

Section 3B – Test Statistics and Critical Values

Vocabulary

Population: The collection of all people or objects to be studied.

Census: Collecting data from everyone in a population.

Sample: Collecting data from a small subgroup of the population.

Statistic: A number calculated from sample data in order to understand the characteristics of the data.
For example, a sample mean average, a sample standard deviation, or a sample percentage.

Parameter: A number that describes the characteristics of a population like a population mean or a population percentage. Can be calculated from an unbiased census, but is often just a guess about the population.

Hypothesis Test: A procedure for testing a claim about a population.

Null Hypothesis (H_0): A statement about the population that involves equality. It is often a statement about “no change”, “no relationship” or “no effect”.

Test Statistic: A number calculated in order to determine if the sample data significantly disagrees with the null hypothesis. There are a variety of different test statistics depending on the type of data.

One-Population Proportion Test Statistic (z): The sample proportion is this many standard errors above or below the population proportion in the null hypothesis.

One-Population Mean Test Statistic (t): The sample mean is this many standard errors above or below the population mean in the null hypothesis.

Two-Population Proportion Test Statistic (z): The sample proportion for group one is this many standard errors above or below the sample proportion for group two. Can also be thought of as how many standard errors the difference between the sample proportions is from zero or from some other specific difference.

Two-Population Mean Average Test Statistic (t): The sample mean for group one is this many standard errors above or below the sample mean for group two. Can also be thought of as how many standard errors the difference between the sample means is from zero or from some other specific difference.

Critical Values: Number or numbers we compare our test statistic to in order to determine significance. In a sampling distribution or a theoretical distribution approximating the sampling distribution, the critical value shows us where the tail or tails are. The test statistic must fall in the tail to be significant.

Critical Value Z-score (Z_c): The number of standard errors the sample proportion should be from the population proportion based on the significance level (or confidence level), tails, and the standard normal distribution. Compare the Z test statistic to these critical values.

Critical Value T-score (T_c): The number of standard errors the sample mean should be from the population mean based on the significance level (or confidence level), degrees of freedom, tails and the student T distribution. Compare the t test statistic to these critical values.



Introduction

A hypothesis test compares random sample data to the null and alternative hypothesis in order to determine the validity of a population claim. Remember, the null hypothesis is the population statement that involves equality, usually to some population parameter. The principle of sampling variability (random chance) tells us that random sample statistics will usually be very different than population parameters in the null hypothesis. We know the sample data will almost always disagree with the null hypothesis. The question is does it significantly disagree?

Significant differences are very difficult to determine. For example, a sample mean is 4.3 kg higher than the population mean in the null hypothesis, but is that significant? A sample percentage may be 3.5% below the population percentage in the null hypothesis, but is that significant? Statisticians and mathematicians put a lot of thought into figuring out ways to determine if a sample statistic significantly disagrees with a population parameter. The result was the invention of the “test statistic”.

Test Statistic: A number calculated in order to determine if the sample data significantly disagrees with the null hypothesis.

Data can take on many forms, categorical or quantitative, one-population or two-population or ten-populations. So there are many types of test statistics to deal with the various types of data and the number of populations. Test statistics have one thing in common. They are trying to see if the random sample data significantly disagrees with the null hypothesis or if it is not a significant disagreement.

Using Test Statistics and Critical Values to Judge Significance

Key question: How can we tell if the test statistic indicates a significant disagreement between the sample data and the null hypothesis?

The answer to this is to compare the test statistic to a critical value.

Critical Value: We compare a test statistic to this number to determine if the sample data significantly disagrees with the null hypothesis.

Critical values and test statistics are always different depending on the situation. Am I dealing with a left-tailed, right-tailed or two-tailed test? What confidence level (or significance level) should I use? Most computer programs calculate the test statistics and critical values for a situation. Your job as a data scientist is to be able to explain these numbers and judge significance.

The most important rule to remember about test statistics is that the following. You may not understand the test statistic calculation or how the critical value was found, but you can at least judge whether your random sample data significantly disagrees with the null hypothesis or not.

Critical values help us judge significance. To understand a critical value you have seen the sampling distribution or a theoretical distribution curve that estimates a sampling distribution. If the test statistic falls in one of the tails of the sampling distribution or in the tail of the theoretical curve, this indicates significance. If the test statistic does not fall in one of the tails of the theoretical distribution curve, the sample data does not significantly disagree with the null hypothesis. A key question is where do the tails start? That is what the critical value is calculating. The critical value is the cutoff for where the tails of the distribution begin. We just need to compare the test statistic to the critical value and see if the test statistic falls in the tail. When a computer program calculates the test statistic and critical value for you, it is always good to draw the theoretical curve and use the critical values to visualize the tails. Now ask yourself if the test statistic is in the tail or not.

Significance Rule for Test Statistics and Critical Values

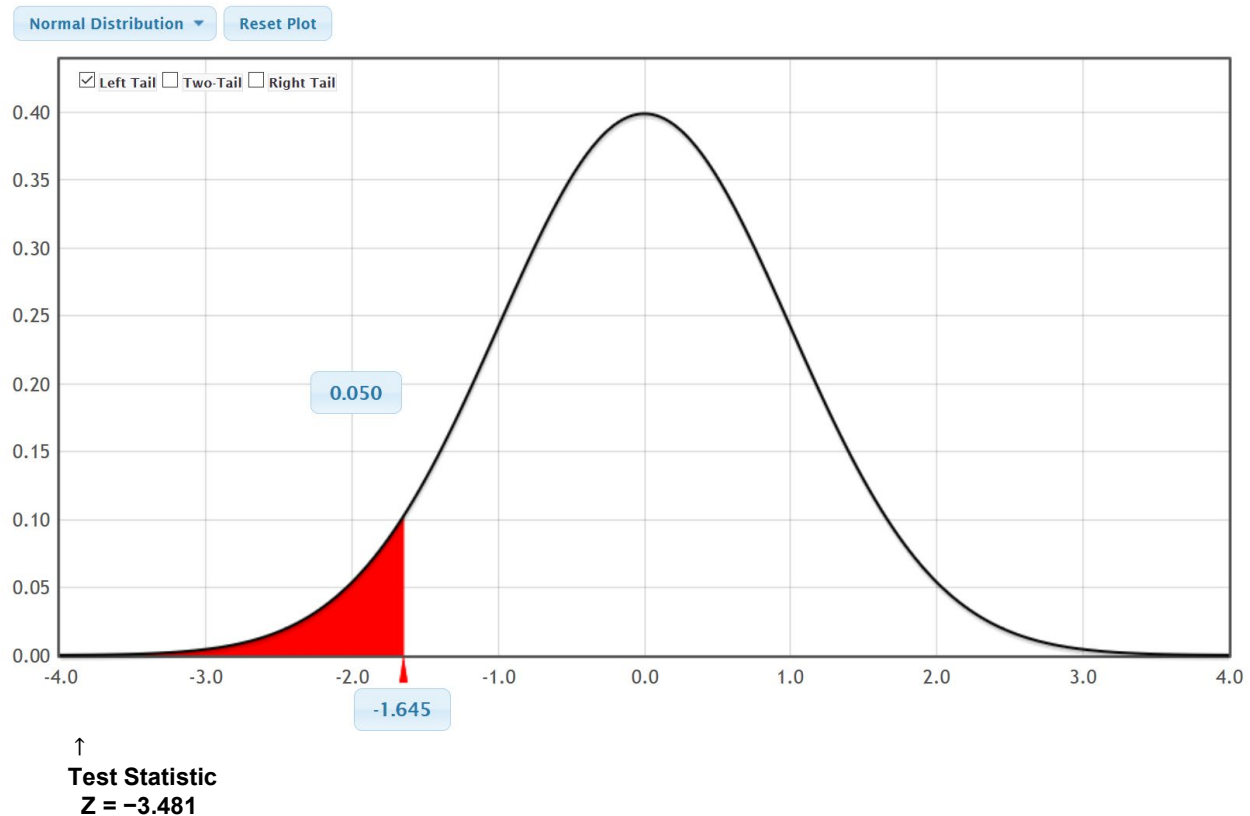
- **If the test statistic falls in the tail determined by the critical value (or values), then the sample data significantly disagrees with the null hypothesis.**
- **If the test statistic does NOT fall in the tail determined by the critical value (or values), then the sample data does NOT significantly disagree with the null hypothesis.**



Left-tailed test statistic example

Suppose we are doing a left tailed hypothesis test for proportions and the computer calculates a test statistic of $Z = -3.481$ and a critical value of -1.645 . The Z-distribution is a normal curve. So we need to see if the test statistic falls in the left tail. Negative numbers can be particularly challenging. Remember on the number line, -3.481 is less than -1.645 , so the test statistic does fall in the left tail. This means that the sample data significantly disagrees with the null hypothesis.

A nice way to visualize this principle is to go to the “Theoretical Distributions” menu in StatKey at www.lock5stat.com. Click on “normal”, then “Left Tail”. Put the critical value of -1.645 in the bottom box.

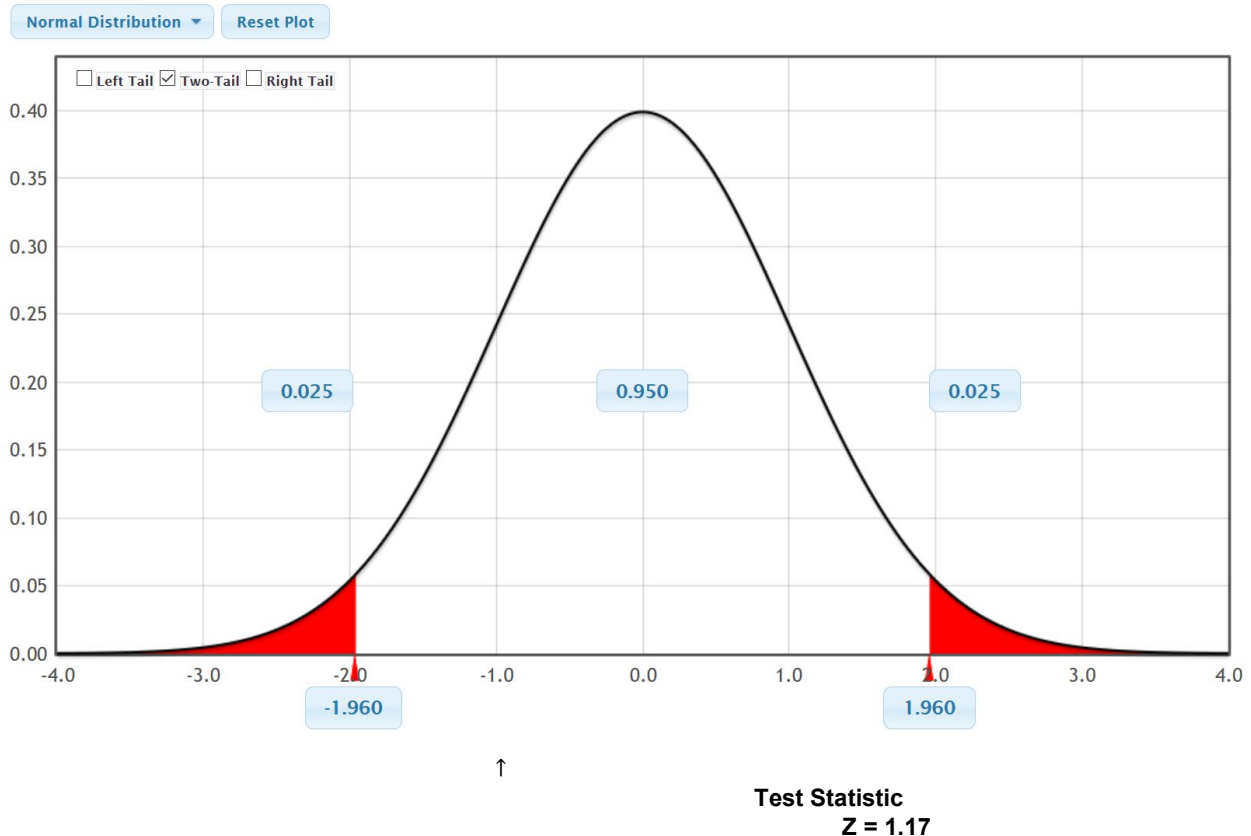


Two-tailed Test Statistic Example

Suppose we are doing a two-tailed hypothesis test for proportions. The computer calculated a Z-test statistic of $Z = +1.17$ and Critical Values of ± 1.96 .

Notice that in a two-tailed test, there are two critical values. In normal distributions like Z or T, two-tailed critical values are a certain number of standard errors above or below and so the positive and negative number of standard errors. In this example we see the “plus or minus” notation for the critical value ± 1.96 . This really means that the lower critical value for the left tail is -1.96 and the upper critical value for the right tail is $+1.96$. In non-normal distributions like F or χ^2 , the upper and lower critical values are not plus or minus. They all work the same way though. Is my test statistic in the tail. If we draw the normal Z-curve with the right tail starting at the upper critical value of $+1.96$ and left tail starting at -1.96 , we can see if the test statistic of $+1.17$ falls in the tail. Again, we can use StatKey to visualize the curve and the critical values. Go to the “Theoretical Distributions” menu in StatKey at www.lock5stat.com. Click on “normal”, then “Two Tail”. Put the critical value of -1.96 in the bottom box in the left tail and $+1.96$ in the bottom box of the right tail. Where does our test statistic fall?





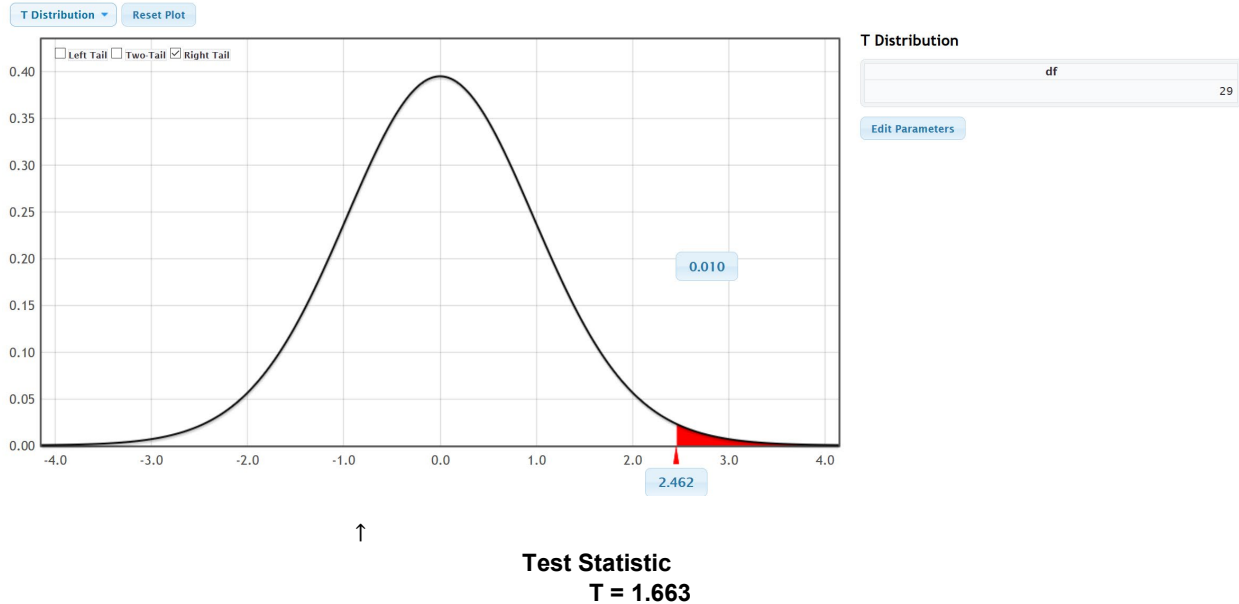
Notice our test statistic did not fall in one of the tails. So our test statistic and critical values are indicating that the sample data does not significantly disagree with the null hypothesis.

Right-tailed Test Statistic Example

Suppose we are doing a right tailed hypothesis test for means and a degrees of freedom of 29. The computer calculated a T-test statistic of $T = +1.663$ and a critical value of $+2.462$.

Like the Z-distribution, the T-distribution is also normal. So if this was a two-tailed test, we would see the “±” notation. This is a right tailed test though, so we see there is one upper critical value for the right tail. Draw the normal T-curve with the right tail starting at the upper critical value of $+2.462$. You can have StatKey draw it for you if you wish. Go to the “Theoretical Distributions” menu in StatKey at www.lock5stat.com. Click on “t”, then put in the degrees of freedom of 29. Now click “Right Tail” and put the critical value of $+2.462$ in the bottom box below the right tail. Does our test statistic fall in the tail?





Notice that our T-test statistic of 1.663 did not fall in the tail determined by the critical value. This again tells us that our sample data does not significantly disagree with the null hypothesis.

Significance Levels (Alpha Levels) (α)

The significance level (or alpha level) is an important number in a hypothesis test. It is often denoted by the Greek letter “alpha” or “ α ”. When scientists perform a hypothesis test, they choose a significance level for the test. This number is very important. It is tied to the critical value and is also very important in understanding the amount of sampling variability in a hypothesis test. Do you recall the confidence levels from the last chapter? Think of the significance level (α) as the opposite of the confidence level ($1 - \alpha$).

<u>Confidence Levels ($1 - \alpha$)</u>	<u>Significance Levels (α)</u>
90% (0.90)	10% (0.10)
95% (0.95)	5% (0.05)
99% (0.99)	1% (0.01)

If you want to be 95% confident, you will use a 5% significance level in your hypothesis test. Similarly, if you want to be 90% confident, you will use a 10% significance level. 99% confident corresponds to a 1% significance level.

Note: Remember that a 95% confidence level is the most common. Not surprisingly, a 5% significance level is also the most common. A good rule of thumb, is that if you do not know what significance level to use, use 5% ($\alpha = 0.05$).

One-population proportion Z-test statistic

The test statistic used for a one-population proportion (%) hypothesis test is the Z-test statistic. The Z-test statistic counts the number of standard errors that the sample proportion (\hat{p}) is above or below the population proportion (π) in the null hypothesis. If you recall from previous chapters, the number of standard errors or standard deviations is often called a “Z-score”. Not surprisingly, the one-population proportion test statistic is a Z-score.

Key question: How can we tell if the Z-test statistic indicates a significant disagreement between the sample proportion and the population proportion in the null hypothesis?



Remember, to know if a test statistic indicates a significant difference, compare the test statistic to a critical value. If the test statistic fall in the tail of the distribution determined by the critical value, then it is a significant difference.

If you remember, there are three famous critical value Z-scores for 90%, 95% and 99% confidence (± 1.645 , ± 1.96 or ± 2.576). These numbers are often used to see if our Z-test statistic is significant.

One-population proportion Z-test statistic example

Suppose we want to test the claim that the population percentage is 25%. Let's look at the percentage problem where the population proportion (π) is 0.25 (25%) and the sample proportion (\hat{p}) is 0.22 (22%). We know the sample value is % lower, but we do not know if that is significant. Another important bit of information is the sample size. In this case, it was 100.

$$H_0: \pi = 0.25 \text{ (Claim)}$$

$$H_A: \pi \neq 0.25$$

To find out if our sample data disagrees with the null hypothesis, let's calculate the test statistic.

Formulas for Test Statistics often follow general patterns. A one-population proportion or mean test statistic counts how many standard errors that the sample statistic is above or below the population parameter in the null hypothesis. Here is the general formula.

$$\frac{(\text{Sample Statistic} - \text{Population Parameter})}{\text{Standard Error}}$$

Remember in the last unit we saw that statisticians often used formulas to approximate the standard error. For one-proportion confidence intervals the standard error was $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. In a hypothesis test we have an idea of what the population proportion (π) is, so we prefer to use π instead of \hat{p} in the standard error estimation formula.

Here is the formula for a one-population proportion Z-test statistic.

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

Let's plug in our numbers. Remember that the sample proportion $\hat{p} = 0.22$, the population proportion $\pi = 0.25$ and the sample size $n = 100$.

$$Z = \frac{(\hat{p} - \pi)}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{(0.22 - 0.25)}{\sqrt{\frac{0.25(1-0.25)}{100}}} = \frac{(0.22 - 0.25)}{\sqrt{\frac{0.25(0.75)}{100}}} \approx \frac{(-0.03)}{0.0433} \approx -0.69$$

It is always very important to explain your test statistic. In this one population proportion hypothesis test, the Z-test statistic is counting how many standard errors the sample proportion (\hat{p}) is above or below the population proportion (π). If the test statistic is negative, it is "below". If the test statistic is positive, it is "above".

Z-test statistic sentence:

The sample proportion (0.22) was 0.69 standard errors below the population proportion (0.25).

OR

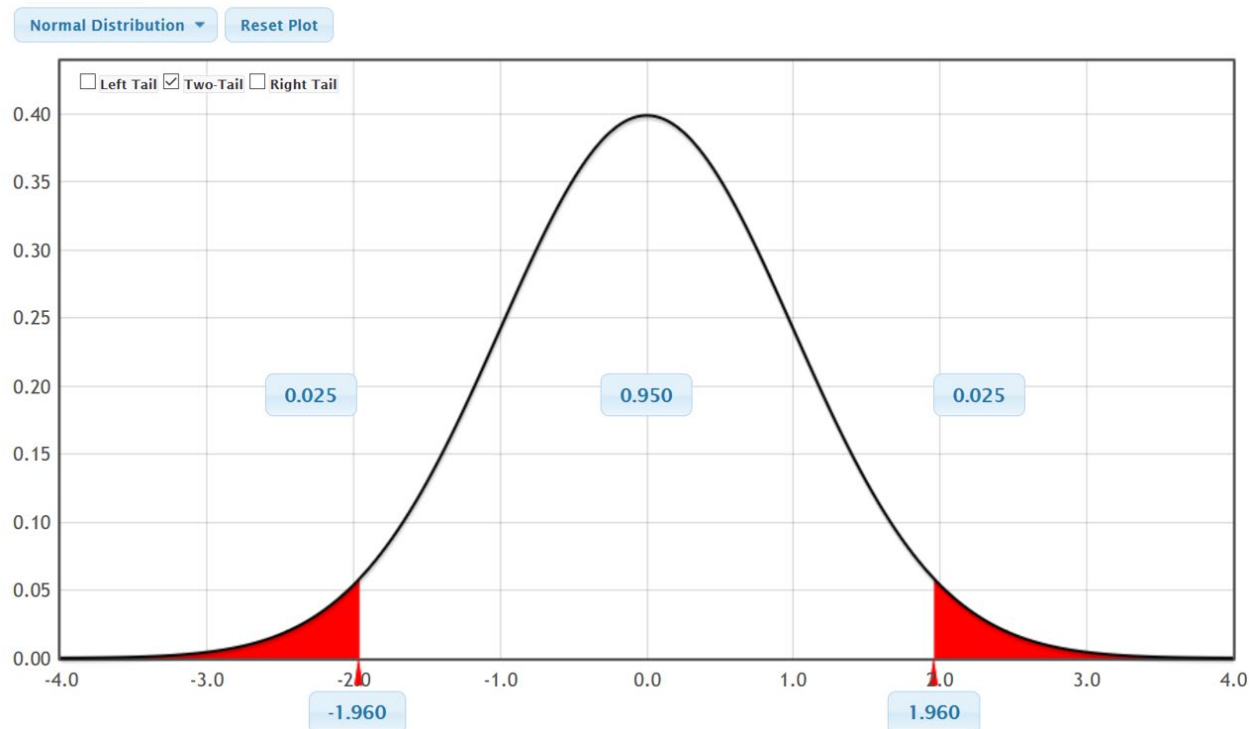
The sample percentage (22%) was 0.69 standard errors below the population percentage 25%.



Is it significant? To determine if this is significant, we should calculate the critical value. Computer programs usually calculate the test statistic and critical value for us. We also learned in the last chapter that we can use StatKey to calculate a critical value.

To calculate a critical value, you must first determine what significance level you plan to use and if your hypothesis test is left-tailed, right-tailed or two-tailed. In this example, the data scientist used a 5% significance level. Notice the alternative hypothesis was “ \neq ”, so this is a two-tailed test.

Go to www.lock5stat.com and click on “StatKey”. Under the theoretical distributions menu click on “normal”. This is the menu for calculating critical value Z-scores. Leave the mean at zero and the standard deviation as one. Click two tail. The significance level is the probability in the tail. Since there are two tails, the significance level will be broken in half with 2.5% in each tail. The numbers on the bottom are the critical values.



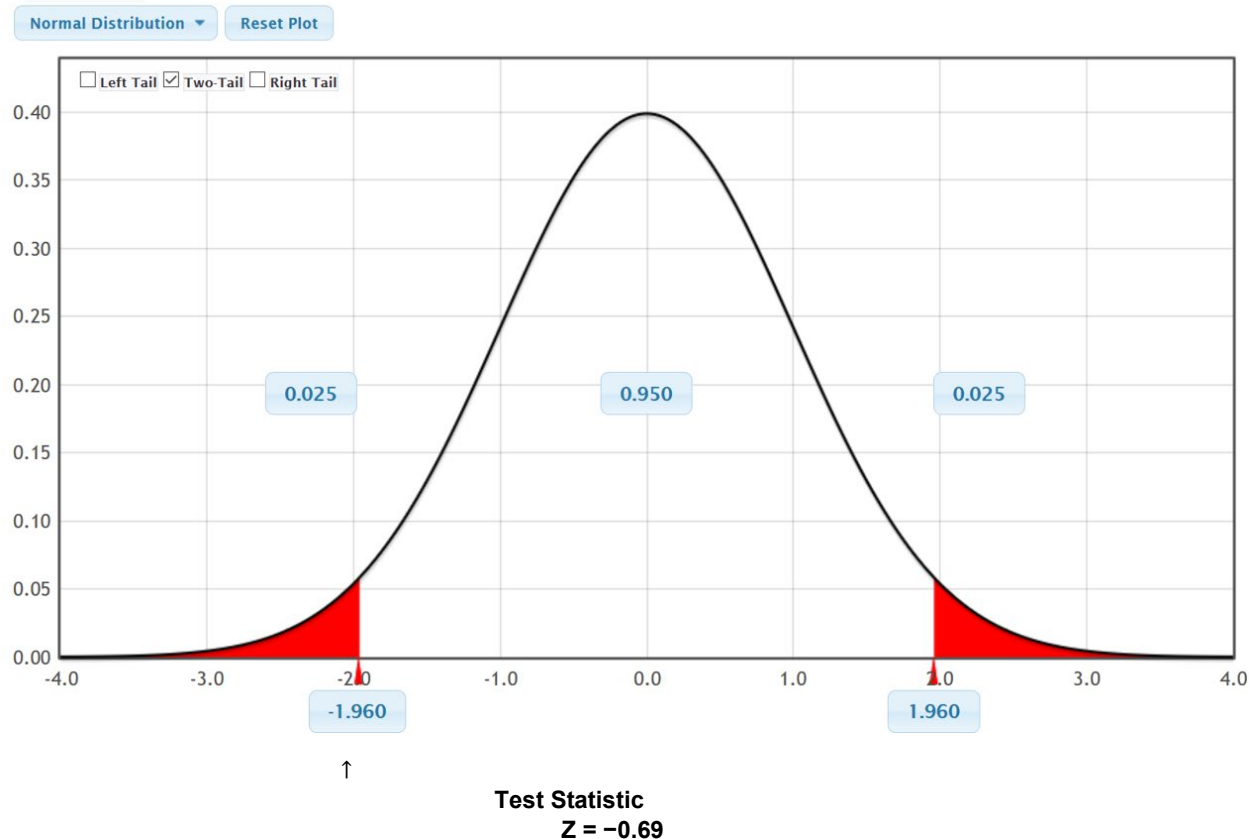
So we see that our critical values for this hypothesis test are ± 1.96 . Remember, our test statistic is $Z = -0.69$. Let's use our significance rule to determine if the sample data significantly disagreed with the null hypothesis.

Significance Rule for Test Statistics and Critical Values

- If the test statistic falls in the tail determined by the critical value (or values), then the sample data significantly disagrees with the null hypothesis.
- If the test statistic does NOT fall in the tail determined by the critical value (or values), then the sample data does NOT significantly disagree with the null hypothesis.

Did our Z-test statistic of $Z = -0.69$ fall in one of the tails?





Our test statistic did not fall in the tail, so is not significant. The sample proportion and population proportion in the null hypothesis are not significantly different. They are actually pretty close. This also tells us that the sample data does not significantly disagree with the null hypothesis.

Statcato: Most people working in statistics today do not calculate test statistics and critical values. Statistics computer programs can calculate them for us. In Statcato, go to the statistics menu, then click on hypothesis test. Click on "one-population proportion". Most computer programs require the number of success (events) to calculate proportions. In the last example, they did not tell us the number of successes. They just told us the total sample size (100) and the sample proportion (0.22). If you multiply them, you get the number of successes or events.

$$\text{Number of Successes (events)} = 100 \times 0.22 = 22$$

So in the one-population proportion hypothesis test menu in Statcato, go to the summarized data and enter 22 for the number of events and 100 for the total number of trials. Put in not equal for the alternative hypothesis, 0.25 for the hypothesized proportion, and a significance level of 0.05. Here is the Statcato printout. Notice the test statistic is close to our calculation and the critical values are the same as what we calculated with StatKey.



Inputs <input type="radio"/> Samples in column: <input checked="" type="radio" value="1"/> Summarized sample data: Number of events: <input type="text" value="22"/> Number of trials: <input type="text" value="100"/>		Significance <input checked="" type="radio"/> Significance Level: <input type="text" value="0.05"/> 0 - 1.00 (e.g. 0.05) <input type="radio"/> Confidence Level: <input type="text" value="0.95"/> 0 - 1.00 (e.g. 0.95)	
		Alternative Hypothesis Alternative Hypothesis: <input type="text" value="Not Equal to"/> Hypothesized Proportion: <input type="text" value="0.25"/>	
<input type="button" value="OK"/> <input type="button" value="Cancel"/>			

Hypothesis Test - One population proportion: confidence level = 0.95

Input: Summary data

Null hypothesis: $p = 0.25$

Alternative hypothesis: $p \neq 0.25$

N	Sample Proportion	Significance Level	Critical Value	Test Statistic Z	p-Value
100	0.22	0.05	-1.96, 1.96	-0.693	0.4884

Note about Sample Size

Sample size plays a key role in significance. Let's look at the previous example but with a sample size of 1000.

$$Z = \frac{(\hat{p} - \pi)}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{(0.22 - 0.25)}{\sqrt{\frac{0.25(1-0.25)}{1000}}} = \frac{(0.22 - 0.25)}{\sqrt{\frac{0.25(0.75)}{1000}}} \approx \frac{-0.03}{0.0137} \approx -2.19$$

If our sample size had been 1000, then the sample percentage of 22% is significantly lower than our population percentage of 25%. In fact, our sample percentage of 22% is 2.19 standard errors below the population percentage of 25%. A test statistic of $Z = -2.19$ would have fallen in the left tail since it is less than the lower critical value of -1.96 . At a sample size of 1000, 22% is significantly less than 25%.

One-population mean T-test statistic

The test statistic used for a one-population mean hypothesis test is the T-test statistic. Like the Z-test statistic, the one-population T-test statistic also counts the number of standard errors that the sample statistic is from the population parameter in the null hypothesis. In this case, it will count the number of standard errors that the sample mean (\bar{x}) is above or below the population mean (μ) in the null hypothesis.

Key question: How can we tell if the T-test statistic indicates a significant disagreement between the sample mean (\bar{x}) and the population mean (μ) in the null hypothesis?

Remember, to know if a test statistic indicates a significant difference, compare the test statistic to a critical values. For a significant difference, the test statistic should fall in the tail of the distribution determined by the critical values.

In our last chapter, we saw that T-score critical values are organized by degrees of freedom. For a one-population hypothesis test, the degrees of freedom is the sample size $- 1$ or " $n - 1$ ".



One-population mean T-test statistic example

An article in a health magazine claims that the mean average weight of all men is less than 180 pounds. A random sample of 40 men found that the sample mean was 172.55 pounds with a sample standard deviation of 26.327 pounds. Use a 10% significance level and the random sample data to test this claim.

$$H_0 : \mu = 180$$

$$H_A : \mu < 180 \text{ (Claim)}$$

Notice that the sample mean of 172.55 pounds is 7.45 pounds less than the population mean of 180 pounds. Is that significant?

The answer again is we don't know yet. We would need calculate a test statistic and a critical value to see if 7.45 pounds is a lot in this situation. Notice that since the alternative hypothesis is less than, this is a left-tailed test.

Here is the formula for calculating test statistics to compare sample and population means. Notice it follows the same general pattern and seeks to count how many standard errors the sample statistic is above or below the population parameter in the null hypothesis. In this case, the one-population T-test statistic calculates the number of standard errors that the sample mean (\bar{x}) is above or below the population mean (μ) in the null hypothesis. If the test statistic comes out negative it will be "below" and if the test statistic is positive, it will be "above". Notice also that we are using the same standard error estimation formula that we used for confidence intervals.

$$T = \frac{\text{(Sample Mean - Population Mean)}}{\text{Standard Error}} = \frac{(\bar{x} - \mu)}{\left(\frac{s}{\sqrt{n}}\right)}$$

To calculate the test statistic, plug in the sample mean $\bar{x} = 172.55$, the population mean $\mu = 180$, the sample standard deviation $s = 26.327$, and sample size $n = 40$.

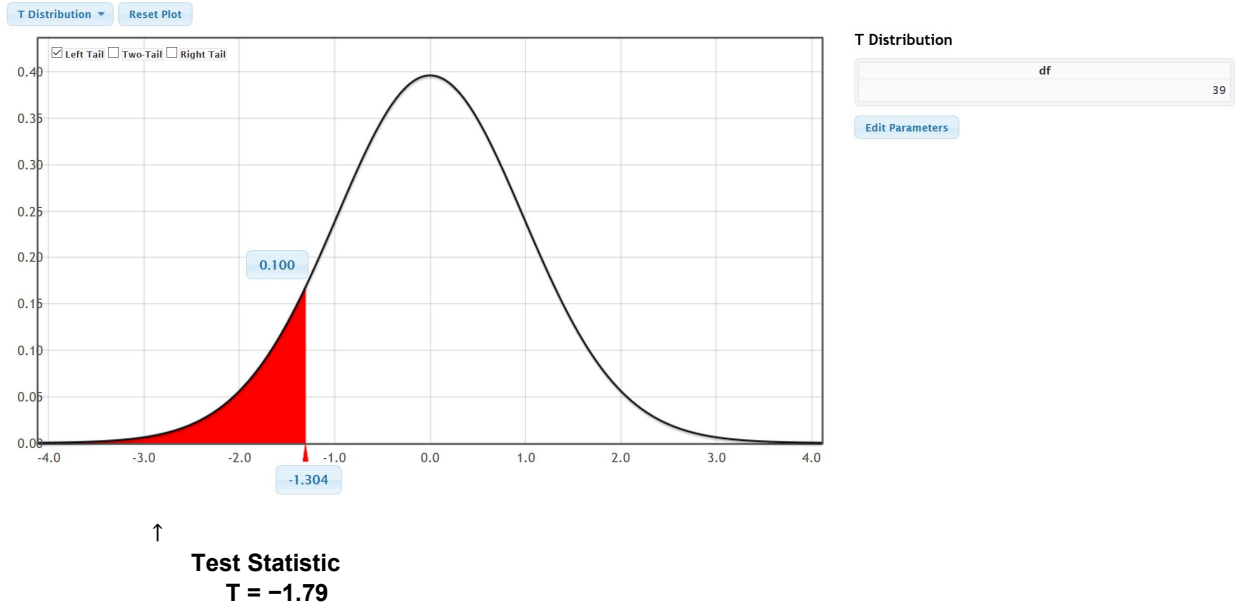
$$T = \frac{(\bar{x} - \mu)}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{(172.55 - 180)}{\left(\frac{26.327}{\sqrt{40}}\right)} \approx \frac{(-7.45)}{4.1627} \approx -1.79$$

Test-statistic sentence: The sample mean of 172.55 pounds is 1.79 standard error below the population mean of 180 pounds.

Is it significant?

Again, we will use StatKey to calculate the critical value. Go to the "Theoretical Distributions" menu and click on "T". In this problem the degrees of freedom is $40 - 1 = 39$. We also used a 10% significance level for this left-tailed test. After entering the degrees of freedom into StatKey, click on left tail, and put in 10% (0.10) for the tail proportion. Does the test statistic of $T = -1.79$ fall in the tail?





Notice that since the test statistic $T = -1.79$ is less than the critical value of -1.304 , the T-test statistic does fall in the left tail. So our sample mean of 172.55 pounds is significantly lower than the population mean of 180 pounds and the sample data significantly disagrees with the null hypothesis.

Statcato: Again, we can calculate the test statistic and critical value with a program like Statcato. Go to the “statistics” menu again and click on “hypothesis tests”. Now click on “one-population mean”. Under summarized data, put in the sample size of 40, sample mean of 172.55, and the sample standard deviation of 26.327. Change the significance level to 0.10, the alternative hypothesis to less than, and the hypothesized mean as 180. Then push ok. The printout is given below. Notice the test statistic is the same as our calculation and the critical value is the same as what we got with StatKey.

Hypothesis Test: 1-Population Mean ×

Help F1

<p>Inputs</p> <p><input type="radio"/> Samples in column: <input checked="" type="radio"/> Summarized sample data:</p> <p>Size: <input type="text" value="40"/></p> <p>Mean: <input type="text" value="172.55"/></p> <p>Standard deviation: <input type="text" value="26.327"/></p>	<p>Population Standard Deviation</p> <p>Population standard deviation:</p> <p><input type="radio"/> Known: <input type="text"/> - use z distribution</p> <p><input checked="" type="radio"/> Unknown - use t distribution</p>
<p>Significance</p> <p><input checked="" type="radio"/> Significance Level: <input type="text" value="0.10"/> 0 - 1.00 (e.g. 0.05)</p> <p><input type="radio"/> Confidence Level: <input type="text" value="0.90"/> 0 - 1.00 (e.g. 0.95)</p>	<p>Alternative Hypothesis</p> <p>Alternative Hypothesis: <input type="text" value="Less than"/></p> <p>Hypothesized Mean: <input type="text" value="180"/></p>



Hypothesis Test - One Population Mean: confidence level = 0.90

Input: Summary data

σ unknown (using t distribution)

Null hypothesis: $\mu = 180.0$

Alternative hypothesis: $\mu < 180.0$

N	Sample Mean	Stdev s	Significance Level	Critical Value	Test Statistic	p-Value
40	172.55	26.327	0.10	-1.304	-1.790	0.0406

One-population variance χ^2 -test statistic

When doing a hypothesis test about a population standard deviation (σ), we often prefer to use the population variance (σ^2). Remember, the variance is the square of the standard deviation. You can also think of the standard deviation as the square root of the variance. In a one population variance or standard deviation hypothesis test, we will compare the sample variance (s^2) to the population variance (σ^2).

One-population variance χ^2 -test statistic Example

While many people believe that the population standard deviation for men's weight is 20 pounds. This also implies that the population variance for men's weight is 400 square pounds. A random sample of 40 men has a sample standard deviation of 26.327 pounds and a sample variance is 693.111 square pounds. Let's use a 1% significance level for this problem. Notice that the test can be done using the standard deviation or the variance.

$$H_0 : \sigma = 20 \ (\sigma^2 = 400) \text{ (Claim)}$$

$$H_A : \sigma \neq 20 \ (\sigma^2 \neq 400)$$

Notice that the sample standard deviation of 26.327 pounds is 6.327 pounds more than the population standard deviation of 20 pounds.

Is that significant?

To determine this, we will need a test statistic. In our discussions about sampling distributions in the last chapter, we learned that the sampling distribution for variance is rarely normal and usually skewed to the right. For this reason we cannot use a Z or T-test statistic. These are based on normal sampling distributions. We saw in the last chapter that for one-population variance and standard deviation, we use the Chi-Squared distribution (χ^2) with a degrees of freedom $n - 1$.

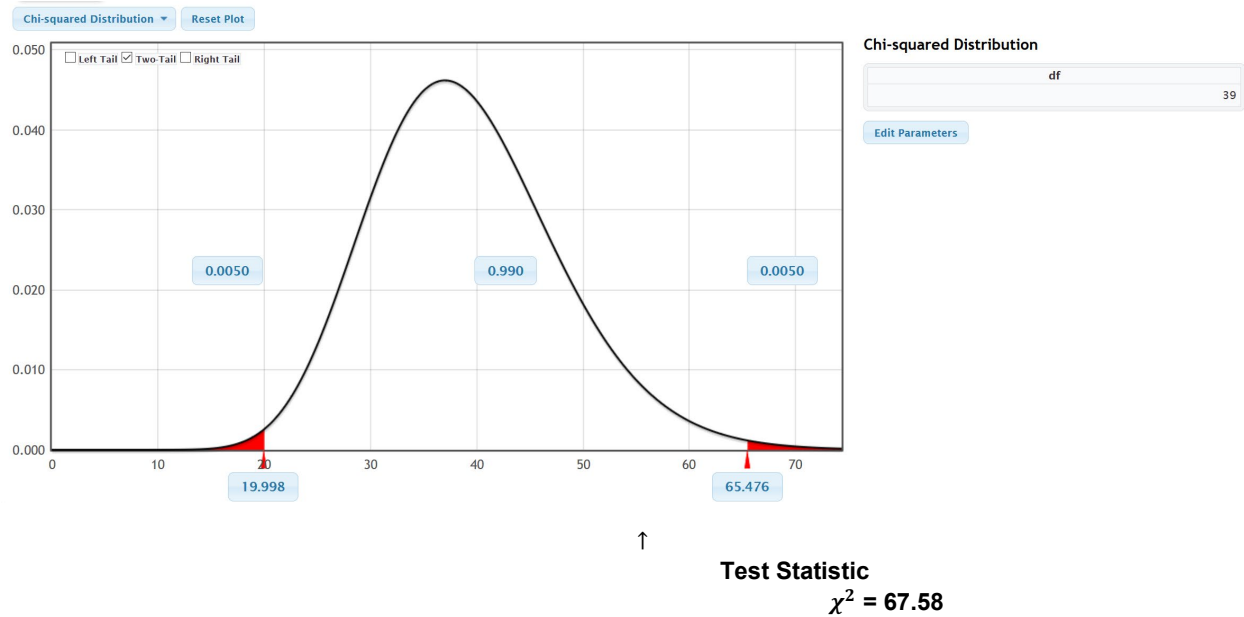
Chi-Squared One-population Variance Test Statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

Let's calculate the Chi-Squared test statistic for this problem. The degrees of freedom for this problem is $df = n - 1 = 40 - 1 = 39$. The sample standard deviation is $s = 26.327$, so the sample variance is $s^2 = 693.111$. The population standard deviation in the null hypothesis is 20 pounds, so the population variance is $\sigma^2 = 400$.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(40-1)26.327^2}{20^2} = \frac{(39)693.11}{400} \approx 67.58$$

Is this significant? Again, we will need to calculate the critical values. Let's go to StatKey at www.lock5stat.com. Under the theoretical distributions menu click on " χ^2 ". Put in the degrees of freedom 39 and click "Two Tail". Since this is a two-tailed test we will need to divide the significance level of 1% in half and put 0.5% (0.005) in each tail. What are the critical values? Does the test statistic fall in the tail?





First notice that the Chi-squared distribution is not perfectly normal. The F-distribution is very similar. Also the upper and lower critical values are not positive or negative like Z or T. It still works the same way though. We see that we have divided the significance level 0.01 in half and put 0.005 in each tail. The upper critical value is 65.476 and lower critical value is 19.998.

We also see that the χ^2 test statistic of 67.58 is higher than the upper critical value of 65.476, so falls in the right tail. The sample standard deviation significantly disagrees with the population standard deviation and the sample data significantly disagrees with the null hypothesis.

Statcato: Remember, it is always preferable to use technology to calculate test statistics. To calculate the χ^2 -test statistic and critical value with Statcato, go to the “statistics” menu, click on “hypothesis tests”, then click on “one-population variance”. Put in the sample size of 40 and either the sample standard deviation (26.327 pounds) or the sample variance. Change the significance to 1% (0.01). We also need to put in the population standard deviation (20) or the population variance and change the alternative hypothesis to “Not Equal to”. Now press “OK”.

Hypothesis Test: 1-Population Variance ×

Help F1

Inputs

Samples in column:

Summarized sample data:

Sample Size:

Variance:

Standard deviation:

Significance

Significance Level: 0 - 1.00 (e.g. 0.05)

Confidence Level: 0 - 1.00 (e.g. 0.95)

Alternative Hypothesis

Alternative Hypothesis:

Hypothesized variance:

Hypothesized standard deviation:



Hypothesis Test - One population variance: confidence level = 0.99

Input: Summary data

Null hypothesis: $\sigma^2 = 400.0$

Alternative hypothesis: $\sigma^2 \neq 400.0$

N	Sample Stdev s	Sample Var s^2	Significance Level	Critical Value	Test Statistic	p-Value
40	26.327	693.111	0.01	19.996, 65.476	67.578	0.0061

Notice the test statistic is close to what we calculated above and the critical values are very close to our StatKey calculation.

Important Notes:

- *It is important to understand how test statistics work and what the formulas represent. It is not important to calculate these by hand with a calculator. Statistics programs can calculate the test statistic quickly and with much better accuracy. It is important that you can explain the test statistic and what it tells us about significance.*
- *Whether or not a test statistic is significant can be difficult to interpret. Computer programs often give critical values to compare the test statistic to so that you can know if it was significant. It is important to draw a picture and visualize where the critical values are and where the tails begin. Here is the significance rule for test statistics and critical values.*

Significance Rule for Test Statistics and Critical Values

- **If the test statistic falls in the tail determined by the critical value (or values), then the sample data significantly disagrees with the null hypothesis.**
 - **If the test statistic does NOT fall in the tail determined by the critical value (or values), then the sample data does NOT significantly disagree with the null hypothesis.**
-

