

6.4: Confidence Interval for Means (Sigma Unknown)

Learning Objectives

- Form a basic intuition regarding the development and shape of t -distributions
- Introduce accumulation functions for t -distributions
- Discuss degrees of freedom
- Find critical values
- Construct confidence intervals for means using sample data
- Estimating necessary sample sizes for desired margin of errors

▮ [Section 6.4 Excel File](#) (contains all of the data sets for this section)

Confidence Intervals for Means

Having developed a construction technique for confidence intervals for mean with σ known, we now drop our simplifying assumption and address the common case when we do not know much about the population; in particular, we do not know the population mean or the population standard deviation. Having an idea of what the sampling distribution of sample means looks like is paramount to our method. We must know that the sampling distribution of sample means is approximately normal. We encouraged the reader to review the section on sampling distributions of sample means in the last chapter; hopefully, the conditions are committed to memory, but we provide them now for quick reference: either the parent population is normal or the sample size n is greater than 30. Recall that this threshold works for many populations but not all.

The sampling distribution of sample means is approximately normal with a mean, $\mu_{\bar{x}} = \mu$, and a standard deviation, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Now, we do not know the value σ in addition to not knowing the value of μ . We set a level of confidence, which we understand as the percent of samples of size n that will produce a sample mean within a certain distance of the population mean; we call this distance the margin of error, ME. In the previous section, we determined the margin of error by transforming the sampling distribution of sample means into the standard normal distribution and finding our critical values. At this stage, we now run into difficulties because we do not know σ . We can approximate with the following transformation, which we call the t -transformation.

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

Notice, that we simply substituted σ with s , the sample standard deviation. Think about the ramifications of this; different samples will naturally produce different values for s . Thus, when we evaluate the t -transformation, we do not expect to get the same values that the z -score transformation would produce since that transformation used σ for every computation. We might expect that the distribution that the sampling distribution of sample means will not be normally distributed.

We can build the theory just as we did with sampling distributions: by examining populations where we have all the data, computing the value of the transformation for each sample, constructing a histogram, and identifying the general shape. The theory is formalized, just as with sampling distributions, with some sophisticated mathematics beyond the scope of this course, but we will, hopefully, build a basic intuition by trying to understand how the t -transformation compares to the z -score transformation.

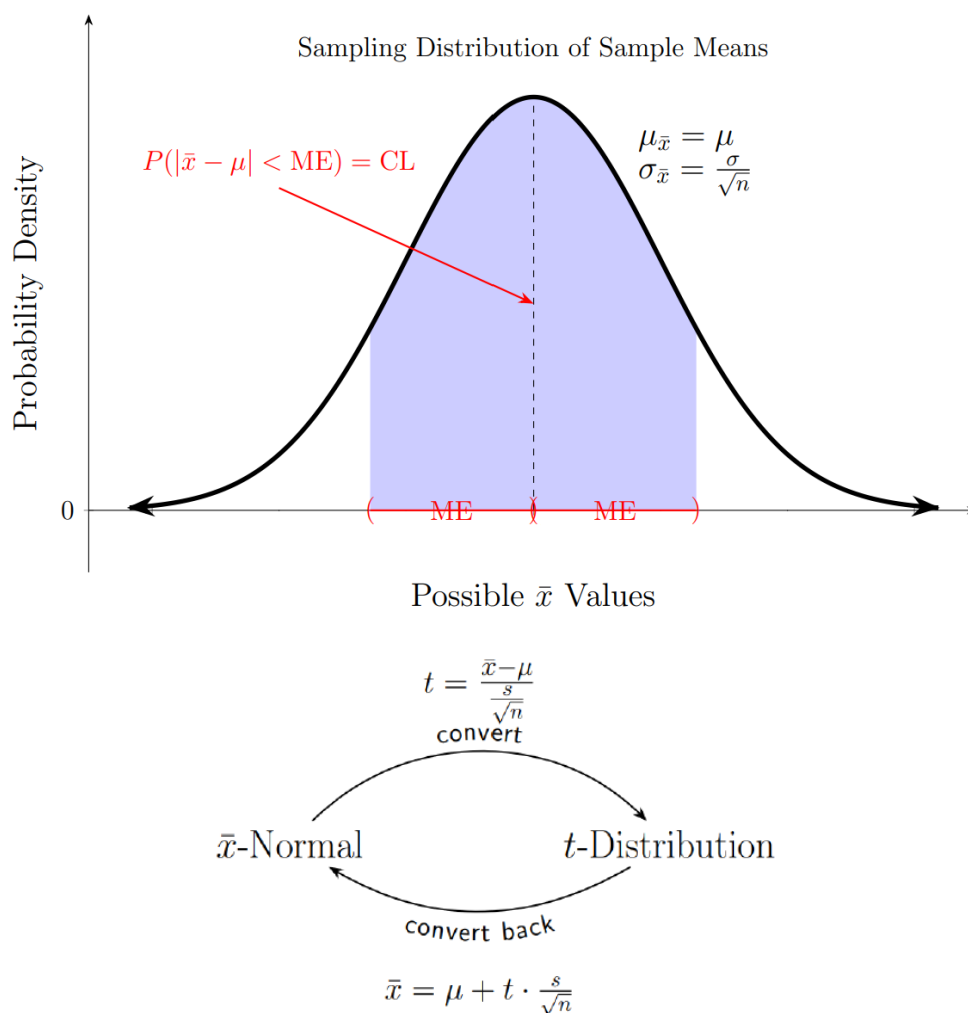
The sampling distribution of sample means is approximately normal, and normal distributions are symmetric, meaning, half of the samples of a given size produce sample means greater than μ and the other half of the samples produce sample means less than μ . We can also have samples with the same sample mean but with quite different standard deviations. So, we need to understand how these sample standard deviations are distributed. We are interested in the probability that our sample standard deviation is less than the population standard deviation. Note that the probability that the sample standard deviation is less than the population standard deviation is the same as the sample variance being less than the population variance. The sampling distribution of population variance is closely associated with the χ^2 -distribution (an interested reader is encouraged to read or reread [the sampling distribution of sample variances](#) section for further details) which we have [seen to be skewed right](#). Recall that when a distribution is skewed right, the mean is greater than the median and the probability that the sample variance is less than the population variance is greater than 50%. So, we are more likely to get sample standard deviations that are smaller than the population standard deviation than to get larger sample standard deviations.

What does this all say about the t -transformation? Since we are just as likely to get sample means larger or smaller than the mean, we will be symmetric about 0. Since we are more likely to get sample standard deviations that are smaller than the population standard

deviation, we will usually be dividing by a smaller number in the t -transformation than in the z -score transformation. When dividing by smaller numbers, the quotient is larger. We expect larger magnitudes under the t -transformation than under the z -score transformation. This indicates that there is a greater probability density in the tails (the distribution has thicker tails). Hopefully, at this point, we recognize these descriptions as our descriptions of the [Student's \$t\$ -distribution](#). This t -distribution will have $n - 1$ degrees of freedom, and we will explain why later in the section.

Determining the Margin of Error (σ unknown)

Now that we have built an intuitive understanding of the t -distribution, we are prepared to determine the necessary margin of error in the context of not knowing the population standard deviation. We provide a similar progression of figures to illustrate the process below.



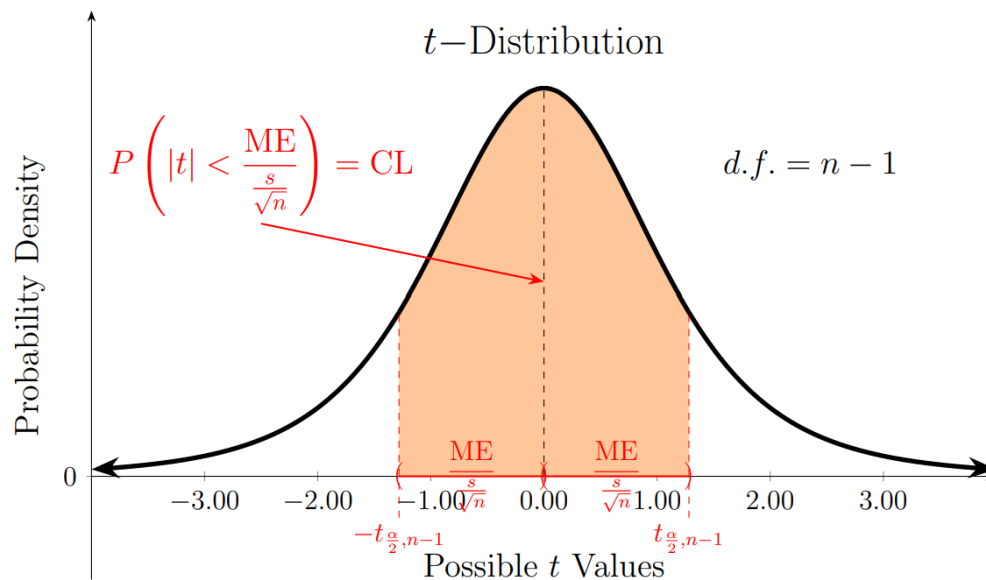


Figure 6.4.1: Sampling distribution of sample means under the t -transformation

Knowing that the sampling distribution of sample means is transformed into a t -distribution with $n - 1$ degrees of freedom enables us to determine the necessary margin of error for our desired confidence level. Just as when σ is known, the margin of error is scaled by a known factor known to us. And once we determine the boundary points of our shaded region, the critical values $\pm t_{\frac{\alpha}{2}, n-1}$, we can compute the margin of error. Notice the additional subscripts in the notation for our critical values of the t -distribution due to the critical values depending on the confidence level and the degrees of freedom. Once again, we must use technology to determine critical values with an accumulation functions specific to the t -distribution.

In Excel, we utilize the T.DIST and T.INV functions which work very similarly to the NORM.DIST and NORM.INV functions that we have been working with for quite some time. We use the distribution function to find the area to the left of a point. Using T.DIST, we enter the point that we want the area to the left of, and the necessary information to describe the distribution; for normal distributions, we used the mean and standard deviation, but for t -distributions, we use the degrees of freedom. Finally, we then tell the function to accumulate.

$$P(t < a) = \text{T.DIST}(a, d.f., 1)$$

The T.INV function is used to find the point in a t -distribution with a certain number of degrees of freedom such that a given area is to the left of the point. Using T.INV, we enter the area and the degrees of freedom.

$$a = \text{T.INV}(\text{area to the left of } a, d.f.) = \text{T.INV}(P(t < a), d.f.)$$

? Text Exercise 6.4.1

Remaining in the context of constructing confidence intervals for population means when σ is unknown, determine the positive critical value for the 95% confidence level given the indicated sample size by roughly sketching the t -distribution and then using technology.

The rough sketches drawn by hand are important to ensure a proper approach to the problems, but we cannot accurately depict what happens to the distributions as the degrees of freedom change. As you complete each part of this text exercise, examine the computer-generated graphics to solidify what happens as the degrees of freedom increase.

1. $n = 4$

Answer

We begin by sketching the t -distribution with 3 degrees of freedom. It is symmetric about 0 and has thicker tails than the standard normal distribution. We then form an interval centered at 0 and label the boundary points. The area under the curve between these two points is our confidence level. Notice that the critical values are equal in magnitude but opposite in sign.

Due to the symmetry of the distribution, the two tails have equal area giving $\frac{\alpha}{2}$ in each tail. Since our confidence level is 95%, the α value is 0.05 meaning the area in each tail is 0.025.

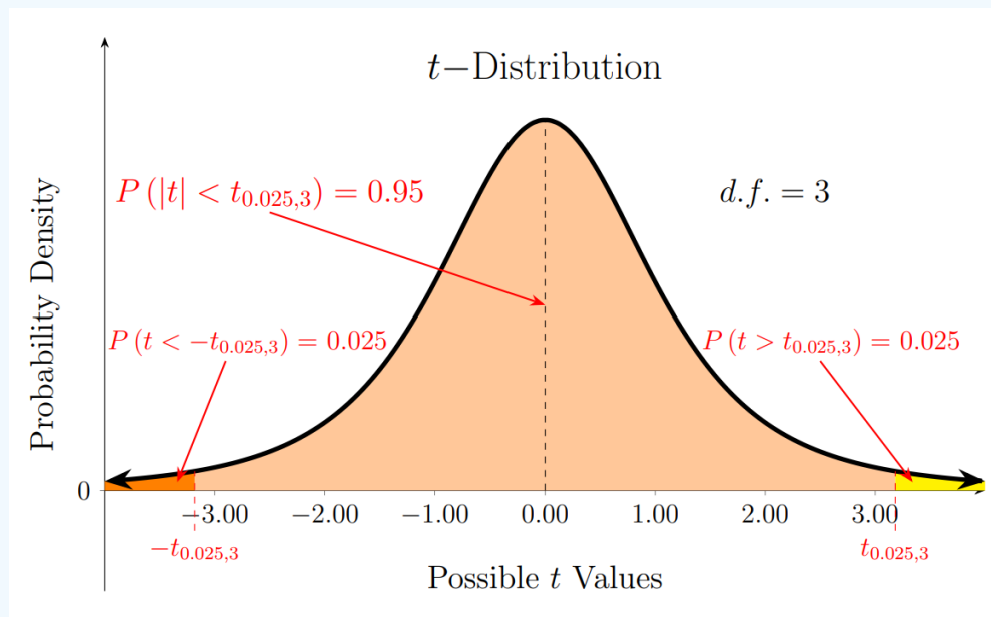


Figure 6.4.2 t -Distribution with 3 degrees of freedom

Since the critical values are equal in magnitude but opposite in sign, we can compute the positive critical value by multiplying the negative critical value by -1 .

$$t_{0.025,3} = -1 \cdot \text{T.INV}(0.025, 3) \approx 3.1825$$

2. $n = 8$

Answer

Keeping the same level of confidence but with a larger sample, our tasks throughout this text exercise are similar. The t -distribution now has 7 degrees of freedom, but the areas and the process remain the same.

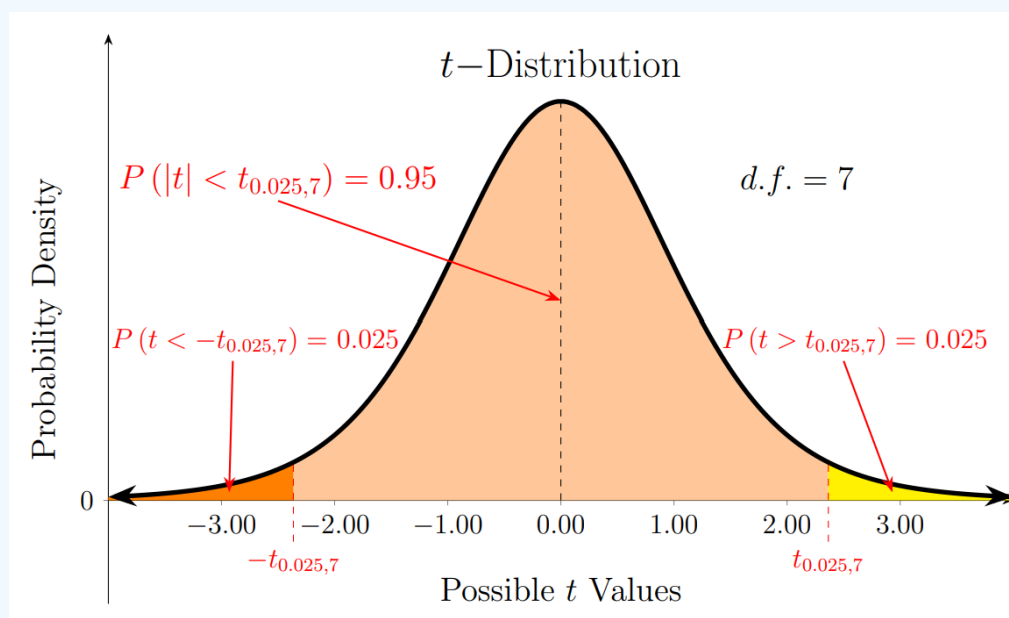


Figure 6.4.3 t -Distribution with 7 degrees of freedom

$$t_{0.025,7} = -1 \cdot \text{T.INV}(0.025, 7) \approx 2.3646$$

3. $n = 16$

Answer

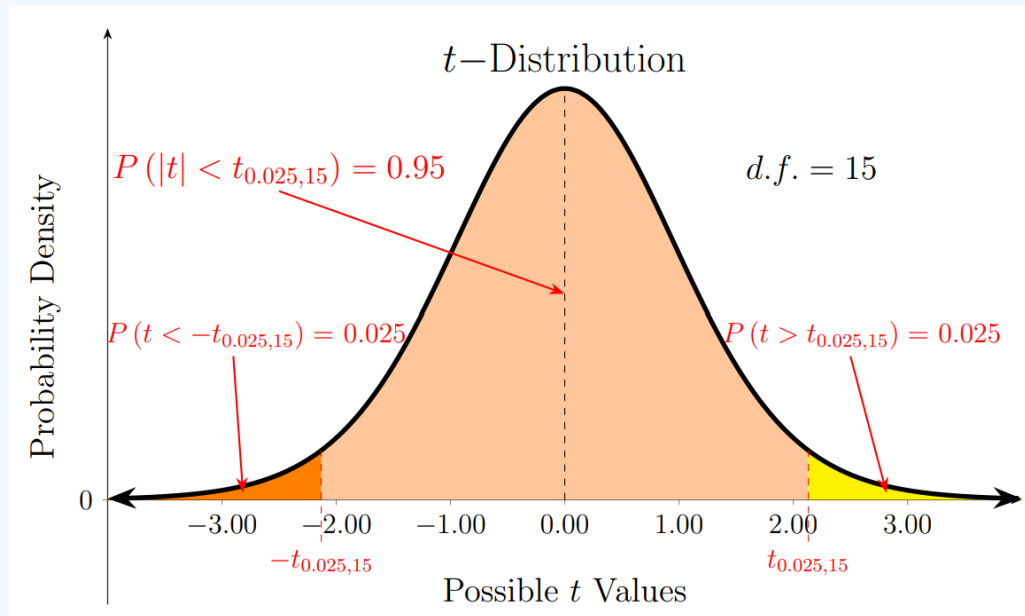


Figure 6.4.4 t -Distribution with 15 degrees of freedom

$$t_{0.025,15} = -1 \cdot \text{T.INV}(0.025, 15) \approx 2.1315$$

4. $n = 32$

Answer

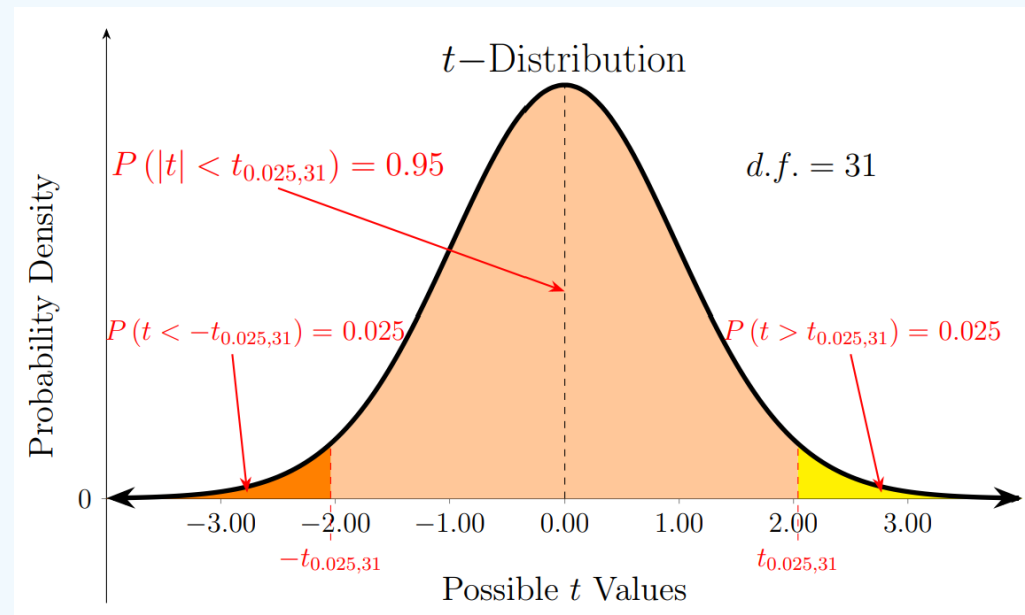


Figure 6.4.5 t -Distribution with 31 degrees of freedom

$$t_{0.025,31} = -1 \cdot \text{T.INV}(0.025, 31) \approx 2.0395$$

5. Describe what is happened to the t -distributions as the degrees of freedom increased. What happened to the positive critical values? What will happen if we use larger and larger sample sizes?

Answer

Since each figure is plotted with the same horizontal axis, we can tell that the area under the curve in the tails decreases fairly noticeably with each subsequent figure. Since the total area is always 1, there is more area in the central portion of the distribution. We can see this faintly as the curve rounds out near its peak and thickens ever so slightly by the labels of the areas in the tails. The critical values are decreasing in magnitude with each increase in the degrees of freedom. This makes sense because the tails are getting thinner. The rate at which the tails are thinning and the critical values are decreasing in magnitude is slowing down. Recall that when we first introduced the t -distribution, we said that the distribution gets closer and closer to the standard normal distribution as the degrees of freedom increase. As such, we can expect the critical values of the t -distributions to approach the critical value of the standard normal distribution $z_{0.025} \approx 1.96$ as the sample size increases. With 61 degrees of freedom, we finally are less than 2. Only after 473 degrees of freedom will answers rounded to two decimal places match for this level of confidence.

Note: Degrees of Freedom

Recall that the shape of the t -distribution is determined by a quantity called degrees of freedom ($d. f.$). Notice that the degrees of freedom in the figure above are related to the sample size, n ; in particular, $d. f. = n - 1$. We shall now discuss the meaning of the terminology. In reading "degrees of freedom", we might naturally think about the extent to which someone is capable of determining and carrying out an action. That is a fine initial intuition; along those lines, we can think of the degrees of freedom as a measure regarding the extent to which the data can vary independently, given any constraints on the data.

Suppose the sample mean of 20 observations was 10. With the current information that we have, we do not know anything about the actual values; all we know is that the sum of the 20 observations is 200. There are infinitely many possible values for our 20 observations that result in such a mean. If it was then revealed to us that the first observation value was 11, we would know the sum of the last 19 observation values is 189. There are again infinitely many possible values for these 19 observations. If it was then revealed that the second observation value was 15, we would know the sum of the last 18 observation values is 171, which can again happen in infinitely many ways. At what point, after revealing many observation values, will we know what the remaining observation values have to be knowing the sample mean is 10? Once 19 of the observation values are revealed, we will be able to deduce the last observation without it being revealed. We thus say that the first 19 observations are free to vary independently, but the 20th observation value depends entirely on the previous observation values since the sample mean is known. Hence, there are 19 ($n - 1$) degrees of freedom.

When collecting data via random sampling, each observation value is independent of the others; the information gleaned from each observation provides information independent of the other observations. However, once we know, say, the sample mean, a constraint is placed on the data. The observation values are now connected through the expression $x_1 + x_2 + \dots + x_n = \bar{x} = 20$. Each bit of information is no longer perfectly independent. Degrees of freedom is a measure for how much independent information remains given the constraint(s) in place.

So, why does the t -distribution have $n - 1$ degrees of freedom in this situation? The sample mean only requires the observation values; there are no constraints connecting the data within the definition of the sample mean. This is not the case when we compute sample standard deviation. Here, we consider all the square deviations from the mean as the fundamental pieces of information, but implicit within this description is an expression that connects all data values: the sample mean. Each square deviation is not perfectly independent from every other square deviation. Since we needed the sample mean to compute the sample standard deviation, we have a constraint within the analysis. As we have seen above, this constraint reduces the amount of independent information by 1.

As a general rule of thumb, we expect the degrees of freedom to be the sample size minus the number of statistics used within computations of the statistics involved in the analysis. Different statistical analyses have different degrees of freedom; so, it is important to understand where the measurement comes from and to pay close attention to the literature explaining any particular statistical analysis.

With the critical values in hand, we again notice that the positive critical value is equal to the length of the scaled margin of error and determine our margin of error from there.

$$\frac{ME}{\frac{s}{\sqrt{n}}} = t_{\frac{\alpha}{2}, n-1}$$

$$ME = t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}}$$

Constructing Confidence Intervals for Means (σ unknown)

We now have all the pieces to construct a confidence interval for the population mean.

$$(\bar{x} - ME, \bar{x} + ME) = \left(\bar{x} - t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}} \right)$$

We often write these confidence intervals as $\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}}$.

? Text Exercise 6.4.2

We have discussed the heights of adult females on several occasions; they are normally distributed with a mean of 64 inches with a standard deviation of 2.5 inches. We have yet to discuss the heights of adult males. It seems reasonable to think that if the heights of adult females are normally distributed, the heights of adult males will also be normally distributed (this is indeed the case). We want to build a 90% confidence interval to catch the average height of adult males. To do so, we randomly sampled 10 adult males. Their heights in inches from shortest to tallest are reported below. Construct and interpret the confidence interval.

64, 66, 67.5, 68, 69, 69.75, 71.25, 72, 73, 73.25

Answer

In order to construct a confidence interval for means, we need the sampling distribution of sample means to be at least approximately normal and the sample to have been randomly selected. Given that the heights of adult males are normally distributed, we have that the sampling distribution of sample means is normally distributed despite the fact that we only have a sample size of 10. The sample was randomly chosen; we can, therefore, proceed.

We need the sample mean, standard deviation, sample size, and positive critical value to construct the confidence interval. We report, without explanation, that $\bar{x} = 69.375$ inches, $s \approx 3.0647$ inches, and $n = 10$. To calculate the positive critical value, we must find α . Since $CL = 0.90$, $\alpha = 0.1$. Since we do not know the population standard deviation, our critical value comes from the t -distribution with 9 degrees of freedom. We sketch the distribution to help us compute the positive critical value.

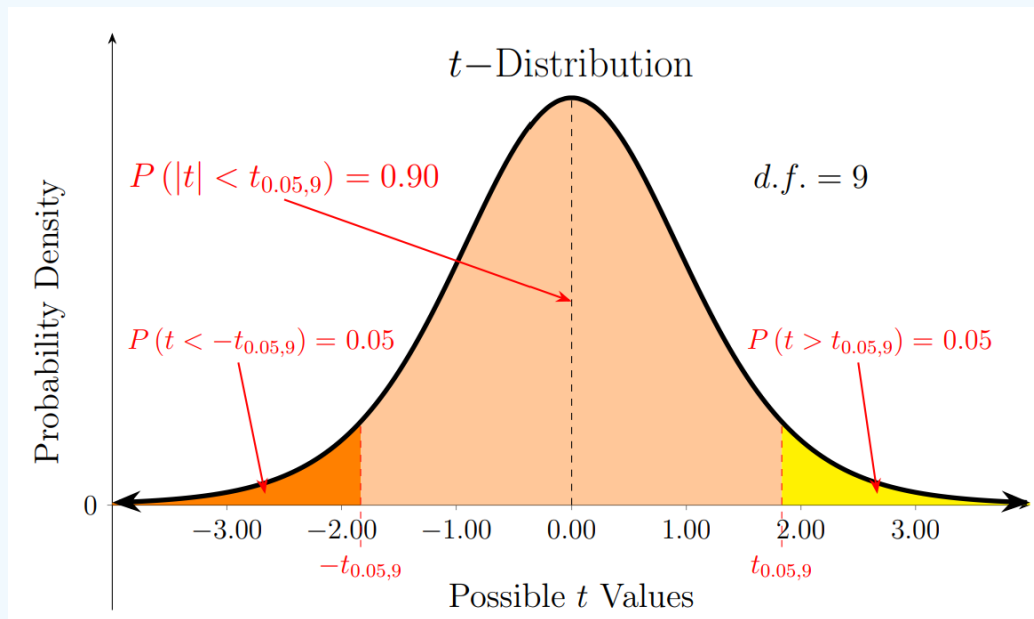


Figure 6.4.6 t -Distribution with 9 degrees of freedom

$$t_{0.05,9} = -1 \cdot T.INV(0.05, 9) \approx 1.8331$$

$$\left(\bar{x} - t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}} \right) \approx \left(69.375 - 1.8831 \cdot \frac{3.0647}{\sqrt{10}}, 69.375 + 1.8831 \cdot \frac{3.0647}{\sqrt{10}} \right) \approx (67.5985, 71.1516)$$

At a confidence level of 90%, the average height of all adult males μ is between 67.5985 inches and 71.1516 inches.

For many, the confidence interval produced in the previous text exercise is rather disappointing; there is a wide range of values that the population mean could be. We might desire to determine how large of a sample would be sufficient to expect that the margin of error is less than some specified value. In our case, we might be interested in determining how large of a sample would be sufficient to expect that at most one whole number falls in the constructed confidence interval while keeping the confidence level at 90%.

A natural place to begin would be with the margin of error formula and trying to solve for n , just as we did in the last section.

$$ME = t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}}$$

Difficulties, however, arise in several places. There are two values in the equation that depend on the value of n , the critical value and the square root of n . We also do not know what our sample standard deviation will be without actually collecting the sample. Unfortunately, these issues cannot be remedied perfectly, but there are paths forward that can arrive at reasonable estimates of a sufficient sample size. These estimates on n are just that, estimates; they will not be guaranteed to work without fail, but in general, we can rely on them to continue forward. The reader is encouraged to take more advanced statistics courses for a thorough explanation. We shall only provide some basic intuition about the considerations. We first solve for the sample size n .

$$n = \left(\frac{t_{\frac{\alpha}{2}, n-1} \cdot s}{ME} \right)^2$$

We have not solved for n explicitly because the critical value from the t -distribution depends on n . Our path forward in estimating n is to replace the values in the numerator with values that we believe are larger than the values that will end up being used when we actually construct the confidence interval. When we replace one value in the numerator with one that is larger in the expression, the product will be larger. If we use this larger product as our estimated n , we have probably chosen a larger than necessary sample size for our desired precision. Doing this will give us a conservative estimate of the sample size. We could always be wrong about whether the values were larger or not and then possibly not reach the desired precision in the confidence interval. Let us go through this value by value.

How do we pick a large enough value to overestimate the critical value? Recall that as the degrees of freedom increase, the t -distributions get closer and closer to the standard normal distribution and that the t -distribution has fatter tails than the standard normal distribution. These two facts imply that for a given confidence level, the critical values of the t -distributions are farther away from 0 than the critical values of the standard normal distribution, but as n increases the critical values of the t -distributions approach the critical values of the standard normal distribution. This means that the critical values in t -distributions get smaller in magnitude as n increases. This gives a conservative overestimate of the critical value by using the positive critical value at the same level of confidence from a t -distribution with fewer degrees of freedom than you expect to have from your future sample. If the underlying distribution is not normal, we can expect to need at least a sample size of 31, so $t_{\frac{\alpha}{2}, 30}$ would be an overestimate because we know we need at least 31 observations in our sample. If the underlying distribution is normal, we could have less in our sample. In this case, we could use the t -distribution with only 1 degree of freedom. It will be at least as large as every possible critical value. This is the most conservative estimate we can make.

Just as with constructing confidence intervals, there is a balancing act at play, not between confidence and precision, but between confidence and work. The more conservative our estimate, the larger the estimated sample which requires more work on the part of the researcher. There are time and financial constraints involved. Sometimes, researchers are satisfied with less conservative estimates which is fine since the most conservative estimates are overestimates by a large margin. Indeed, some textbooks recommend using the critical value from the standard normal distribution, which is guaranteed to be an underestimate of the critical value used when constructing the confidence interval.

The decision to choose a large enough value to overestimate the sample standard deviation is a more difficult question. In essence, the sample standard deviation is an estimate for the population standard deviation. We know that the sample standard deviation tends to be an underestimate of the population standard deviation but that it can be larger. We have computed a sample standard deviation from an initial study, but we do not know if it is larger than the population standard deviation or smaller. We expect that for larger samples our

computed standard deviation is closer to the population standard deviation but that is not necessarily the case. As we might see (if you study all of the bonus material, you will), it is sometimes possible to construct a confidence interval for the population standard deviation. With such an estimate, we could estimate the upper bound of where we are confident the population standard deviation falls, but even that could fail us. As such, some researchers construct confidence intervals from the pilot data. Others just use the sample standard deviation found in the sample data. And, yet others add a certain amount to the computed sample standard deviation. Here is where our guarantee must fail, but again that does not mean the estimate is not useful for continuing. We will be satisfied using the sample standard deviation from the pilot study in this text.

Within this text, we adopt the practice of using the pilot study sample standard deviation, comparing both $t_{\frac{\alpha}{2},1}$ and $t_{\frac{\alpha}{2},30}$ when the underlying distribution is normal, and otherwise just using $t_{\frac{\alpha}{2},30}$ in our estimates. In this last case, if our estimated value is less than 31, we select 31 instead. Let us now estimate the sample size such that we expect to have only one whole number in confidence interval.

We have a sample of size 10 and computed its sample standard deviation. We do not expect this standard deviation to be equal to the population standard deviation or the future sample to have the same standard deviation as this sample, but we will use $s = 3.0647$ to approximate what the sample standard deviation of a future sample might be. Since the underlying distribution is normal, we will compare $t_{0.05,1} \approx 6.3138$ and $t_{0.05,30} \approx 1.6973$ as overestimates of the critical value.

We also need to determine the margin of error specified in the problem. If there is to be at most one whole number in the interval, the interval length must be less than 1 for the distance between consecutive whole numbers is 1. Thus if the length of our interval is larger than 1, it is possible to have two whole numbers in our confidence interval. If the length is less than or equal to 1, then it is possible that we do not have any whole numbers in our interval, but that is permitted in the phrasing of the question. Thus the maximal length of the interval so that we have at most one whole number in the interval is 1. Since the margin of error is half of the interval length, our desired margin of error is confirmed to be 0.5 inches. We thus consider the following two cases.

$$n \leq \left(\frac{t_{0.05,1} \cdot s}{\text{ME}} \right)^2 \approx \left(\frac{6.3138 \cdot 3.0647}{0.5} \right)^2 \approx 1497.652$$

$$n \leq \left(\frac{t_{0.05,30} \cdot s}{\text{ME}} \right)^2 \approx \left(\frac{1.6973 \cdot 3.0647}{0.5} \right)^2 \approx 108.2264$$

In the first case, the estimated sample size is 1498; in the second case, the estimated sample size is 109. In both cases, the estimated sample size is larger than 30. As such, we use the estimate from case 2. We expect that a random sample of 109 adult males will produce a confidence interval that contains at most 1 whole number.

? Text Exercise 6.4.3

A pilot study of 9 randomly selected college students revealed that the average time spent scrolling on social media per day was 2.2 hours with a standard deviation of 2.1 hours. The researchers plan to conduct a larger study to construct a confidence interval at a confidence level of 99 with a margin of error of 1.25 hours. Estimate the number of college students that should be randomly sampled in the larger study to produce the desired results.

Answer

The goal is to build a confidence interval at the 99% confidence level with a margin of error that is no more than 1.25 hours. The sample standard deviation from the pilot study is 2.1 hours. We do not know anything specific about the population. We, therefore, use $t_{0.005,30} \approx 2.75$ to replace our critical value in the estimation.

$$n \leq \left(\frac{t_{0.005,30} \cdot s}{\text{ME}} \right)^2 \approx \left(\frac{2.75 \cdot 2.1}{1.25} \right)^2 \approx 21.3443$$

Using the critical value for a sample size of 31 produced an estimated sample size of 22. We cannot use this result for two reasons, in order to build our confidence interval in this situation, we need a sample size of at least 31. The critical value at the 99% confidence level with a sample of 22 will be larger than the critical value used in our estimate; we, therefore, must proceed with caution. We are not guaranteed that the value would not be sufficient given the possibility that the sample standard deviation in the large study might be smaller than in the pilot study. Luckily for us, the decision was forced with the original consideration. If the underlying distribution is not normal and we do not know any more information about its

distribution, we expect the sampling distribution of sample means to be approximately normal when $n > 30$ which is an important requisite for constructing confidence intervals.

6.4: Confidence Interval for Means (Sigma Unknown) is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by The Math Department at Fort Hays State University.

- **10.7: Confidence Interval for Mean** by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.