


7.2: Claims on Population Means

Learning Objectives

- Test claims on population means both when σ is known and when σ is unknown
- Generalize the three forms of hypothesis tests
- Introduce, motivate, and utilize test statistics in computing p -values

 **Section 7.2 Excel File:** (contains all of the data sets for this section)

Means and Hypothesis Testing

Now that we have been introduced to the general logic of hypothesis testing, we will begin to address the particulars found within hypothesis testing based on the parameters of interest. We begin with testing hypotheses about the population mean. Just like our considerations of confidence intervals for means, we will have two cases to consider based on whether the population standard deviation is known or unknown. The latter case is the more frequently occurring case as we discussed in the chapter on confidence intervals, but we again begin with the case that the population standard deviation is known because of pedagogical considerations.

Recall the general process of hypothesis testing. A claim is made that warrants testing. This hypothesis may be derived from simple observation, past experimental data, a person, or an institution, and it tends to be the alternative hypothesis. A competing hypothesis, which tends to be the null hypothesis, is constructed. The statement of the alternative hypothesis determines the type of test to be conducted (either a one-tailed test or a two-tailed test). At this point, researchers generally settle on the level of the test (the probability of a type I error given that the null hypothesis is true). The next step in the process is designing the experiment to ensure that the test can actually be conducted. The calculation of the p -value depends on the sampling distribution of the sample statistics used to estimate the population parameter, assuming the null hypothesis is true. We need to ensure certain conditions are met. When testing claims on population means, we utilize the sampling distribution of sample means, which is normal when the underlying population is normal, and approximately normal for the most common distributions (recall our previous discussion on sampling distributions) when the sample size is larger than 30. We also need to ensure that our sample is randomly selected. Once we have designed the experiment, it must be conducted and analyzed. We then compute the p -value, the probability that at least as extreme as what was observed happens. If the p -value $< \alpha$ value, then we conclude that there is sufficient evidence to reject the null hypothesis. If the p -value $\geq \alpha$ value, then we conclude that there is not sufficient evidence to reject the null and hence fail to reject the null hypothesis. When sharing the results of the hypothesis test, include the p -value so others can also assess at the desired level of significance. With this succinct review, let us enter into testing claims on population means when σ is known.

Claims on Population Means (σ known)

Text Exercise 7.2.1

A bottling company is responsible for bottling 2 liter bottles of Dr. Pepper. The company policy regarding quality assurance is required to randomly sample 100 bottles each week to assess how well the bottles are being filled. The company assesses the test at a significance level of 0.01. If the company that built the bottling equipment guarantees that the machinery operates with a standard deviation of 0.1 liters and the last sample of 100 randomly chosen bottles had a sample mean of 1.98 liters, determine the hypotheses, make a conclusion regarding the test, and interpret the meaning within the context of the problem.

Answer

Since we are dealing with the amount of Dr. Pepper filled in the 2 liter bottles, we are dealing with population means. Each bottle has a certain amount of Dr. Pepper. Each bottle is supposed to have 2 liters of soda. So the population mean should be 2 liters. One hypothesis would be $\mu = 2$ liters. The bottling company wants to make sure that it is not overfilling (it does not want to shrink its profit margin) nor underfilling (it does not want to upset its customers and reface any false advertising lawsuits). We conclude that the other hypothesis would be $\mu \neq 2$ liters.

We now need to determine which hypothesis is to be the null hypothesis. If the company assumes that the machinery is not filling properly from the outset, it would automatically recalibrate the machinery every time; the collecting of evidence would be unnecessary. If the company acts as if the machines are not working properly when they really are, the company would unnecessarily waste production time. On the other hand, if the company acts as if the machines are filling properly

when they really are not, the company would produce bottles without the proper amount in them. If it was enough that customers would notice, the employees would likely notice before shipping them out. We, therefore, set the hypotheses as follows.

$$H_0 : \mu = 2 \text{ liters}$$

$$H_1 : \mu \neq 2 \text{ liters}$$

Since we had a random sample of 100 bottles, we have that the sampling distribution of sample means is approximately normal. Since we assume that the null hypothesis is true for the computation of the p -value, we have that the mean of the sampling distribution of sample means is $\mu_{\bar{x}} = 2$ liters. We also have that the standard deviation of the sampling distribution of sample means is $\sigma_{\bar{x}} = \frac{0.1}{\sqrt{100}} = \frac{0.1}{10} = 0.01$. Since our alternative hypothesis is that the mean is not equal to 2 liters, we are looking for evidence in two directions; we have a two-tailed test. Recall that the sample mean was 1.98 liters, which is less than 2 liters. We need to find the value that would be just as extreme except in the opposite direction. Since 1.98 liters is 0.02 liters below the hypothesized value, the other value we are looking for is 0.02 liters above the hypothesized value, namely 2.02 liters. To find the p -value, we compute the area in the left tail ending at 1.98 liters and the right tail starting at 2.02 liters. See the figure below for a visual.

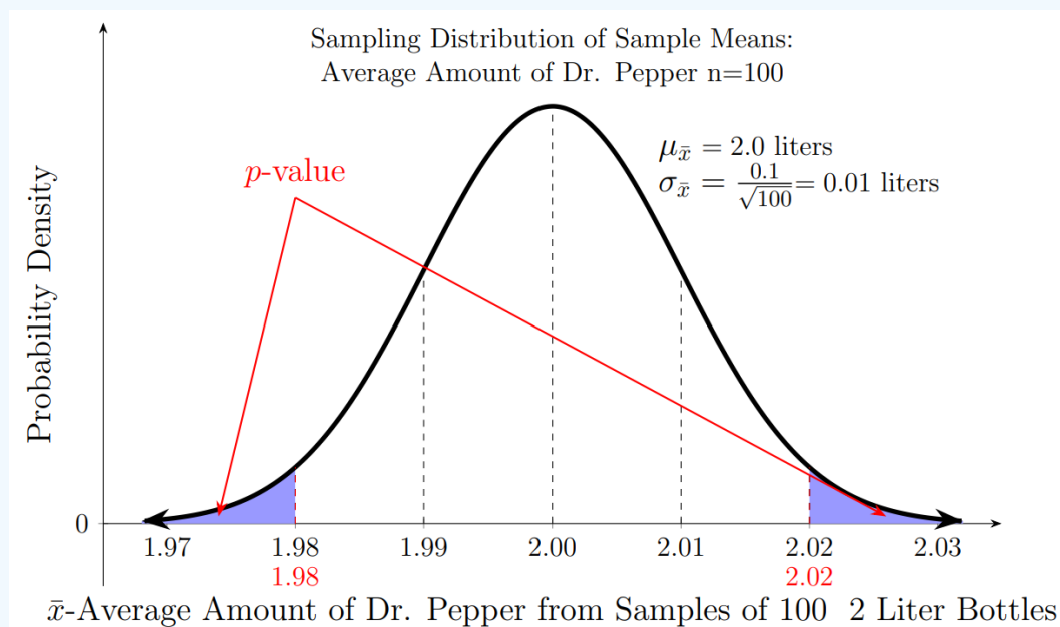


Figure 7.2.1 Sampling distribution of sample means

We compute the p -value using technology. Note that since the sampling distribution of sample means is approximately normal, the area in the two tails is equal due to symmetry. This eases the calculation. For this first exercise, we compute it both ways.

$$\begin{aligned} p\text{-value} &= \text{NORM.DIST}(1.98, 2, 0.01, 1) + (1 - \text{NORM.DIST}(2.02, 2, 0.01, 1)) \\ &\approx 0.02275 + (1 - 0.97725) \\ &\approx 0.0455 \\ p\text{-value} &= 2 \cdot \text{NORM.DIST}(1.98, 2, 0.01, 1) \\ &\approx 2 \cdot 0.02275 \\ &\approx 0.0455 \end{aligned}$$

Since 0.0455 is not less than 0.01, we do not have sufficient evidence to reject the null hypothesis that the machines are filling the 2 liter bottles properly. The machines, therefore, pass the weekly test for quality assurance.

? Text Exercise 7.2.2

In 2024, researchers at the University of Maryland edited the genes of poplar trees to reduce the amount of lignin naturally present in the tree. This is desirable because the process of strengthening wood involves heat and compression, and the more lignin present in the tree, the harder the tree is to compress. The amount of lignin present in a poplar tree is often reported as a percent of the tree's dry weight. The average poplar tree has 27% of its dry weight due to the weight of lignin.

Suppose that independent researchers want to verify the claim. They request to randomly sample 45 of the many thousands of genetically altered poplar trees growing across the various stations the University of Maryland researchers utilize. For the sake of open and honest scientific research, their requests are granted. If it is known that the standard deviation of the percent of weight of lignin in all poplars, including this genetically altered poplar, is 4.2% and the sample mean was 25.3%, conduct the hypothesis test at the 0.005 level of significance.

Answer

The measure of interest for each tree is the percent of dry weight that is due to the presence of lignin. Despite the fact that this measurement returns a percent, we are not considering proportions as our parameter of interest. We are interested in the population mean of the percent weight due to lignin measurements in the genetically altered poplar trees. It is known that the typical poplar tree has 27% of its dry weight due to the weight of lignin, and the researchers at the University of Maryland think that they have reduced the amount of lignin naturally present in the genetically altered poplar tree. We naturally obtain the hypothesis that $\mu < 27$ percent. From here, we identify the opposing hypothesis that $\mu \geq 27$ percent. Note that this is a one-tailed test. We conduct a one-tailed test because regardless of whether the genetically altered poplar trees have the same amount of lignin or more lignin than the regular poplar trees, interest in these genetically altered trees would fade. There is no need to distinguish between no change and a change for the worse.

To select which hypothesis is to be considered the null hypothesis, we note that we are conducting the experiment to test that the claims of the University of Maryland researchers are true; so, we do not want to assume their conclusion from the beginning. We will set the null hypothesis to say that the genetically altered poplars have at least as much lignin present as a percent of their dry weights as regular poplar trees.

$$H_0 : \mu \geq 27 \text{ percent}$$

$$H_1 : \mu < 27 \text{ percent}$$

We now begin to look at the sampling distribution of sample means given the size of the random sample and the assumption that the null hypothesis is true. Since the sample size is 45, we expect the sampling distribution of sample means to be approximately normal. Since there are many possible population means under the assumption that the null hypothesis is true, we conduct our study with the value of μ that will produce the largest p -value. This occurs when $\mu = 27$ percent.

We have $\mu_{\bar{x}} = 27$ percent and $\sigma_{\bar{x}} = \frac{4.2}{\sqrt{45}} \approx 0.6261$ percent. We need to determine what would be considered at least as extreme as the evidence from the sample which produced a sample mean of 25.3 percent. In this case, the more extreme would be smaller and smaller percentages. So we are looking for the area in the tail on the left side of the sampling distribution that ends at 25.3 percent. Notice how the direction of the tail matches the direction of the inequality in the alternative hypothesis; we call this a left-tailed test. See the figure below for a visualization.

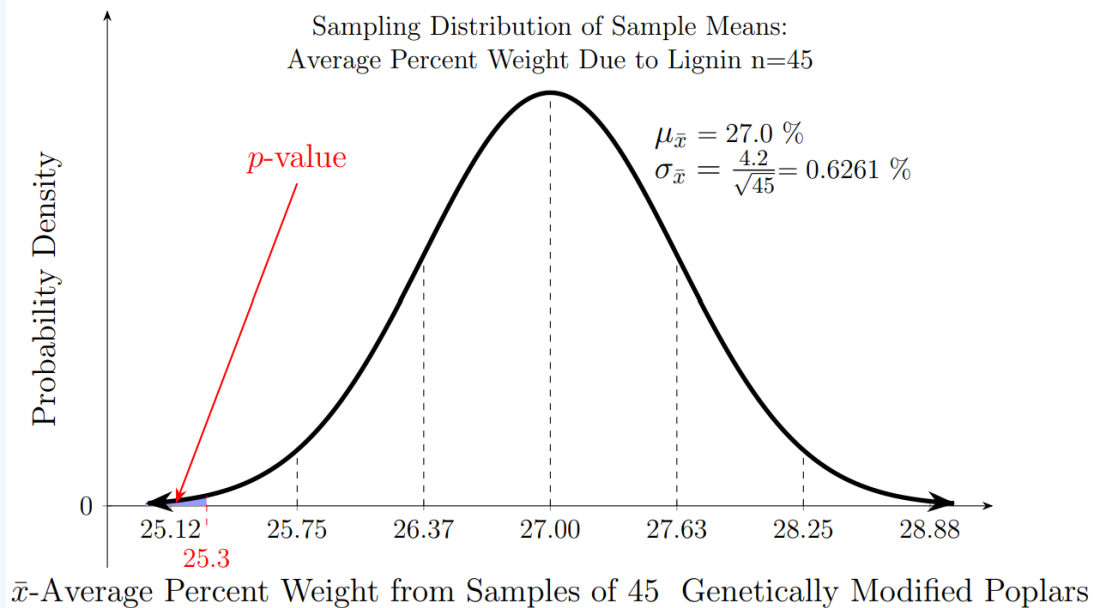


Figure 7.2.2 Sampling distribution of sample means

We thus compute the p -value using technology.

$$p\text{-value} = \text{NORM.DIST}(25.3, 27, \frac{4.2}{\sqrt{45}}, 1) \approx 0.0033$$

Since $0.0033 < 0.005$, there is sufficient evidence to reject the null hypothesis in favor of the alternative hypothesis that the genetically altered poplar trees have less lignin naturally present. There is sufficient evidence in support of the claims of the researchers at the University of Maryland.

Note: Hypotheses and p -values

At this stage, we have conducted both types of one-tailed tests: right-tailed (recall the last section with corn yields) and left-tailed (this section with lignin). When we operated under the assumption of the truth of the null hypothesis, we had to decide which value of μ to use in computing the p -value. We chose the value that would be the hardest to find sufficient evidence (the value that would produce the largest p -value). This turned out to be when μ equaled the common or accepted value from the problem (the 70 bushels per acre of the commonly used corn and the 27% of the regular poplar trees). This will always be the case and contributes to some textbooks making the pedagogical choice to say that the null hypothesis is always of the form: the parameter equals the standard value. We do not adopt this simplification for the form of the null hypothesis, but we do emphasize that when the parameter equals the standard value, the p -value produced will be the largest, which is essential for our analyses.

We also want to emphasize that there are three tests to consider: left-tailed tests, right-tailed tests, and two-tailed tests. These three tests correspond to the three possible forms of hypotheses. Consider looking at the alternative hypotheses. In this last text exercise about poplar trees, the alternative hypothesis was $\mu < 27$ percent, and we computed the p -value by calculating the area in the left tail of the distribution. In the text exercise about Dr. Pepper, the alternative hypothesis was $\mu \neq 2$ liters, and we considered the area in both tails. And finally, in the exercise about corn yields, the alternative hypothesis was $\mu > 70$ bushels per acre, and we looked only at the area in the right tail. Notice that the direction of the tail matches the direction of the inequality sign in the alternative hypothesis. All of this is provided in the figure below within the context of means but applies to any parameter (μ_0 indicates the common or accepted value of the population mean).

left-tailed test	two-tailed test	right-tailed test
$H_0 : \mu \geq \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu \leq \mu_0$
$H_1 : \mu < \mu_0$	$H_1 : \mu \neq \mu_0$	$H_1 : \mu > \mu_0$
use $\mu = \mu_0$ to compute p -value	use $\mu = \mu_0$ to compute p -value	use $\mu = \mu_0$ to compute p -value

Test Statistics

Recall that we can transform any normal distribution into the standard normal distribution and that this transformation preserves area. An implication of these facts is that the computation of p -values can be done within the context of the standard normal distribution once we transform the particular sample mean. We call this transformed value of the calculated sample statistic the **test statistic**. As discussed, different transformations can change sampling distributions of particular sample statistics to particular common distributions. For now, understand that the basic idea of the test statistic is that it is a value that represents the value of the sample statistic computed from the actual sample collected which facilitates computing the p -value.

Recall that the z -score transformation sends any normal distribution with a mean μ and a standard deviation σ to the standard normal distribution given by the formula below.

$$z = \frac{x - \mu}{\sigma}$$

? Text Exercise 7.2.3

Repeat the hypothesis tests from Text Exercises 7.2.1 and 7.2.2 using test statistics to compute the p -values. Verify that the same p -values are computed which in turn yield the same conclusions as before.

1. The first text exercise considered filling 2 liter bottles of Dr. Pepper. A sample of 100 2 liters was randomly chosen which produced a sample mean of 1.98 liters. The population standard deviation was 0.1 liters. The hypothesis test was to be conducted on the hypotheses below at the $\alpha = 0.01$ level of significance.

$$H_0 : \mu = 2 \text{ liters}$$

$$H_1 : \mu \neq 2 \text{ liters}$$

Answer

All conditions to conduct a hypothesis test are met; for details, review Text Exercise 7.2.1. We need to compute our test statistic. We assume that the null hypothesis is true and, therefore, know that the sampling distribution is approximately normal with $\mu_{\bar{x}} = 2$ liters and $\sigma_{\bar{x}} = \frac{0.1}{\sqrt{100}} = 0.01$ liters. Since we are transforming the sampling distribution of sample means with σ known into the standard normal distribution our z -transformation takes on the form below.

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

We can insert the values from our particular context to arrive at the following.

$$z = \frac{1.98 - 2}{\frac{0.1}{\sqrt{100}}} = \frac{-0.02}{0.01} = -2$$

Understand the -2 value to mean that under the assumption of the truth of the null hypothesis the evidence that we collected from the sample mean is 2 standard deviations below the hypothesized population mean.

The alternative hypothesis contains the \neq sign implying a two-tailed test. We need to consider what is equally extreme in the opposite direction. Since the standard normal distribution is centered at 0, all we have to do is take the value equal in magnitude and opposite in sign, 2. Two values are equally extreme if they are the same number of standard deviations away from the population mean. We have the following visualization for computing the p -value.

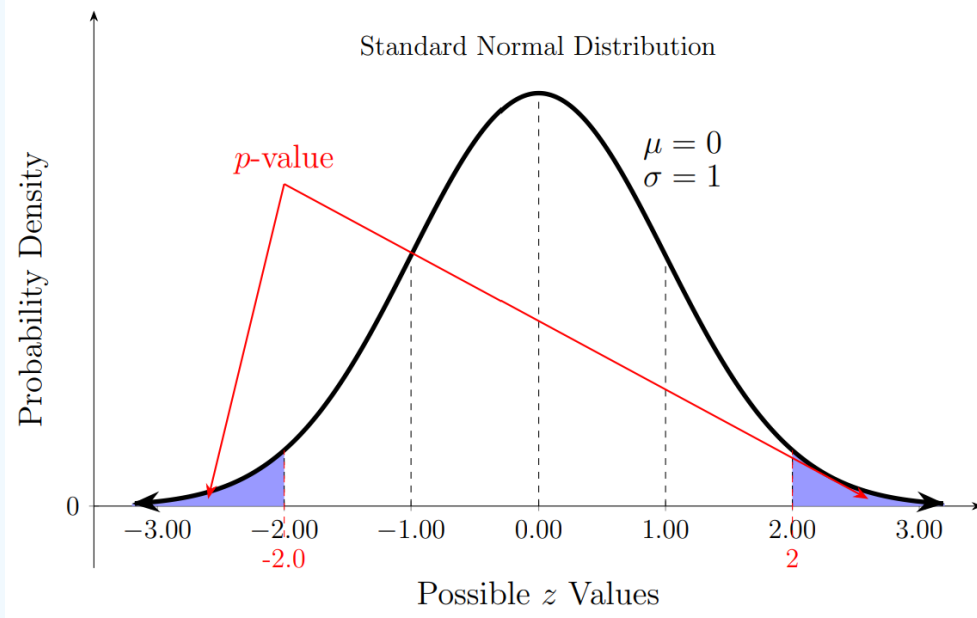


Figure 7.2.3 Standard normal distribution

$$\begin{aligned}
 p\text{-value} &= \text{NORM.S.DIST}(-2, 1) + (1 - \text{NORM.S.DIST}(2, 1)) \\
 &\approx 0.02275 + (1 - 0.97725) \\
 &\approx 0.0455 \\
 p\text{-value} &= 2 \cdot \text{NORM.S.DIST}(-2, 1) \\
 &\approx 2 \cdot 0.02275 \\
 &\approx 0.0455
 \end{aligned}$$

We arrive at the same p -value as before and since 0.0455 is not less than 0.01, we fail to reject the null hypothesis.

- The second text exercise considered the percent weight of poplar trees due to lignin. A sample of 45 genetically altered poplar trees was randomly chosen, producing a sample mean of 25.3 percent. The population standard deviation was 4.2 percent. The hypothesis test was conducted on the hypotheses below at the $\alpha = 0.005$ significance level.

$$\begin{aligned}
 H_0 &: \mu \geq 27 \text{ percent} \\
 H_1 &: \mu < 27 \text{ percent}
 \end{aligned}$$

Answer

All of the conditions to conduct a hypothesis test are met; for details, review Text Exercise 7.2.2. We need to compute our test statistic. We assume that the null hypothesis is true and, therefore, know that the sampling distribution is approximately normal with $\mu_{\bar{x}} = 27$ percent and $\sigma_{\bar{x}} = \frac{4.2}{\sqrt{45}} \approx 0.6261$ percent. We can compute our test statistic.

$$z = \frac{25.3 - 27}{\frac{4.2}{\sqrt{45}}} \approx \frac{-1.7}{0.6261} \approx -2.7152$$

We can understand the -2.7152 value to mean that assuming the truth of the null hypothesis the evidence that we collected from the sample mean is 2.7152 standard deviations below the hypothesized population mean.

Since the alternative hypothesis contains the $<$ sign, this hypothesis test is a one-tailed test, and more extreme values would be values to the left, resulting in a left-tailed test. We have the following visualization for computing the p -value.

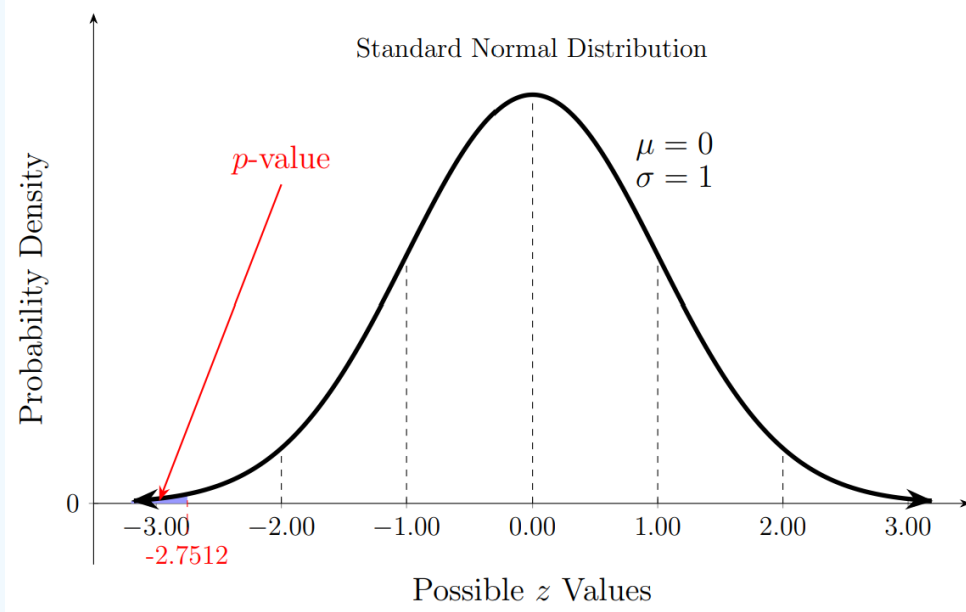


Figure 7.2.4 Standard normal distribution

$$p\text{-value} = \text{NORM.S.DIST} \left(\frac{-1.7}{\frac{4.2}{\sqrt{45}}}, 1 \right) \approx \text{NORM.S.DIST} (-2.7152, 1) \approx 0.0033$$

We again produce the same p -value which yields that there is sufficient evidence to reject the null hypothesis in favor of concluding the alternative hypothesis: the genetically altered popular trees have less lignin naturally present.

Claims on Population Means (σ unknown)

We are now prepared to move to the more common situation: testing hypotheses about population means when the population standard deviation is unknown. The added complication comes from the fact that we do not know the standard deviation of the sampling distribution. We can estimate the population standard deviation using the sample standard deviation from our collected sample, but using this estimate has ramifications.

Recall constructing confidence intervals for population means when σ was unknown, we considered what happens under the t -transformation.

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

We concluded that the t variable followed a particular distribution, the Student's t -distribution with $n - 1$ degrees of freedom. Notice how similar the formula for the variable t is to the formula for calculating the test statistic when σ is known. The only difference is that one formula has an s while the other has a σ . Just as the z -score transformation provided the formula to calculate the test statistic when σ is known, the t -transformation provides the formula we use to compute the test statistic when σ is unknown. We, therefore, know the distribution of test statistics when σ is unknown is the t -distribution with $n - 1$ degrees of freedom. We use this fact to compute the p -value for testing claims on population means when σ is unknown. For a refresher on the t -distribution see [Section 6.4](#). The two processes for testing hypotheses about population means are very similar. The main difference is the distribution in which we calculate the p -value. When σ is known, we use the standard normal distribution. When σ is unknown, we use the t -distribution with $n - 1$ degrees of freedom.

Before we test hypotheses about population means with σ unknown, let us review the overall process of hypothesis testing to reinforce the procedure and highlight the distinctions between various situations.

1. Use natural observation, previous experimental results, or the claims of others to formulate a hypothesis that warrants testing.
Within the context of means, each observation must admit some quantitative fact that can be measured and averaged. This

excludes considerations of whether or not observations have a particular quality; that will be studied in the section on claims on population proportions.

2. Identify a competing hypothesis and consider the ramifications of acting as if one of the hypotheses is true when, in fact, it is not. Name the hypothesis with the less drastic ramifications as the null hypothesis. The novel or claimed hypothesis is generally the alternative hypothesis. See note in the previous section regarding other ways to help distinguish between null and alternative hypotheses.
3. Determine the methodology of collecting evidence against the null hypothesis and determine what constitutes sufficient evidence by setting the level of significance. Make sure the design meets the requirements of the test intended to be conducted. For claims on population means, ensure that the sample is randomly selected and that either the underlying population is normally distributed or that the sample is large enough that the sampling distribution of sample means is approximately normal. In most cases, $n > 30$ will be sufficient.
4. Conduct the experiment and collect the evidence.
5. Compute the test statistic. Be sure to make the distinction between sample and population standard deviations. The most common situation is that we only have access to the sample standard deviation s and, therefore, must use the t -transformation to compute our test statistic. We often denote the test statistic based on which transformation is used. If the population standard deviation is known, the z -score transformation is used, and the test statistic is denoted with a z . If the population standard deviation is unknown, the t -transformation is used and the test statistic is denoted with a t .
6. Use the hypotheses to determine whether a test is a left-tailed, right-tailed, or two-tailed test. Note that the directions match the sign in the alternative hypothesis.
7. Determine the p -value by considering the test statistic, the appropriate distribution, and the type of test and then using technology to make an appropriate calculation.
8. Compare the p -value to the α value. If the p -value $< \alpha$ value, then we reject the null hypothesis in favor of the alternative hypothesis. If the p -value $\geq \alpha$ value, we fail to reject the null hypothesis.

? Text Exercise 7.2.4

A guest speaker at a local library presented on the change in human physical characteristics over the last two centuries in the United States. The presenter claimed that male height has consistently increased over that time and will continue to do so. The last evidence cited in this regard was in 1970, stating that the average height of adult males was 176 centimeters. Given the span of over 50 years, we decide to test the hypothesis that the average height of adult males in 2024 has increased since 1970. Suppose we collect a sample of 16 adult males randomly selected and measure their heights in centimeters. The data is presented below. We decide to conduct the hypothesis test at an α value of 0.05.

169, 171, 171.5, 173, 173, 174.25, 174.5, 175, 176, 177, 177.75, 178, 178, 179, 179.5, 180

Answer

We must confirm that our circumstances enable a hypothesis test to be conducted; we need a random sample and reasonable confidence that the shape of the sampling distribution of sampling means is approximately normal. The first component, the random sample, is easily confirmed. The sample size chosen is 16, which does not meet the typical threshold of more than 30. We must recall that male and female adult height are normally distributed. We, therefore, have that the sampling distribution of sample means is normally distributed. We can conduct the test. Note that no population standard deviations are given. We must utilize the recently discussed method that involves the t -distribution!

The guest speaker made the claim that the average height of adult males has increased over time. The population mean was 176 centimeters back in 1970; so, we identify one of the hypotheses as $\mu > 176$ centimeters. The opposite hypothesis would thus be $\mu \leq 176$ centimeters. This is a one-tailed test because both the average being the same or smaller than before are equally detrimental to the claims of the guest speaker. In both cases, the ramifications of acting as if one hypothesis is true when it is false seem to be mild. So, we pick the null hypothesis to be the one contrary to the claimed hypothesis. We settle on the hypotheses as follows.

$$H_0 : \mu \leq 176 \text{ centimeters}$$

$$H_1 : \mu > 176 \text{ centimeters}$$

Given the hypotheses, we have a right-tailed test. We compute our test statistic.

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

To do this, we must compute the sample mean and sample standard deviation using the values collected from our random sample. We produce the following results: $n = 16$, $\bar{x} \approx 175.4063$ centimeters, and $s \approx 3.2887$ centimeters.

At this stage, we notice that our sample mean is less than the hypothesized population mean. Since we have a right-tailed test, we are looking for evidence against the null hypothesis in the form of sample means larger than the hypothesized value. We can thus immediately conclude that there is not sufficient evidence to reject the null hypothesis. We will show the remainder of the computation to solidify the process and strengthen our conclusion.

$$t \approx \frac{175.4063 - 176}{\frac{3.2887}{\sqrt{16}}} \approx -0.7222$$

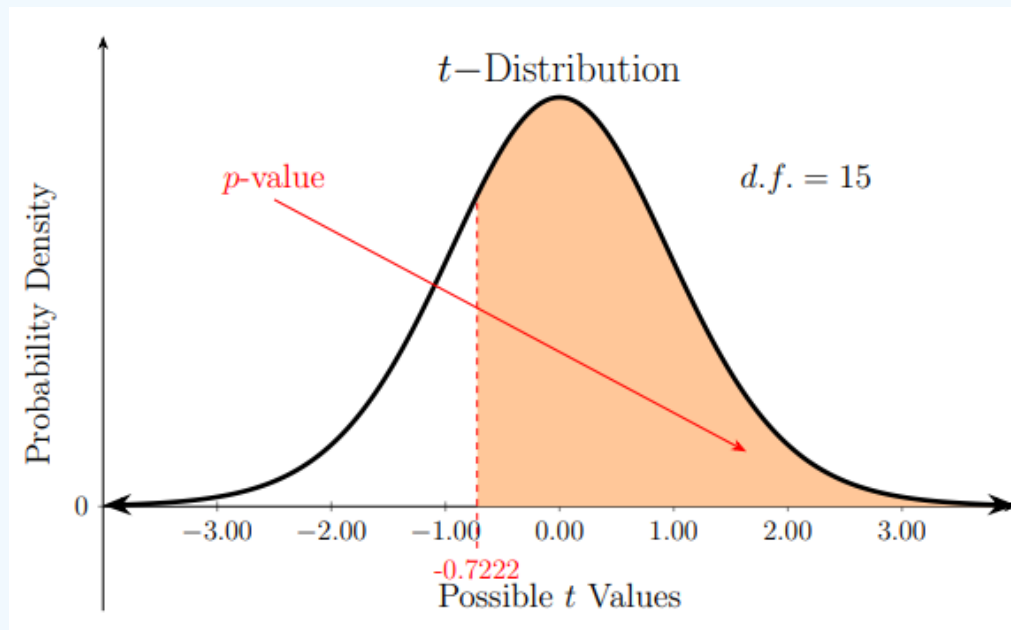


Figure 7.2.5 Right-tailed test with $t = -0.7222$ using t -distribution with 15 degrees of freedom

As we can see, the shaded area is over half of the area due to the symmetry of the t -distribution. We now compute the p -value using technology.

$$p\text{-value} \approx 1 - \text{T.DIST}(-0.7222, 15, 1) \approx 1 - 0.2406 \approx 0.7594$$

The p -value is larger than the α value of 0.05. We conclude the test by failing to reject the null hypothesis. There is not sufficient evidence to support the guest speaker's claim.

It is still possible that the average height of adult males has increased, and something rare occurred in the act of sampling. It is also possible that the average height is the same as it was. Based on the evidence, we may also be open to the idea that the average height may be smaller. Reviewing the guest speaker's reasoning behind why the heights have been increasing may be prudent. Is there a faulty assumption? Is there an explanation as to why it may have been increasing and now possibly decreasing? Perhaps we can conduct another random experiment to test whether the mean is now less than it was before. Hypothesis tests that fail to reject the null hypothesis can still inform further research and inquiry.

? Text Exercise 7.2.5

A [study](#) published in 2021, concluded that the average weekly recreational screen time of 18 – 29 year olds (emerging adults) increased from 2018 to 2020 during the pandemic estimating the average weekly recreational screen time with the confidence interval 28.5 ± 11.6 hours. Recreational screen time does not include screen time associated with work or school.

The pandemic is now behind us, but the effects of the pandemic are still playing out. A researcher is still interested in the weekly recreational screen time of emerging adults and conducts a study on 56 randomly selected emerging adults with a sample mean of 30.4 hours and a sample standard deviation of 6.3 hours. The researcher adopts the following hypotheses to be tested at the 0.05 level of significance. Conduct the test.

$$H_0 : \mu = 28.5 \text{ hours}$$

$$H_1 : \mu \neq 28.5 \text{ hours}$$

Note that this sample data is completely fabricated.

Answer

Having been given the hypotheses, we note that we are conducting a two-tailed test. The researcher adopted the central value of confidence interval from the study to compare the current data. We do not know the population standard deviation and thus operate within the realm of the t -transformation and t -distribution.

$$t = \frac{30.4 - 28.5}{\frac{6.3}{\sqrt{56}}} \approx 2.2569$$

Since we have a two-tailed test, we determine the test-statistic that is equally as extreme as the computed test statistic in the opposite direction. Again due to symmetry, this is the value of equal magnitude but opposite sign. We have the following visualization for computing the p -value.

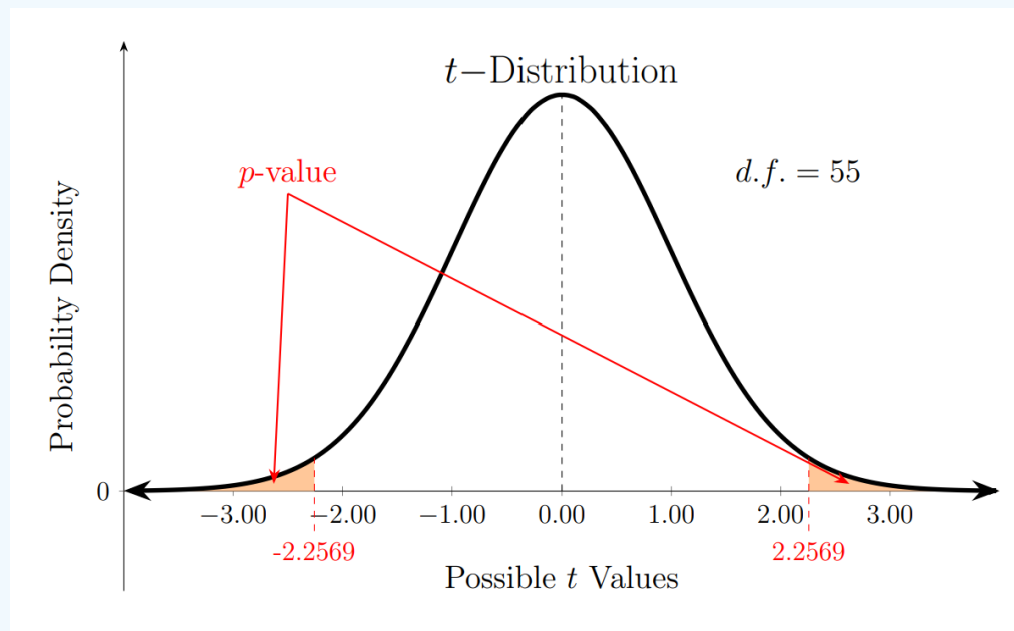


Figure 7.2.6 Two-tailed test with $t = \pm 2.2569$ using t -distribution with 55 degrees of freedom

$$p\text{-value} = 2 \cdot T.DIST(-2.2569, 55, 1) \approx 2 \cdot 0.0140 \approx 0.0280$$

We again used symmetry, noting that the boundaries of the tails are equidistant from 0, so that we can double the area found in one of the tails. The left tail can be computed directly using the negative value of the two test statistics.

We compare 0.0280 with 0.05 and find that $0.0280 < 0.05$. We have sufficient evidence to reject the null hypothesis that the average weekly recreational screen time for emerging adults in 2024 is 28.5 hours and conclude that the average weekly recreational screen time for emerging adults in 2024 is not 28.5 hours. The evidence indicates that the average may actually be higher, but to reach such a conclusion, another study must be conducted.

- [12.1: Testing a Single Mean](#) by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.