Lesson 11
Hypothesis Tests and Confidence Intervals for Means

Outline of the Lesson

As we have discussed for several lessons now, *inferential statistics* has two important features:

- Information is obtained from a *sample*.
- The information from the sample is used to draw a conclusion (an *inference*) about the entire *population* from which the sample was drawn.

Up to this point we have dealt exclusively with categorical variables. In the current lesson we will begin the process of learning about inferential statistics for numerical variables. We begin with the simplest situation: we have a single population we are interested in studying, and we have a numerical variable that we wish to study for that population. We will study that variable by choosing a random sample from the population, and recording the variable's values for the individuals in that sample. For example, the population of interest might consist of all Pennsylvania adults, and the variable might be the amount of state income tax paid for the tax year 2010.

When a variable is categorical, the primary method used to summarize the data for an entire population, or for an entire sample, is to calculate a proportion. Thus, in Lessons 6 to 8 you learned how to use the sample proportion to make inferences about the corresponding population proportion. Recall that the sample proportion is an example of a *statistic* – a fact about the sample – and that the population proportion is an example of a *parameter* – a fact about the entire population.

For a numerical variable, there are several possible statistics/parameters we could study. For example, the mean, median, IQR, and standard deviation are various ways in which we summarize the results from a numerical variable. In this course, we will focus on the *mean* as the statistic/parameter of interest. So, the parameter of interest might be the mean state income tax paid by all Pennsylvania adults in the tax year 2018. To make inferences about that parameter, we would choose a random sample of Pennsylvania adults, and calculate the mean state income tax for the people in that sample.

In some ways, this lesson should be one of the easier lessons for you to master. For a categorical variable, you already know how to use a sample proportion to carry out a hypothesis test for a population proportion, and to calculate a confidence interval for a population proportion. The methods we use for a numerical variable are nearly identical to those you already know. The one complicating factor is that we have to do some thinking, before we begin, in the following form: *Is this an analysis of a categorical variable, or is it an analysis of a numerical variable*? Once we make the correct decision about this issue, we will know how to proceed. If the variable is categorical, we will calculate a *proportion* for our sample and use it to make statements about the population *proportion*. If the variable is numerical, we will calculate a *mean* for our sample and use it to make statements about the population *mean*.

11.1 – Hypothesis Tests for a Population Mean

Review of hypothesis tests for proportions

Because the new methods are very similar to the methods we already know, it will be good to do a very quick review of what we already know. Here, then, are some key ideas related to carrying out a hypothesis test for a population proportion.

- We are studying a single population, and the variable we want to study is categorical.

- We want to test a claim about *p*, the proportion of "success" in the population.

- We describe the claim to be tested by writing null and alternative hypotheses. The null hypothesis is always a claim that the population proportion is equal to some particular value, written symbolically as:

  $$H_0: p = p_0$$

  The alternative hypothesis states that the null hypothesis is false. For most situations, we simply state that the population proportion is *not* equal to the claimed value in the null hypothesis, writing:

  $$H_a: p \neq p_0$$

  This is a two-tail hypothesis test. Today's statisticians almost always use two-tail tests.

  In a few situations, when the researcher has a preconceived notion that the population proportion is larger than the claimed proportion, we will use a one-tail (right tail) test, writing the alternative hypothesis as:

  $$H_a: p > p_0$$

  Similarly, we may do a one-tail (left tail) test if there is a preconceived notion that the population proportion is less than the claimed proportion, writing the alternative hypothesis as:

  $$H_a: p < p_0$$

- We use the proportion from our sample, written as $\hat{p}$, to calculate a test statistic and *P*-value. For a one-sample proportion hypothesis test, the sampling distribution is approximately normal, so the test statistic is a *z*-score. The *P*-value is based on areas in the normal distribution.

- If the *P*-value is less than α (the chosen significance level), we *reject* the null hypothesis, and conclude that there *is* enough evidence to support the alternative hypothesis.
  If the *P*-values is not less than α, we *fail to reject* the null hypothesis, and conclude that there *is not* enough evidence to support the alternative hypothesis.

- The significance level α is chosen to limit the likelihood of Type 1 errors (rejecting a true null hypothesis).

- The calculations involve three steps, although of course it is possible to combine the steps. First, we calculate the standard error using the formula

$$se = \sqrt{\frac{p_0(1-p_0)}{n}}$$

We then use this to calculate the test statistic:

$$z = \frac{\hat{p}-p_0}{se}$$

In words, we calculate the difference between the proportion in the sample and the proportion claimed by the null hypothesis, divided by the standard error.
We do a left-tail, right-tail, or two-tail *P*-value calculation depending on the $<$, $>$, or $\neq$ in the alternative hypothesis. Although we can use Table A for this calculation, it is preferable to use technology such as the calculator supplied with these lessons to calculate the *p*-value.

Exercise 1 (Review): Carry out a two-tail test to examine the claim that the population proportion is 43%, given that your sample of size 956 has a sample proportion of 47.5%.

Hypothesis tests for means

Our description of carrying out a hypothesis test involving a population mean is written in exactly the same form as our earlier review of hypothesis tests for population proportions. We do this in order to emphasis how very similar the two methods are.

- We are studying a single population, and the variable we want to study is *numerical*.

- We want to test a claim about the *mean* (*average*) for the variable for the entire population. Recall that the commonly used symbol for a population mean is μ.

- We describe the claim to be tested by writing null and alternative hypotheses. The null hypothesis is always a claim that the population mean is equal to some particular value. Keeping in mind that μ is the symbol for a population mean, we can write the null hypothesis symbolically as:

$$H_0: \mu = \mu_0$$

- The alternative hypothesis states that the null hypothesis is false. For most situations, we simply state that the population mean is *not* equal to the claimed value in the null hypothesis, writing:

$$H_a: \mu \neq \mu_0$$

  This is a two-tail hypothesis test. Today's statisticians almost always use two-tail tests.
  In a few situations, when the researcher has a preconceived notion that the population mean is larger than the claimed mean, we will use a one-tail (right tail) test, writing the alternative hypothesis as:

$$H_a: \mu > \mu_0$$

  Similarly, we may do a one-tail (left tail) test if there is a preconceived notion that the population mean is less than the claimed mean, writing the alternative hypothesis as:

$$H_a: \mu < \mu_0$$

- We use the mean and standard deviation from our sample, written as $\bar{x}$ and *s* respectively, to calculate a test statistic and *P*-value. As we know, for proportions we use a z score and areas in a normal distribution for these calculations. When working with means, we must use a distribution known as a *t* distribution. There are many different *t* distributions, and the one we use depends on the size of our sample. All the distributions are mound-shaped, very similar to the normal distribution, but with more data in the tails than for the normal distribution. The particular distributions are identified by what is called the **degrees of freedom** (abbreviated **df**). For large values of **df**, the distribution is very close to the normal distribution. For a hypothesis test for a population mean, the distribution we use is the one given by the formula $df = n - 1$, where *n* is the size of the sample.
  To summarize, for a one-sample mean hypothesis test the test statistic is a *t*-score. The *P*-value is based on areas in an appropriate *t* distribution.

- If the *P*-value is less than α (the chosen significance level), we *reject* the null hypothesis, and conclude that there *is* enough evidence to support the alternative hypothesis.
  If the *P*-values is not less than α, we *fail to reject* the null hypothesis, and conclude that there *is not* enough evidence to support the alternative hypothesis.

- The significance level α is chosen to limit the likelihood of Type 1 errors (rejecting a true null hypothesis).

- The calculations involve three steps, although of course it is possible to combine the steps. First, we calculate the standard error using the formula

$$se = \frac{s}{\sqrt{n}}$$

In this formula, *s* is the standard deviation of the sample. We then use this to calculate the test statistic:

$$t = \frac{\bar{x} - \mu_0}{se}$$

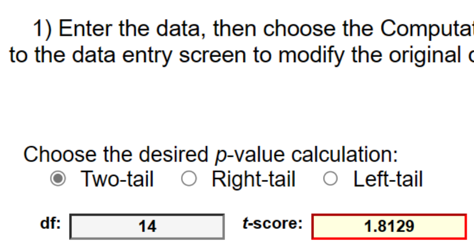
In words, we calculate the difference between the mean in the sample and the mean claimed by the null hypothesis, divided by the standard error.

We do a left-tail, right-tail, or two-tail *P*-value calculation depending on the <, >, or ≠ in the alternative hypothesis. The calculation is based on the particular *t* distribution whose degrees of freedom (df) is given by the formula $df = n - 1$.

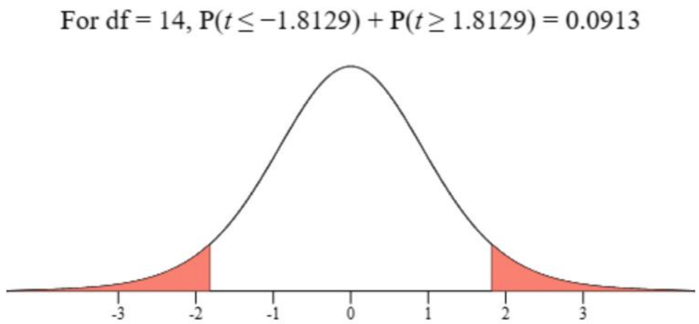
For the normal distribution, we were able to use Table A to calculate the *P*-value, although we rather quickly switched to using technology. For our situation, there is a table (Table B) from which we could determine a possible range of *P*-values, but not the actual *P*-value. However, it is far superior to use technology. Using the calculator supplied with these lessons, we use menu option *Distributions*, submenu *t distribution p-value*, as illustrated in the following example.

Example. To illustrate the mechanics of the calculations, we will carry out a two tail test to investigate the claim that the population mean is 65. We take a sample of size 15 and obtain a mean of 67.2, with a standard deviation of 4.7, for the variable we are measuring. Here are the steps:

1. $H_0: \mu = 65$
   $H_a: \mu \neq 65$
2. $se = \frac{s}{\sqrt{n}} = \frac{4.7}{\sqrt{15}} = 1.2135$
3. $t = \frac{67.2 - 65}{1.2135} = 1.8129$
4. Because *n* is 15, we use $df = n - 1 = 14$ when we calculate the *p*-value. Using menu option *Distributions*, submenu *t distribution p-value*, we enter our *t*-score and *df* as shown here:



5. When we press the *Computations* button we obtain these results:



The two-tail *p*-value is 0.0913.

6. Using either $\alpha = 0.05$ or $\alpha = 0.01$, we do not reject the null hypothesis. There is not enough evidence to conclude that the population mean is not 65. (This does not imply that the mean definitely *is* 65, only that it *might be* 65 – there is not enough evidence to say that it *isn't* 65.)

Exercise 2: Carry out a left-tail test to examine the claim that the population mean is 514, given that your sample of size 56 has a sample mean of 490 with a standard deviation of 98.

The applets at the following links provide additional practice. The first applet includes all the calculation steps. The second is identical to the first, except that the calculations are done for you.

Mean hypothesis test - calculations

Mean hypothesis test - conclusions

Choosing hypotheses and interpreting results in the context of the problem

Here is a further example which contains a real-world context, illustrating the type of situation for which these methods are applicable.

Example. In a certain municipality which collects an income tax from its residents, the average tax paid last year was $342.11, As part of its budgeting process, the governing body assumes the average will be the same for the current tax year. A statistician worries that this may overestimate the tax revenue for the current year. She randomly samples 27 taxpayers in the municipality, and calculates the tax they will be paying this year. The sample yields a mean of $329.37 with a standard deviation of $40.17. Carry out an appropriate hypothesis test, and report on the results using significance level $\alpha = 0.05$.

Step 1 - hypotheses. The general guidelines for choosing the null hypothesis are the same as we have encountered before. The null hypothesis will always be a statement that the population mean is equal to some particular value. Also, the null hypothesis is generally a statement of "no difference" or "no change." In this case, the null hypothesis states that the population mean this year (the mean for the entire municipality) will be no different from the mean last year. In symbols:

$H_0: \mu = 342.11$

In general, the alternative hypothesis will simply state that the population mean is not 342.11. However, in this example, the statistician has a preconceived notion (or at least a worry) that the mean will be lower this year. This is a situation where a one-tail test might be indicated – the statistician would want to report a study that indicates a lower mean (hence a revenue shortfall) but if the study indicates a higher mean that will be of no concern. The statistician would likely use a left-tail test, with this alternative hypothesis:

$H_a: \mu < 342.11$

Notice that this decision happens *before* taking the sample, so it is not influenced by what the sample mean turns out to be.

Step 2 – calculations. By now, if you have used the applets to practice, the calculations should be relatively routine.

$se = \frac{s}{\sqrt{n}} = \frac{40.17}{\sqrt{27}} = 7.7307$

$$t = \frac{329.37 - 342.11}{7.7307} = -1.6480$$

p-value using technology = 0.0557

Step 3 – conclusions. Since 0.0557 is not less than 0.05, we fail to reject the null hypothesis. As we learned in the lesson on proportion hypothesis tests, in general when we reject a null hypothesis we report something along these lines:

At significance level _____ there was enough evidence to conclude that <u>*description of alternative hypothesis* .</u>

If we do not reject the null hypothesis, we simply negate this to write:

At significance level _____ there was ***not*** *enough evidence to conclude that* <u>*description of alternative hypothesis* .</u>

Using this template, we can describe our results as: At significance level $\alpha = 0.05$, there was not enough evidence to conclude that the mean tax paid this year will be less than the $342.11 reported last year.

Alternatively, we can use the word "significant" to write our conclusion. The general format involves describing the alternative hypothesis using the word "significant" if we rejected the null hypothesis and "not significant" if we did not reject the null hypothesis. In this example, we might write, "The mean tax paid this year will not be significantly less than the $342.11 average last year." Keep in mind, however, that the word significant in this context does not mean "large," it simply indicates whether or not we rejected the null hypothesis.

Exercise 3: For a certain brand of soft drink, the bottles are supposed to (on average) contain 16 ounces. Quality control routinely conducts sampling to ensure the proper functioning of the machines that fill the bottles. In a recent sample of 85 bottles, the bottles had an average content of 15.87 ounces with a standard deviation of 0.59 ounces.

- **a.** Write the null and alternative hypotheses for a suitable hypothesis test.
- **b.** The results of the calculations for the hypothesis test are $t = -2.0314$, p-value = 0.0454. Using $\alpha = 0.01$, write the conclusion in the context of the problem.
- **c.** With these results, what should quality control do?

The applet at this link gives additional practice in interpreting various population mean hypothesis test situations.

Mean hypothesis test - interpretation

11.2 – Confidence Intervals for a Population Mean

Review of confidence intervals for proportions

Just as we did for hypothesis tests, we begin with a review of the method used to calculate a confidence interval for a population proportion. Then we will use that as a way to discuss how working with means differs from working with proportions.

- We are studying a single population, and the variable we want to study is categorical.
- We want to estimate p, the proportion of "success" in the population. For example, we might want to estimate the proportion of all introductory statistics students who would answer yes to the question, "Have you ever smoked more than two packs of cigarettes a day?"
- For the point estimate we use the proportion from our sample, written as $\hat{p}$.
- The interval estimate is called a *confidence interval*. It is the point estimate, plus or minus the margin of error. The margin of error is chosen in order to provide a certain confidence level. For example, if we calculate a 95% confidence interval, this means we have used a method which yields a correct answer for 95% of the possible samples from that population. (An answer is correct if the calculated interval does contain the actual population proportion.)
- The margin of error is $z^* \cdot se$. In this calculation, z^* is based on the fact that the sampling distribution for the sample proportions is approximately normal, and se is the standard error (standard deviation) of that sampling distribution.
- The calculations are done in three steps, although of course those steps can be combined. First, we calculate the standard error using the formula

$$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Next, we calculate the margin of error as

$$m.e. = z^* \cdot se$$

The z* can be determined using Table B – the bottom line, labeled z^* (or, in some texts, labeled ∞). Finally, we calculate the interval as

$$(\hat{p} - m.e., \hat{p} + m.e.)$$

- The interval indicates a range of values that we are confident the population proportion lies within. Recall that we use the word "plausible" in this connection. The interval contains the plausible values for the population proportion.

Exercise 4 (Review): Calculate a 99% confidence interval for a population proportion if the proportion for the sample is 65% and the sample size is 1500.

Confidence intervals for means

Our description of calculating a confidence interval for a population mean is written in exactly the same form as our earlier review of confidence intervals for population proportions. We do this in order to emphasis how very similar the two methods are.

- We are studying a single population, and the variable we want to study is numerical.

- We want to estimate μ, the mean of that variable for the entire population. For example, we might want to estimate the average (mean) length of mature great white sharks.

- For the point estimate we use the mean from our sample, written as $\bar{\boldsymbol{x}}$.

- The interval estimate is called a *confidence interval*. It is the point estimate, plus or minus the margin of error. The margin of error is chosen in order to provide a certain confidence level. For example, if we calculate a 95% confidence interval, this means we have used a method which yields a correct answer for 95% of the possible samples from that population. (An answer is correct if the calculated interval does contain the actual population mean.)

- For proportions, the margin of error is $z^* \cdot se$, where *se* is the standard error of the sampling distribution, and z^* is based on the fact that the sampling distribution for the sample proportions is approximately normal. For confidence intervals for means, there is a subtle difference in calculating the margin of error. Just as was the case for hypothesis tests, we cannot use the normal distribution, but instead must use a *t* distribution. There are many different *t* distributions, and the one we use depends on the size of our sample. All the distributions are mound-shaped, very similar to the normal distribution, but with more data in the tails than for the normal distribution. The particular distributions are identified by what is called the **degrees of freedom** (abbreviated **df**). For large values of **df**, the distribution is very close to the normal distribution. For a confidence interval for a population mean, the distribution we use is the one given by the formula $df = n - 1$, where *n* is the size of the sample.

We use the notation t^* rather than z^* to indicate that the value comes from a *t* distribution rather than the normal distribution. So the margin of error is $t^* \cdot se$ rather than $z^* \cdot se$. The method for determining the correct t^* value will be described shortly.

- The calculations are done in three steps, although of course those steps can be combined. First, we calculate the standard error using the formula

  $$se = \frac{s}{\sqrt{n}}$$

  In this formula, *s* is the standard deviation of the sample. (This is the same formula we used for hypothesis tests.) Next, we calculate the margin of error as

  $$m.e. = t^* \cdot se$$

  The t^* comes from Table B. Which row of the table you use depends on your sample size. The rows of the table are labeled with the *df* column (*df* stands for "degrees of freedom"). For the method we are studying here, *df* is always one less than the sample size. (If there is no entry in the table for your *df*, use the next smaller *df* that is in the table.)

  Finally, we calculate the interval as

  $$(\bar{x} - m.e., \bar{x} + m.e.)$$

- The interval indicates a range of values that we are confident the population mean lies within. Recall that we use the word "plausible" in this connection. The interval contains the plausible values for the population mean.

CAUTION: This is a source of frequent confusion for many students. We are making a statement about the *mean* (the *average*) value of the variable. We are NOT making a statement about any of the individual values of the variable. For example, if we say that the average GPA at the university lies in the interval (2.73, 3.13), this does not in any way imply that your GPA or that of your friend is within that interval. In fact, it is quite likely that a lot of students have GPAs that are not in the interval. All we are saying is that if we took the average for the entire student body, the answer would lie in that interval.

Example. To illustrate the mechanics of the calculations, suppose we take a sample of size 15 and obtain a mean of 67.2, with a standard deviation of 4.7, for the variable we are measuring. Here are the calculations for a 95% confidence interval:

1. $\bar{x} = 67.2$

2. $se = \frac{s}{\sqrt{n}} = \frac{4.7}{\sqrt{15}} = 1.2135$

3. Because *n* is 15, we use *df* = 14. In Table B, in the *df* = 14 row and the confidence level = 95% column, we find the entry 2.145. Here is a snapshot of the pertinent portion of the table:

Table B – *t* distribution critical values

	Confidence Level							
	80%	90%	95%	96%	98%	99%	99.5%	99.9%
	Right-tail Probability							
df	0.10	0.05	0.025	0.02	0.01	0.005	0.0025	0.0005
1	3.078	6.314	12.706	15.895	31.821	63.657	127.321	636.619
2	1.886	2.920	4.303	4.849	6.965	9.925	14.089	31.599
3	1.638	2.353	3.182	3.482	4.541	5.841	7.453	12.924
12	1.356	1.782	2.179	2.303	2.681	3.055	3.428	4.318
13	1.350	1.771	2.160	2.282	2.650	3.012	3.372	4.221
14	1.345	1.761	2.145	2.264	2.624	2.977	3.326	4.140
15	1.341	1.753	2.131	2.249	2.602	2.947	3.286	4.073

Therefore, we use $t^* = 2.145$.

4. The margin of error is 2.145 * 1.2135 = 2.6030

5. The confidence interval, with the final answer rounded to the nearest tenth, is (67.2 – 2.6030, 67.2 + 2.6030) = (64.6, 69.8).

Example: Based on the interval we just calculated, which if any of these are plausible values for the population mean? 66, 69.7, 70

Solution: Both 66 and 69.7 are values within the interval, so both these values are plausible. On the other hand, although 70 is very close to the upper end of the interval, it is not within the interval, so *based on this interval it is not plausible that the population mean is 70.*

Exercise 5: Calculate a 99% confidence interval for a population mean for each situation described. Round your final answer to the nearest whole number. For each situation, is it plausible that the population mean is 450?

a. sample size 28, mean for the sample is 497, standard deviation for the sample is 93

b. sample size 55, mean for the sample is 497, standard deviation for the sample is 93 (**Reminder**: if the table does not contain a row for the desired *df* value, use the next smaller *df* that is in the table.)

The applets at the following links provide additional practice. The first applet includes all the calculation steps. The second is identical to the first, except that the calculations are done for you.

Mean confidence interval - calculations

Mean confidence interval - conclusions

Interpreting results in the context of the problem

Here is a further example which contains a real-world context, illustrating the type of situation for which these methods are applicable. We have seen a similar example in the earlier coverage of hypothesis tests for population means.

Example. In a certain municipality which collects an income tax from its residents, the average tax paid last year was $342.11, As part of its budgeting process, the governing body hires a statistician to estimate the tax revenue for the current tax year. She randomly samples 27 taxpayers in the municipality, and calculates the tax they will be paying this year. The sample yields a mean of $329.37 with a standard deviation of $40.17. Calculate and explain the significance of an appropriate 99% confidence interval (rounded to the nearest cent). Is it plausible that the average will be the same as last year?

Solution. We want to estimate the average income tax that will be paid by all the residents this year. As a single number (a point estimate) we simply use the average from the sample – namely, $329.37 per taxpayer. To calculate the corresponding interval estimate (confidence interval), these are the steps:

$$se = \frac{s}{\sqrt{n}} = \frac{40.17}{\sqrt{27}} = 7.7307$$
$$m.e. = t^* \cdot se = 2.779(7.7307) = 21.4836 \text{ (from Table B, confidence level 99\%, df 26)}$$
$$(\bar{x} - m.e., \bar{x} + m.e.) = (329.37 - 21.4836, 329.37 + 21.4836) = (307.89, 350.85)$$

We are 99% confident that, for the population consisting of all taxpayers in that municipality, the average tax paid this year will be somewhere between $307.89 and $350.85). Since last year's average ($342.11) is one of the numbers in this interval, it is plausible that the average will be the same as last year.

Notes:

1. As we pointed out earlier, the interval is a statement about the **mean (average)** for the **entire population**. We are predicting that later this year, when all the tax returns have been filed, if we were to calculate the average tax paid for all those returns that average would lie between $307.89 and $350.85. It is ***not*** a statement about the tax paid by any particular individual.

2. Refer to note 1. For students just learning about confidence intervals for means, here are two ***incorrect*** interpretations which are frequently given:
    - I am 99% confident that all taxpayers this year will pay somewhere between $307.89 and $350.85. This is ***incorrect*** because the interval is a statement about the average, not about the individual taxpayers.
    - I am 99% confident that the average tax this year for the 27 taxpayers is somewhere between $307.89 and $350.85. This is ***incorrect*** because the interval is a statement about the entire population, not about the sample.

3. We have calculated a 99% confidence interval. This means that we have used a method which generates a correct result for 99% of the possible samples from that population.

Exercise 6: For a certain brand of soft drink, the bottles are supposed to (on average) contain 16 ounces. Quality control routinely conducts sampling to ensure the proper functioning of the machines that fill the bottles. In a recent sample of 85 bottles, the bottles had an average content of 15.87 ounces with a standard deviation of 0.59 ounces. The corresponding 99% confidence interval is $(15.701, 16.039)$.

a. Write a sentence explaining the meaning of this interval.

b. Which of the following are correct statements to make based on this interval?

- I am 99% confident that all the bottles contain between 15.701 ounces and 16.039 ounces.
- There is a 99% probability that the next bottle sampled will contain between 15.701 ounces and 16.039 ounces.
- It is plausible that the mean amount of soda being dispensed is only 15.7 ounces.
- It is plausible that the mean amount of soda being dispensed is only 15.8 ounces.
- It is plausible that the mean amount of soda being dispensed is the correct value (namely, 16 ounces).

c. With these results, what should quality control do?

The applet at this link gives additional practice in interpreting various population mean confidence interval situations.

Mean confidence interval - interpretation

11.3 – Summary

One-sample mean – hypothesis test

- We are studying a single population, and the variable we want to study is numerical.
- We want to test a claim about μ, the mean for that variable in the population.
- The null and alternative hypotheses take the form:

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$
(sometimes > or < rather than ≠ in H_a)

- We calculate a test statistic, which in this case is a t-score, in two steps. First, we calculate the standard error:

$$se = \frac{s}{\sqrt{n}}$$

Then, we calculate the t-score:

$$t = \frac{\bar{x} - \mu_0}{se}$$

In words, we calculate the difference between the means for the sample and the null hypothesis, divided by the standard error.

- The *p*-value is based on areas in the *t* distribution, with degrees of freedom one less than the sample size. The *t* distribution looks quite similar to the normal distribution but has more area in the tails. The larger the sample size, the closer it is to the normal distribution.

- We can do a left-tail, right-tail, or two-tail calculation depending on the <, >, or ≠ in the alternative hypothesis.

One-sample mean – confidence interval

- We are studying a single population, and the variable we want to study is numerical.
- We want to estimate μ, the mean for that variable in the population.
- For the point estimate we use the mean from our sample, written as $\bar{x}$.
- The confidence interval is the point estimate, plus or minus the margin of error. We calculate the interval in three steps. First, the standard error, then the margin of error, and finally the interval:

$$se = \frac{s}{\sqrt{n}}$$

$$m.e. = t^* \cdot se$$

$$(\bar{x} - m.e., \bar{x} + m.e.)$$

- The *t** comes from Table B – which line depends upon the sample size (degrees of freedom *df* is one less than the sample size). This is based on the *t* distribution, which looks quite similar to the normal distribution but has more area in the tails. The larger the sample size, the closer it is to the normal distribution.

***z* or *t*?**

The methods we have learned for proportions and those we have learned for means are quite similar. For confidence intervals, we can summarize the steps as follows:

- Calculate the standard error.
- Calculate the margin of error.
- The interval is the value obtained from the sample, plus or minus the margin of error.

For hypothesis tests, we can summarize the steps this way:

- Calculate the standard error.
- Calculate the test statistic, by subtracting the value in the null hypothesis from the value obtained from the sample, then dividing by the standard error.
- Use the test statistic to calculate a *P*-value.
- Draw a conclusion by comparing the *P*-value to the significance level α.

However, they do have this very important difference: *one uses the normal distribution, the other a t distribution.* Put another way, one uses *z* and the other uses *t*.

- When working with proportions, we use the normal distribution.
  - For confidence intervals, the margin of error is calculated as $z^* \cdot se$, where z^* depends on the desired confidence level.
  - For hypothesis tests, the test statistic is a *z* score, and we use the normal distribution to calculate the *P*-value.

- When working with means, we use an appropriate *t* distribution, with degrees of freedom one less than the sample size.
  - For confidence intervals, the margin of error is calculated as $t^* \cdot se$, where t^* depends on the degrees of freedom as well as the confidence level.
  - For hypothesis tests, the test statistic is a *t* score, and we use a *t* distribution to calculate the *P*-value.

Notes:

1. If you happen to know the standard deviation for the entire population, you can use that in the calculation of the standard error. In this case, one can use the normal distribution instead of a *t* distribution to calculate the margin of error or the *p*-value. However, in practice it is very unlikely that you would know the standard deviation for the entire population. Therefore, statisticians nearly universally use the *t* distribution methods we have presented when working with population means.
2. Refer to note 1. The larger the sample size, the closer the *t* distribution comes to matching the normal distribution. As a result, if you accidentally use the normal distribution for a large sample, your results will not be off by much.

11.4– Using Technology to Assist in Calculations

In this section we will show you how to use technology – specifically, the online calculator[1] provided by the author of these lessons – to do the calculations for hypothesis tests and confidence intervals for numerical variables. Here again is a link to that calculator:

Statistical calculator

To illustrate the process, we will use two examples we have already solved "by hand" in Sections 11.1 and 11.2.

Note: The key to determining the appropriate menu option can be summarized as follows:

- First, determine whether you are doing a hypothesis test or a confidence interval. We use a confidence interval to *estimate* the value for the population mean, and a hypothesis test to *examine/test a claim* about the population mean.
- We have a numerical variable, so we are working with means rather than proportions.
- We have data from a single sample – that is, we have "one mean" that we are examining.
- Does the problem as stated provide the actual data for the sample, or does it instead provide summary information (mean, standard deviation, size, …) about that data?

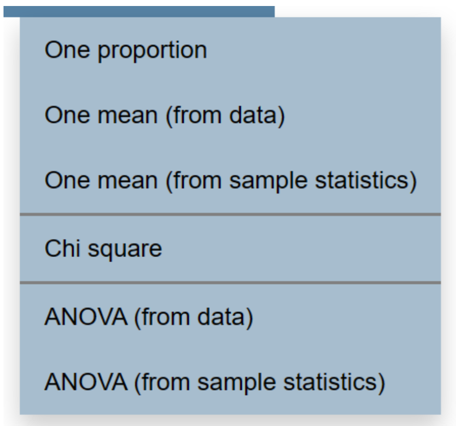
[1] In this section we describe the calculator designed for solving problems such as those typically encountered in an introductory statistics class. In Section 11.6 we will use the datafile-based calculator to analyze numerical variables in a data file.

Example. Carry out a two tail test to investigate the claim that the population mean is 65. We take a sample of size 15 and obtain a mean of 67.2, with a standard deviation of 4.7, for the variable we are measuring.

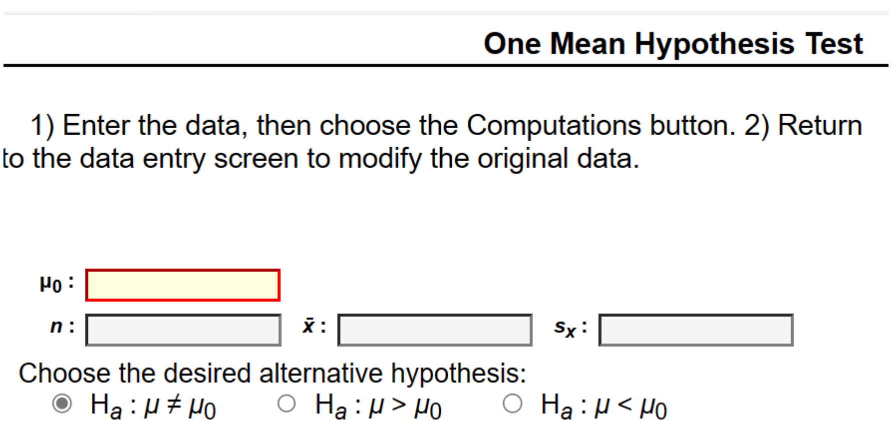
Solution. We are doing a hypothesis test for a single mean, given summary information about the sample. The hypotheses are:

$$H_0: \mu = 65$$
$$H_a: \mu \neq 65$$

We begin by choosing the *Tests* option in the calculator, yielding this list of submenu options:



The submenu option we require is *One mean (from sample statistics)* – we are examining the mean for a single population, and we have been given summary information about the sample data. When we choose that option, we obtain this screen:

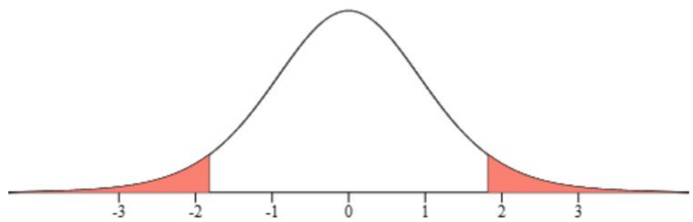


We enter 65 for μ_0, 15 for n, 67.2 for $\bar{x}$, and 4.7 for s_x. The default for the alternative hypothesis is what we want for a two-tail test, so we choose *Computations* to obtain the results:

$H_0 : \mu = 65 \quad H_a : \mu \neq 65$

$n = 15 \quad \bar{x} = 67.2 \quad s_x = 4.7$

$t = 1.8129 \quad p\text{-value} = 0.0913$



The *t* test statistic is 1.8129, with a p-value of 0.0913. We do not reject the null hypothesis.

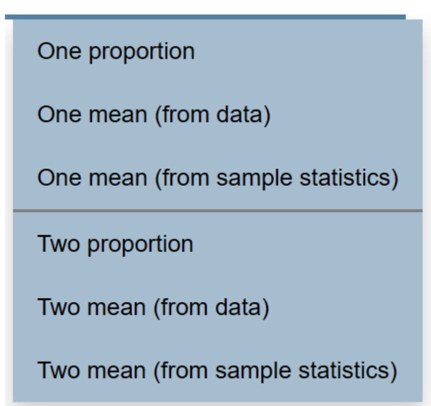
> **Exercise 7:** In Exercise 2 you solved the following problem "by hand." This time you should use the calculator to solve the problem. Your answers may vary slightly from those you got before, due to the rounding you did when you worked "by hand."
>
> Carry out a left-tail test to examine the claim that the population mean is 514, given that your sample of size 56 has a sample mean of 490 with a standard deviation of 98.

The applet at the following link provides additional practice. This is the same applet you have already used to carry out the calculations "by hand." This time, use the calculator for the calculations.

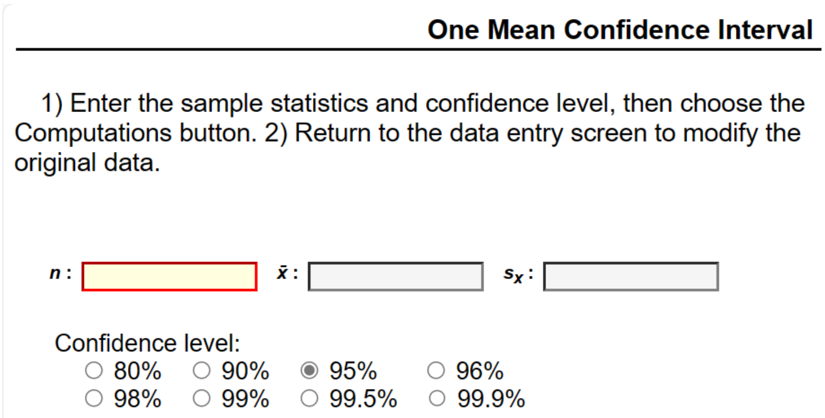
Mean hypothesis test - calculations

Example. Suppose we take a sample of size 15 and obtain a mean of 67.2, with a standard deviation of 4.7, for the variable we are measuring. Find a 95% confidence interval for the population mean.

Solution. We are doing a confidence interval for a single mean, given summary information about the sample. We begin by choosing the *Intervals* option in the calculator, yielding this list of submenu options:



The submenu option we require is *One mean (from sample statistics)* – we are examining the mean for a single population, and we have been given summary information about the sample data. When we choose that option, we obtain this screen:



We enter 15 for n, 67.2 for $\bar{x}$, and 4.7 for s_x. The default for the confidence level is what we want (a 95% confidence interval), so we choose *Computations* to obtain the results:

Point estimate ($\bar{x}$) = 67.2000

95% confidence interval: (64.5972, 69.8028)

Details:
$n = 15$, $s_x = 4.7$
Standard error = 1.2135
$t^* = 2.1448$
Margin of error = 2.6028

In interval notation, rounded to the nearest tenth, the confidence interval is (64.6, 69.8).

Exercise 8: In Exercise 5 you solved the following problem "by hand." This time you should use technology to solve the problem. Your answers may vary slightly from those you got before, due to the rounding you did when you worked "by hand."

Calculate a 99% confidence interval for a population mean for each situation described. Round your final answer to the nearest whole number. For each situation, is it plausible that the population mean is 450?

a. sample size 28, mean for the sample is 497, standard deviation for the sample is 93
b. sample size 55, mean for the sample is 497, standard deviation for the sample is 93 (Note: in addition to rounding differences, your answer will differ from the "by hand" answer for the following reason. Due to limitations in Table B, when you did the problem by hand you had to use 50 as the degrees of freedom. Using technology, the correct value of 54 will be used by the calculator.)

The applet at the following link provides additional practice. This is the same applet you have already used to carry out the calculations "by hand." This time, use the calculator for the calculations.

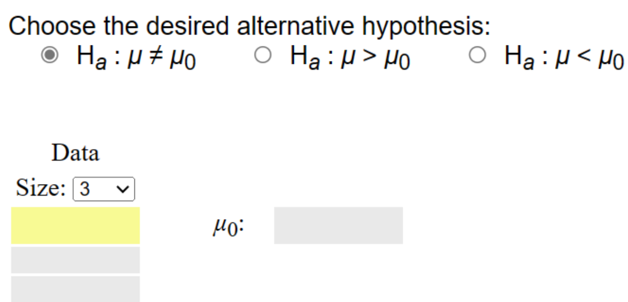
[Mean confidence interval - calculations]

What if I have actual data?

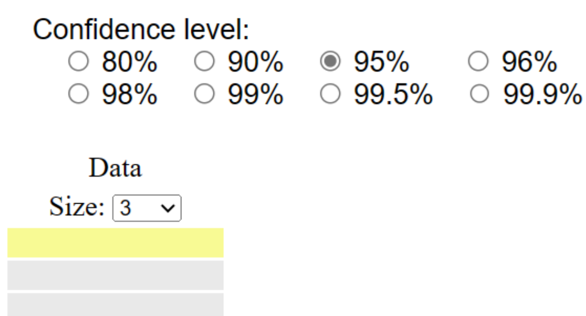
In both these examples, we have been given summary statistics (sample mean, sample standard deviation, and sample size). What would we do differently given a set of data instead of the summary information? An obvious solution would be to use the methods developed in Lesson 2 to enter the data, calculate the statistics, then use the methods developed in the two previous examples, by entering those summary statistics as described in the examples.

However, the calculator provides a more direct capability. Both the *Tests* and the *Intervals* menus contain submenu option *One mean (from data)*. If we choose this submenu option, we obtain a screen where we can enter the data, along with additional information about the test or interval to be calculated, then press the *Computations* button. Here are the corresponding screens:

For hypothesis test:



For confidence interval:



Exercise 9:

a. Test the claim that the mean for a particular population is 10, given the following set of sample data from that population.

10 9 43 17 20 28 35 11

b. For the same data, give an appropriate 99% confidence interval for the population mean. *Hint: You can avoid re-entering the data by using the* Save to file *and* Load from file *options that are supplied.*

11.5 – Assumptions

Just as was true for working with population proportions, there are certain "assumptions" associated with the methods we have described. (By "assumptions" we mean conditions which must be met if we wish to use the methods.) There are three assumptions; we will list them, then discuss each briefly.

1. The variable in question must be numerical.
2. The results must be based on a random sample drawn from the population being studied.
3. The variable for the entire population must have a normal distribution.

In some ways, the first of these three conditions need hardly be mentioned at all. After all, if the variable is categorical, it makes no sense to talk about the population mean for the variable – instead, we should examine a population proportion. As a practical matter for the student just learning about statistics, however, this is a crucial point. We now have methods to use for categorical variables and proportions, and similar but distinct methods to use for numerical variables and means. Your first task when encountering a new problem is to analyze whether the variable being studied is categorical or numerical.

The second condition is one we have discussed earlier, beginning in Lesson 4. Throughout these notes, we are assuming that the researcher has obtained a simple random sample from the population being studied. In practice, it may be very difficult to achieve a simple random sample when carrying out a study. When you read about statistical studies, keep in mind that the results are trustworthy only to the extent that randomness has been achieved in carrying out the study, and to the extent that the sample has been drawn from the same population the researcher is making claims about.

On the surface, the third condition appears to be a significant problem. In practice, it is difficult (impossible?) to guarantee that the population being studied is normal. We can look at a histogram for the data in our sample. If it is mound-shaped, that is an indication – but not proof – that the population could be normal. Fortunately, even when the population is not normal, the method yields reasonably good results (we say it is ***robust***). The larger your sample, the less important it is to have a normal population. For samples of size 30 or larger, the method gives good results for most population distributions, unless the distribution is extremely skewed or has extreme outliers.

By "good results" we mean that the advertised confidence level or significance level is reasonably accurate. For a 95% confidence interval, this means that the method does indeed give an interval containing the population mean for about 95% of the possible samples. For a hypothesis test using α = 0.01, the means that the method used does indeed restrict the likelihood of a Type 1 error to about 1%.

Note: If there are indications that the population may be skewed or contain outliers, you should avoid using a one-tail test. The comments on robustness apply quite well for confidence intervals and two-tail tests, but not as well for one-tail tests. Of course, as we have noted earlier, statisticians are increasingly avoiding one-tail tests for other important reasons.

11.6 – Intervals and Tests for Data Files

In Lessons 2, 3, 9, and 10 we have examined the use of the second calculator provided with these lessons to analyze the data in a data file. In this lesson we will see that this same calculator can be used to carry out hypothesis tests and confidence interval calculations for numerical variables present in a data file. The data file we use is the same used in those earlier lessons. You should have it saved to your own computing device, but in any case here again is a link to that file:

First day survey

As a reminder, the file contains student responses to a first day survey containing these questions:

1. What is your gender? (M) Male (F) Female
2. What is your class year? (FR) Freshman (SO) Sophomore (JR) Junior (SR) Senior
3. How many states have you visited?
4. Do you currently smoke? (Y) Yes (N) No
5. How tall are you (in inches)?
6. How many days per week do you read a newspaper?

To begin, open the data file calculator using the following link, then use the *Load vertical file* button to load the data file containing the student responses.

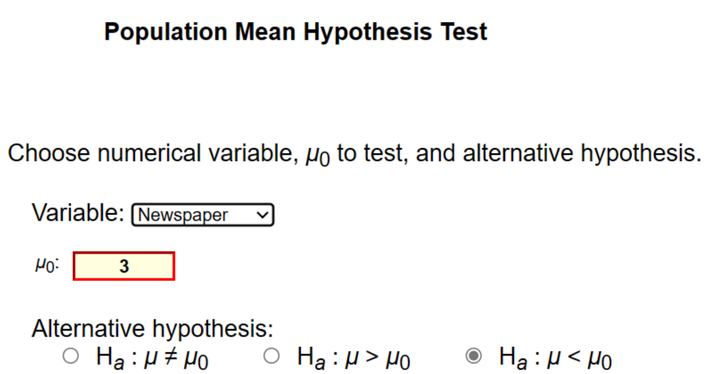
Data file calculator

Example. A researcher suspect that the students who take that class read the paper, on average, fewer than 3 days a week. Do an appropriate test.

Solution. These are the hypotheses:

$$H_0: \mu = 3$$
$$H_a: \mu < 3$$

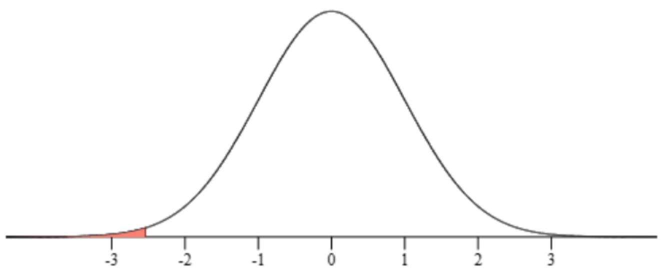
Use the menu option *Tests*, submenu *One mean hypothesis test*, and enter the information as shown here:



Here are the results when you press the *Computations* button:

Hypothesis test for population mean
Variable: *Newspaper*

$H_0 : \mu = 3$
$H_a : \mu < 3$
$\bar{x} = 2.3492$
p-value = 0.007



We have strong evidence ($p = 0.007$) that for the entire population of students who take that course, the mean number of times reading the newspaper per week is less than 3.

Example. Calculate a 99% confidence interval for the number of states visited.

Solution. Use the menu option *Intervals*, submenu *One mean hypothesis test*, and enter the information as shown here:

Population Mean Confidence Interval

Choose numerical variable and confidence level.

Variable: States_Visited

Confidence level:
○ 80% ○ 90% ○ 95% ○ 96%
○ 98% ◉ 99% ○ 99.5% ○ 99.9%

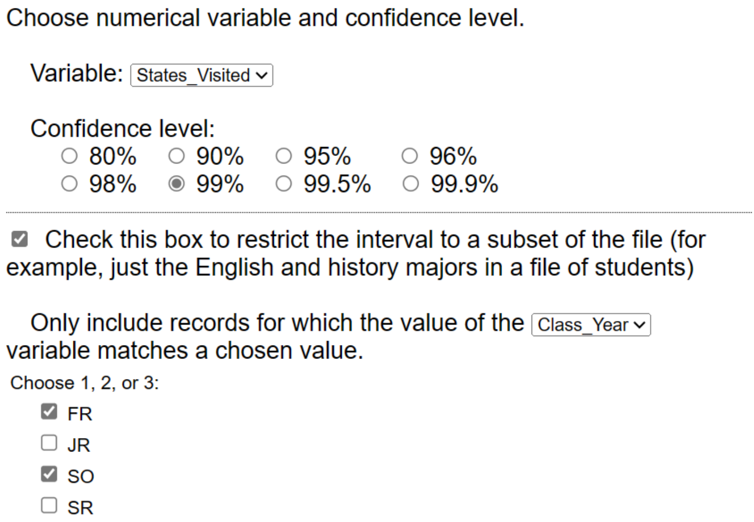
Here are the results when you press the *Computations* button:

Confidence interval for population mean
Variable: *States_Visited*
Point estimate ($\bar{x}$) = 15.8571
99% confidence interval: (13.3359, 18.3784)

Rounded to the nearest tenth, we are 99% confident that the average number of states visited for the population of all students who take that course is between 13.3 and 18.4.

Example: Do the same, but for the underclassmen (freshman/sophomore) students who take the course.

Solution: We simply restrict the calculation just as we have in earlier lessons, as shown here:



Here is the result:

Confidence interval for population mean
Variable: *States_Visited*
Restricted to records where *Class_Year* is one of:
FR, SO
Point estimate ($\bar{x}$) = 13.8529
99% confidence interval: (11.1096, 16.5963)

Exercise 10:

- ***a.*** Test the claim that the average number of states visited for the females who take the course is 11.
- ***b.*** Find a 90% confidence interval for the average height of the males.

Exercise 11: In Exercise 16 of Lesson 2 you created a data file containing the data presented originally in Lesson 1. This data was collected in a statistics course at a public university. Use that data file to:

a. Test the claim that the average height for the males who take the course is 70 inches.
b. Calculate a 98% confidence interval for the average amount of time spent per week on the course.

Solutions to Exercises

1: Carry out a two-tail test to examine the claim that the population proportion is 43%, given that your sample of size 956 has a sample proportion of 47.49%.
$H_0: p = 0.43$
$H_a: p \neq 0.43$
$se = \sqrt{\frac{.43(1-.43)}{956}} = 0.0160$
$z = \frac{.4749-.43}{.0160} = 2.8063$
p-value = 0.0050 (using menu option *Distributions*, submenu *Standard normal p-value,*)
At significance level 0.05 or 0.01 we would reject the null hypothesis, and conclude that the population proportion is not 43%. (It appears to be larger than 43%, based on the sample.)

2: Carry out a left-tail test to examine the claim that the population mean is 514, given that your sample of size 56 has a sample mean of 490 with a standard deviation of 98.
$H_0: \mu = 514$
$H_a: \mu < 514$
$se = \frac{98}{\sqrt{56}} = 13.0958$
$t = \frac{490-514}{13.0958} = -1.8326$
The degrees of freedom is *df* = 55. The (one-tail, left tail) *p*-value from our calculator is 0.0361. The conclusion depends on our chosen significance level. At significance level 0.05 we would reject the null hypothesis, concluding that there is enough evidence to say that the population mean is less than 514. On the other hand, at 0.01 we do not reject the null hypothesis, and we conclude that there is not enough evidence to say that the population mean is less than 514. Recall that using 0.05 rather than 0.01 increases the possibility that you might commit a Type 1 error, rejecting a true null hypothesis.

3: For a certain brand of soft drink, the bottles are supposed to (on average) contain 16 ounces. Quality control routinely conducts sampling to ensure the proper functioning of the machines that fill the bottles. In a recent sample of 85 bottles, the bottles had an average content of 15.87 ounces with a standard deviation of 0.59 ounces.

a. Write the null and alternative hypotheses for a suitable hypothesis test.
The null hypothesis states "no difference," or put another way, "things are just as they should be." In this case, that states that the mean content is 16 ounces, written symbolically:
$H_0: \mu = 16$
The alternative hypothesis simply negates this:
$H_0: \mu \neq 16$
Notice that nothing in the statement implies that quality control suspects either less or more than there should be. In fact, both outcomes would be important to uncover – the former to avoid penalties by oversight agencies, the latter to avoid extra expense for the extra soda being dispensed.

b. The results of the calculations for the hypothesis test are $t = -2.0314$, *p*-value = 0.0454. Using $\alpha = 0.01$, write the conclusion in the context of the problem.
Since 0.0454 is not less than 0.01, we do not reject the null hypothesis. We did not find evidence to conclude that the mean amount of soda being dispensed per bottle is any different from the 16 ounces it is supposed to be.

c. With these results, what should quality control do? Although this particular sample averaged slightly less than 16 ounces, it was not enough to conclude that there is any problem with the dispensing equipment. The sample is satisfactory.

(However, note that the conclusion would be different using $\alpha = 0.05$. One of the decisions facing a quality control department is this: *How willing are we to make a Type 1 error, that is, to reject a true null hypothesis?* Put another way: *How willing are we to shut down a machine that is actually dispensing the correct amount on average?* Remember that the choice of significance level is tied to the probability of Type 1 error.)

4: Calculate a 99% confidence interval for a population proportion if the proportion for the sample is 65% and the sample size is 1500.

$se = \sqrt{\frac{.65(1-.65)}{1500}} = 0.0123,\ z^* = 2.576,\ m.e. = 2.576(.0123) = 0.0317$

Interval is $(.65 - .0317, .65 + .0317)$ = (0.6183,0.6817)

5: Calculate a 99% confidence interval for a population mean for each situation described. Round your final answer to the nearest whole number. For each situation, is it plausible that the population mean is 450?

a. sample size 28, mean for the sample is 497, standard deviation for the sample is 93

$se = \frac{93}{\sqrt{28}} = 17.5753,\ t^* = 2.771,\ m.e. = t^* \cdot se = 2.771(17.5753) = 48.7012$

Interval is $(497 - 48.7012, 497 + 48.7012) = (448, 546)$ to the nearest whole number

Based on this interval, it is plausible that the population mean is 450.

b. sample size 55, mean for the sample is 497, standard deviation for the sample is 93 (**Reminder**: if the table does not contain a row for the desired *df* value, use the next smaller *df* that is in the table.)

$se = \frac{93}{\sqrt{55}} = 12.5401,\ t^* = 2.678$ (we get this from the df = 50 row, since there is no df=54 row), $m.e. = t^* \cdot se = 2.678(12.5401) = 33.5824$

Interval is $(497 - 33.5824, 497 + 33.5824) = (463, 531)$ to the nearest whole number

Based on this interval, it is not plausible that the population mean is 450.

6: For a certain brand of soft drink, the bottles are supposed to (on average) contain 16 ounces. Quality control routinely conducts sampling to ensure the proper functioning of the machines that fill the bottles. In a recent sample of 85 bottles, the bottles had an average content of 15.87 ounces with a standard deviation of 0.59 ounces. The corresponding 99% confidence interval is (15.701,16.039).

a. Write a sentence explaining the meaning of this interval.

I am 99% confident that for the population consisting of all bottles being filled by this machine, the average amount of soda dispensed is between 15.701 ounces and 16.039 ounces.

b. Which of the following are correct statements to make based on this interval?

- I am 99% confident that all the bottles contain between 15.701 ounces and 16.039 ounces. No, the confidence interval is not a statement about the individual bottles, only about the *average* for all the bottles.
- There is a 99% probability that the next bottle sampled will contain between 15.701 ounces and 16.039 ounces. No, for the same reason. The confidence interval is not a statement about any individual bottle.
- It is plausible that the mean amount of soda being dispensed is only 15.7 ounces. No, because 15.7 is not within the interval; it is slightly smaller than the left endpoint of the interval.

- It is plausible that the mean amount of soda being dispensed is only 15.8 ounces. Yes, 15.8 is one of the values in the interval, so that value is plausible.
- It is plausible that the mean amount of soda being dispensed is the correct value (namely, 16 ounces). Yes, 16 is one of the values in the interval, so that value is plausible.

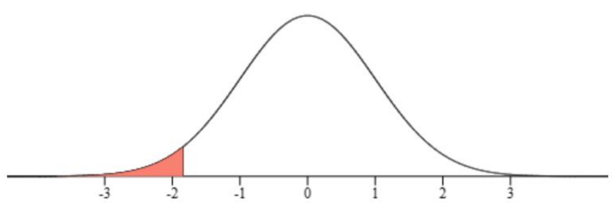
c. With these results, what should quality control do? Since it is plausible that the average is 16 ounces, the results are consistent with a machine that is working correctly. (There is no evidence that it is *not* working correctly.)

7. (Using technology) Carry out a left-tail test to examine the claim that the population mean is 514, given that your sample of size 56 has a sample mean of 490 with a standard deviation of 98.
The menu option is *Tests*, submenu option *One mean (from sample statistics)*. Fill in the information as shown – notice we chose the alternative hypothesis for a left-tail test:

μ_0 : 514
n : 56 $\bar{x}$: 490 s_x : 98
Choose the desired alternative hypothesis:
○ $H_a : \mu \neq \mu_0$ ○ $H_a : \mu > \mu_0$ ◉ $H_a : \mu < \mu_0$

Then press *Computations* to obtain the results:

$H_0 : \mu = 514 \quad H_a : \mu < 514$
$n = 56 \quad \bar{x} = 490 \quad s_x = 98$
$t = -1.8326 \quad p\text{-value} = 0.0361$



The test statistic is $t = -1.8326$ with p-value 0.0361.

8. (Using technology) Calculate a 99% confidence interval for a population mean for each situation described. Round your final answer to the nearest whole number. For each situation, is it plausible that the population mean is 450?
a. sample size 28, mean for the sample is 497, standard deviation for the sample is 93
The menu option is *Intervals*, submenu option *One mean (from sample statistics)*. Fill in the information as shown – notice we chose the desired confidence level (99%):

n : 28 $\bar{x}$: 497 s_x : 93

Confidence level:
○ 80% ○ 90% ○ 95% ○ 96%
○ 98% ◉ 99% ○ 99.5% ○ 99.9%

Use the *Computations* option to obtain the results:

Point estimate $(\bar{x}) = 497.0000$

99% confidence interval: (448.3043, 545.6957)

Details:
$n = 28, \quad s_x = 93$
Standard error = 17.5753
$t^* = 2.7707$
Margin of error = 48.6957

Rounded to the nearest whole number, the confidence interval is $(448, 546)$. It is plausible that the population mean is 450.

b. sample size 55, mean for the sample is 497, standard deviation for the sample is 93

Using the same menu options, enter 55 for n, 497 for $\bar{x}$, and 93 for s_x, and choose confidence level 99%. Press *Computations*; the results include the following:

99% confidence interval: (463.5181, 530.4819)

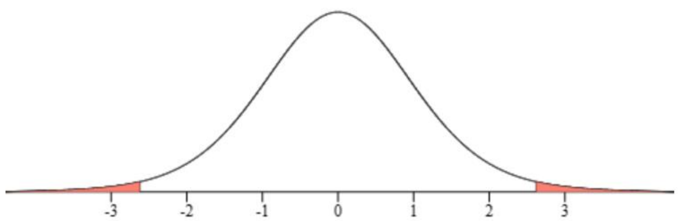
Rounded to the nearest whole number, the confidence interval is $(464, 530)$. It is not plausible that the population mean is 450.

9:

a. Test the claim that the mean for a particular population is 10, given the following set of sample data from that population.

10	9	43	17	20	28	35	11

$H_0 : \mu = 10 \qquad H_a : \mu \neq 10$
$n = 8 \quad \bar{x} = 21.625 \quad s_x = 12.5805$
$t = 2.6136 \quad p\text{-value} = 0.0347$



At α = 0.05, we have evidence that the population mean is not 10. (Since the sample mean is larger than 10, we believe the population mean is larger than 10.)

Note: At α = 0.01, we do **not** have evidence that the population mean is not 10 – it is **possible** that it **is** 10.

b. For the same data, give an appropriate 99% confidence interval for the population mean. *Hint: You can avoid re-entering the data by using the* Save to file *and* Load from file *options that are supplied.*

Point estimate $(\bar{x}) = 21.6250$

99% confidence interval: (6.0598, 37.1902)

Details:
$n = 8, \quad s_x = 12.5805$
Standard error = 4.4479
$t^* = 3.4995$
Margin of error = 15.5652

Rounded to the nearest tenth, we are 99% confident that the population mean is between 6.1 and 37.2. Notice that at a 99% confidence level, 10 is a plausible value for the population mean.

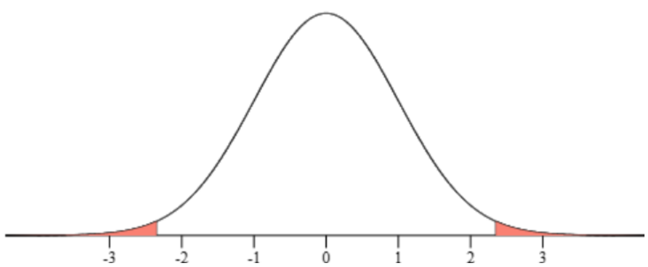
Note: The 95% confidence interval is (11.1075, 32.1425). At the 95% confidence level, 10 is **not** a plausible value for the population mean.

10:

a. Test the claim that the average number of states visited for the females who take the course is 11.

Hypothesis test for population mean
Variable: *States_Visited*
Restricted to records where *Gender* equals:
F

$H_0 : \mu = 11$
$H_a : \mu \neq 11$
$\bar{x} = 14.9583$
p-value = 0.0279



At a 5% significance level (but not at 1%) we have evidence to support the alternative hypothesis: The mean number of states visited for all the females who take the course is not 11 (we believe it is greater than 11).

b. Find a 90% confidence interval for the average height of the males.

Confidence interval for population mean
Variable: *Height(in)*
Restricted to records where *Gender* equals:
M
Point estimate ($\bar{x}$) = 70.8718
90% confidence interval: (70.1078, 71.6358)

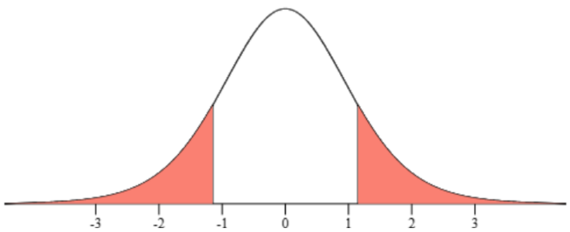
We are 90% confident that the average height for all the males who take the course is between 70.1 and 71.6 inches (to the nearest tenth of an inch).

11: In Exercise 16 of Lesson 2 you created a data file containing the data presented originally in Lesson 1. This data was collected in a statistics course at a public university. Use that data file to:

a. Test the claim that the average height for the males who take the course is 70 inches.

Hypothesis test for population mean
Variable: *Height*
Restricted to records where *Gender* equals:
Male

$H_0 : \mu = 70$
$H_a : \mu \neq 70$
$\bar{x} = 68.6$
p-value = 0.2711



We do not have enough evidence to dispute the claim that the average height for all the males who take the course is 70 inches.

b. Calculate a 98% confidence interval for the average amount of time spent per week on the course.

Confidence interval for population mean
Variable: *Course Time*
Point estimate ($\bar{x}$) = 5.4000
98% confidence interval: (4.0600, 6.7400)

We are 98% confident that on average all the students who take the course will spend between 4.06 and 6.74 hours per week on the course.