of 53 life expectancies of European men in 2011 was taken. The data is actually all of the life expectancies for every country that is considered part of Europe by the World Health Organization. However, the information is still sample information since it is only for one year that the data was collected. It may not be a random sample, but that is probably not an issue in this case.
   b. The distribution of life expectancies of European men in 2011 is normally distributed. To see if this condition has been met, look at the histogram, number of outliers, and the normal probability plot. (If you wish, you can look at the normal probability plot first. If it doesn't look linear, then you may want to look at the histogram and number of outliers at this point.)

Figure #7.3.2: Histogram for Life Expectancies of European Men in 2011



Not normally distributed

Number of outliers:

$IQR = 79 - 69 = 10$

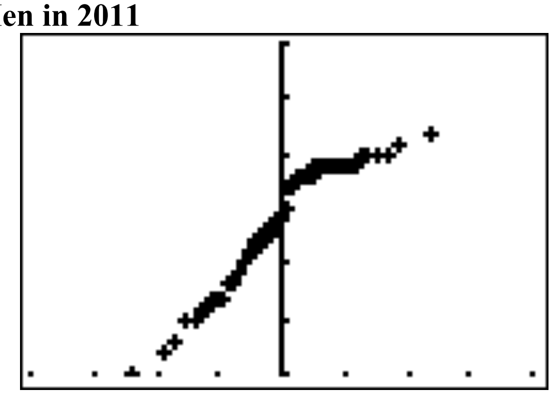
$1.5 * IQR = 15$

$Q1 - 1.5 * IQR = 69 - 15 = 54$

$Q3 + 1.5 * IQR = 79 + 15 = 94$

Outliers are numbers below 54 and above 94. There are no outliers for this data set.

Figure #7.3.3: Normal Probability Plot for Life Expectancies of European Men in 2011



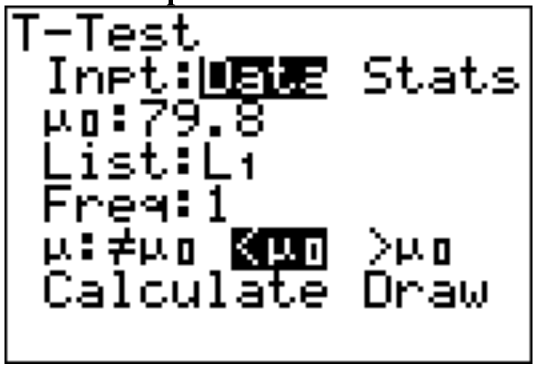
Not linear

This population does not appear to be normally distributed. This sample is larger than 30, so it is good that the t-test is robust.

4. Find the sample statistic, test statistic, and p-value

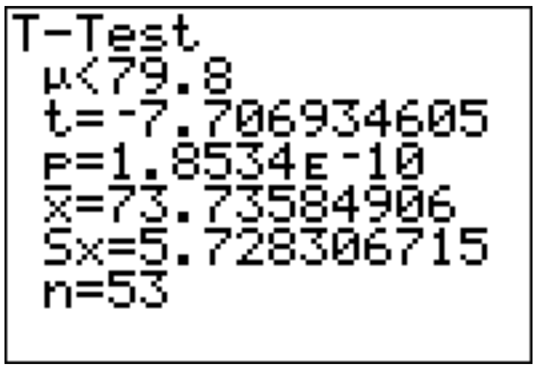
   The calculations will be conducted on the TI-83/84 calculator. Go into STAT and type the data into L1. Then go into STAT and move over to TESTS. Choose T-Test. The setup for the calculator is in figure #7.3.4.

Figure #7.3.4: Setup for T-Test on TI-83/84 Calculator



Once you press ENTER on Calculate you will see the result shown in figure #7.3.5.

Figure #7.3.5: Result of T-Test on TI-83/84 Calculator



The $t = -7.707$ is the test statistic. The $p = 1.8534E - 10$ is the p-value which is actually 1.8534×10^{-10}. You also are given the sample mean, sample standard deviation, and sample size.

5. Conclusion
   Since the p-value is less than 1%, then reject H_o.

6. Interpretation
   There is enough evidence to show that the mean life expectancy for European men in 2011 was less than the mean life expectancy for European women in 2011 of 79.8 years.

Section 7.3: Homework

In each problem show all steps of the hypothesis test. If some of the assumptions are not met, note that the results of the test may not be correct and then continue the process of the hypothesis test.

1.) The Kyoto Protocol was signed in 1997, and required countries to start reducing their carbon emissions. The protocol became enforceable in February 2005. In 2004, the mean CO_2 emission was 4.87 metric tons per capita. Table 7.3.3 contains a random sample of CO_2 emissions in 2010 ("CO2 emissions," 2013). Is there enough evidence to show that the mean CO_2 emission is lower in 2010 than in 2004? Test at the 1% level.

Table #7.3.3: CO_2 Emissions (in metric tons per capita) in 2010

1.36	1.42	5.93	5.36	0.06	9.11	7.32
7.93	6.72	0.78	1.80	0.20	2.27	0.28
5.86	3.46	1.46	0.14	2.62	0.79	7.48
0.86	7.84	2.87	2.45			

2.) The amount of sugar in a Krispy Kream glazed donut is 10 g. Many people feel that cereal is a healthier alternative for children over glazed donuts. Table #7.3.4 contains the amount of sugar in a sample of cereal that is geared towards children ("Healthy breakfast story," 2013). Is there enough evidence to show that the mean amount of sugar in children's cereal is more than in a glazed donut? Test at the 5% level.

Table #7.3.4: Sugar Amounts in Children's Cereal

10	14	12	9	13	13	13
11	12	15	9	10	11	3
6	12	15	12	12		

3.) The FDA regulates that fish that is consumed is allowed to contain 1.0 mg/kg of mercury. In Florida, bass fish were collected in 53 different lakes to measure the amount of mercury in the fish. The data for the average amount of mercury in each lake is in table #7.3.5 ("Multi-disciplinary niser activity," 2013). Do the data provide enough evidence to show that the fish in Florida lakes has more mercury than the allowable amount? Test at the 10% level.

Table #7.3.5: Average Mercury Levels (mg/kg) in Fish

1.23	1.33	0.04	0.44	1.20	0.27
0.48	0.19	0.83	0.81	0.71	0.5
0.49	1.16	0.05	0.15	0.19	0.77
1.08	0.98	0.63	0.56	0.41	0.73
0.34	0.59	0.34	0.84	0.50	0.34
0.28	0.34	0.87	0.56	0.17	0.18
0.19	0.04	0.49	1.10	0.16	0.10
0.48	0.21	0.86	0.52	0.65	0.27
0.94	0.40	0.43	0.25	0.27	

4.) Stephen Stigler determined in 1977 that the speed of light is 299,710.5 km/sec. In 1882, Albert Michelson had collected measurements on the speed of light ("Student t-distribution," 2013). His measurements are given in table #7.3.6. Is there evidence to show that Michelson's data is different from Stigler's value of the speed of light? Test at the 5% level.

Table #7.3.6: Speed of Light Measurements in (km/sec)

299883	299816	299778	299796	299682
299711	299611	299599	300051	299781
299578	299796	299774	299820	299772
299696	299573	299748	299748	299797
299851	299809	299723		

5.) Table #7.3.7 contains pulse rates after running for 1 minute, collected from females who drink alcohol ("Pulse rates before," 2013). The mean pulse rate after running for 1 minute of females who do not drink is 97 beats per minute. Do the data show that the mean pulse rate of females who do drink alcohol is higher than the mean pulse rate of females who do not drink? Test at the 5% level.

Table #7.3.7: Pulse Rates of Woman Who Use Alcohol

176	150	150	115	129	160
120	125	89	132	120	120
68	87	88	72	77	84
92	80	60	67	59	64
88	74	68			

6.) The economic dynamism, which is the index of productive growth in dollars for countries that are designated by the World Bank as middle-income are in table #7.3.8 ("SOCR data 2008," 2013). Countries that are considered high-income have a mean economic dynamism of 60.29. Do the data show that the mean economic dynamism of middle-income countries is less than the mean for high-income countries? Test at the 5% level.

Table #7.3.8: Economic Dynamism of Middle Income Countries

25.8057	37.4511	51.915	43.6952	47.8506	43.7178	58.0767
41.1648	38.0793	37.7251	39.6553	42.0265	48.6159	43.8555
49.1361	61.9281	41.9543	44.9346	46.0521	48.3652	43.6252
50.9866	59.1724	39.6282	33.6074	21.6643		

7.) In 1999, the average percentage of women who received prenatal care per country is 80.1%. Table #7.3.9 contains the percentage of woman receiving prenatal care in 2009 for a sample of countries ("Pregnant woman receiving," 2013). Do the data show that the average percentage of women receiving prenatal care in 2009 is higher than in 1999? Test at the 5% level.

Table #7.3.9: Percentage of Woman Receiving Prenatal Care

70.08	72.73	74.52	75.79	76.28	76.28
76.65	80.34	80.60	81.90	86.30	87.70
87.76	88.40	90.70	91.50	91.80	92.10
92.20	92.41	92.47	93.00	93.20	93.40
93.63	93.68	93.80	94.30	94.51	95.00
95.80	95.80	96.23	96.24	97.30	97.90
97.95	98.20	99.00	99.00	99.10	99.10
100.00	100.00	100.00	100.00	100.00	

8.) Maintaining your balance may get harder as you grow older. A study was conducted to see how steady the elderly is on their feet. They had the subjects stand on a force platform and have them react to a noise. The force platform then measured how much they swayed forward and backward, and the data is in table #7.3.10 ("Maintaining balance while," 2013). Do the data show that the elderly sway more than the mean forward sway of younger people, which is 18.125 mm? Test at the 5% level.

Table #7.3.10: Forward/backward Sway (in mm) of Elderly Subjects

19	30	20	19	29	25	21	24	50

Data Sources:
Australian Human Rights Commission, (1996). *Indigenous deaths in custody 1989 - 1996*. Retrieved from website: http://www.humanrights.gov.au/publications/indigenous-deaths-custody

CDC features - new data on autism spectrum disorders. (2013, November 26). Retrieved from http://www.cdc.gov/features/countingautism/

Center for Disease Control and Prevention, Prevalence of Autism Spectrum Disorders - Autism and Developmental Disabilities Monitoring Network. (2008). *Autism and developmental disabilities monitoring network-2012*. Retrieved from website: http://www.cdc.gov/ncbddd/autism/documents/ADDM-2012-Community-Report.pdf

CO2 emissions. (2013, November 19). Retrieved from http://data.worldbank.org/indicator/EN.ATM.CO2E.PC

Federal Trade Commission, (2008). *Consumer fraud and identity theft complaint data: January-December 2007*. Retrieved from website: http://www.ftc.gov/opa/2008/02/fraud.pdf

Gallup news service. (2013, November 7-10). Retrieved from http://www.gallup.com/file/poll/165896/JFK_Conspiracy_131115.pdf

Healthy breakfast story. (2013, November 16). Retrieved from http://lib.stat.cmu.edu/DASL/Stories/HealthyBreakfast.html

IQ of famous people. (2013, November 13). Retrieved from http://www.kidsiqtestcenter.com/IQ-famous-people.html

Maintaining balance while concentrating. (2013, September 25). Retrieved from http://www.statsci.org/data/general/balaconc.html

Morgan Gallup poll on unemployment. (2013, September 26). Retrieved from http://www.statsci.org/data/oz/gallup.html

Multi-disciplinary niser activity - mercury in bass. (2013, November 16). Retrieved from http://gozips.uakron.edu/~nmimoto/pages/datasets/MercuryInBass - description.txt

Pregnant woman receiving prenatal care. (2013, October 14). Retrieved from http://data.worldbank.org/indicator/SH.STA.ANVC.ZS

Pulse rates before and after exercise. (2013, September 25). Retrieved from http://www.statsci.org/data/oz/ms212.html

SOCR data 2008 world countries rankings. (2013, November 16). Retrieved from http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data_2008_World_CountriesRankings

Student t-distribution. (2013, November 25). Retrieved from
http://lib.stat.cmu.edu/DASL/Stories/student.html

WHO life expectancy. (2013, September 19). Retrieved from
http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends/en/index.html