

Chapters 8 and 9: Sections 8.1, and 9.2--Confidence Intervals

Section 8.1 – The Sampling Distribution of the Mean

THEOREM : $\mu_{\bar{X}} = E(\bar{X}) = \mu$ and $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

Note: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ is also called the **standard error of the mean**.

Case 1: Shape of the Distribution of \bar{X} if X is Normal i.e; $X \sim N(\mu, \sigma)$.

If a random variable X is normally distributed, then the sampling distribution of the sample mean, \bar{X} , is also Normally distributed, i.e;

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right).$$

Case 2: (The Central Limit Theorem-CLT)

Regardless of the shape of the underlying population, sampling distribution of the sample mean, \bar{X} , becomes approximately normal as the sample size, **n**, increases regardless of the distribution of X. The approximation becomes better and better with increasing sample size. **Note:** A sample size of **n=30** or more is generally considered large.

Example 1: Weights of men are normally distributed with a mean of 170 lbs and a standard deviation of 20 lbs. A sample of 16 men is taken and their weights were observed (a) What is the sampling distribution of \bar{X} , the sample average? (b) Find the probability that a randomly selected man will weigh less than 166 lbs? (c) What is the probability that the average weight for the sample selected will be less than 166 lbs? (d) What is the standard error of the mean? Ans. (a) $N(170, 5)$ (b) 0.4207 (c) 0.2119 (d) 5

a. Based on **Case 1**: Since $X \sim N(\mu = 170, \sigma = 20)$ **then** $\bar{X} \sim N\left(\mu = 170, \sigma = \frac{20}{\sqrt{16}} = 5\right)$

b. $P(X < 166) = \text{Normalcdf}(-E99, 166, 170, 20) = 0.4207$

c. $P(\bar{X} < 166) = \text{Normalcdf}(-E99, 166, 170, 5) = 0.2119$

d. Standard error of the mean or standard deviation of \bar{X} ; $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{16}} = \frac{20}{4} = 5$

Example 2: One hundred small bolts are packed in a box. These bolts are randomly selected from a population with a mean weight of 1 ounce and a standard deviation of 0.01 ounces. Find the probability that the average weight of the bolts in the box is less than 1.0015 ounces. Ans. (0.9332)

$$n=100; \text{ Based on Case 2: } \bar{X} \sim N\left(\mu=1, \sigma = \frac{0.01}{\sqrt{100}} = 0.001\right)$$

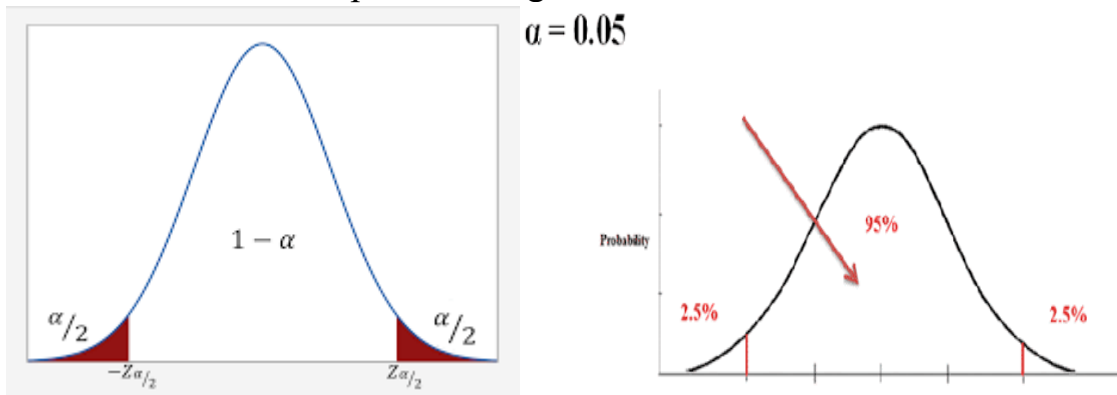
$$P(\bar{X} < 1.0015) = \text{Normalcdf}(-E99, 1.0015, 1, 0.001) = 0.9332$$

Homework-Section 8.1 Online - MyStatLab

Section 9.2 Estimating Population Mean

From section 8.1 cases 1 and 2 show that the sampling distribution of \bar{X} is Normal ; i.e. $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$. We can use it to build a $(1 - \alpha)\%$ confidence interval for μ ; i.e. if $\alpha = 0.05$ then $1 - \alpha = 1 - 0.05 = 0.95$ or 95% confidence interval. It means, we will use the sampling distribution of \bar{X} ,

$\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$, to find the 95% middle range of the \bar{X} 's. Then we are 95% that the true population mean μ should be in this range. We will not use the formulas below to compute the range. We will use the calculator.



Confidence Interval: A confidence interval for estimating μ is

$[\bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}]$ or $[\bar{X} \pm Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}]$ where $1 - \alpha$ is the called the confidence level and $Z_{\frac{\alpha}{2}}$ is the z-value providing an area of $\frac{\alpha}{2}$ in the upper tail of the standard normal probability distribution.

There is a problem using this method. Please NOTE: It's not realistic to claim we know σ but we don't know μ . Therefore we have to use another sampling distribution to build a $(1 - \alpha)\%$ confidence interval for μ .

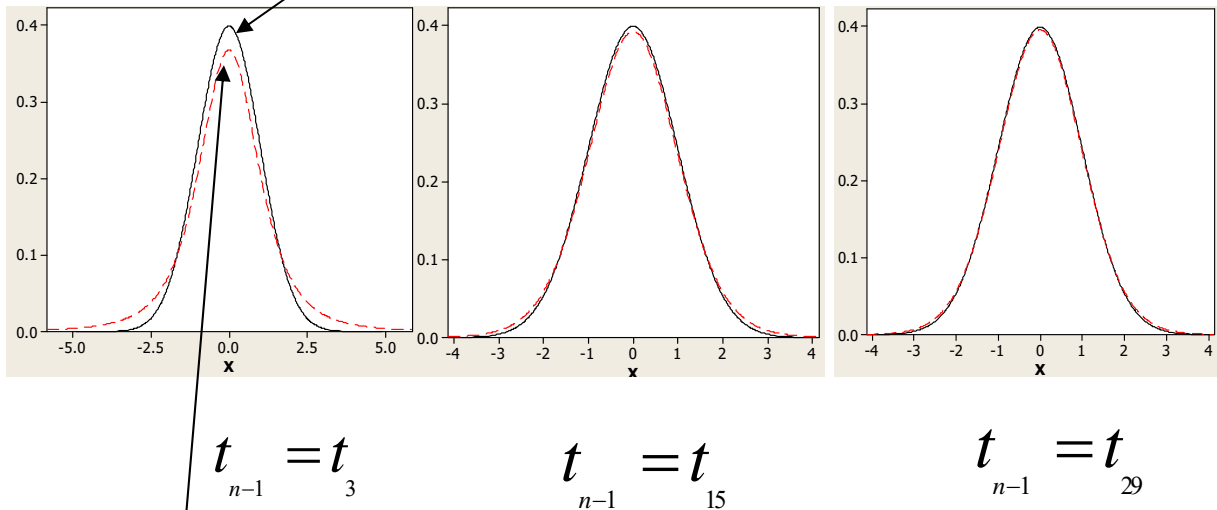
Since we don't know the population standard deviation σ , we will use its estimate S , the sample standard deviation.

When we use S in place of σ the sampling distribution of \bar{X} changes to a new sampling distribution called the student's t-distribution or just the t-distribution.

Note: **Case 1** in section 8.1 becomes **Case 3** in this section and **Case 2** in section 8.1 becomes **Case 4** in this section. Under cases 3 and 4 the sampling distribution of \bar{X} is a t-distribution with $n-1$ degrees of freedom; i.e. $\bar{X} \sim t_{n-1}$.

The t distribution is actually a family of probability distributions. Each specific distribution depends upon a parameter known as the degrees of freedom (df) where $df = n - 1$.

As the number of degrees of freedom increases, the t-distribution and the standard normal distribution, $Z \sim N(0,1)$, look the same (see the graphs below).



The first (red curve) has $df=3$, the second (red curve) $df=15$, and the third (red curve) $df=29$. All three graphs are compared to the standard Normal distribution, $Z \sim N(0,1)$.

Note: As the degrees of freedom increases the t-distribution is approaching the standard Normal distribution, $Z \sim N(0,1)$.

Confidence Interval for μ using the t-distribution

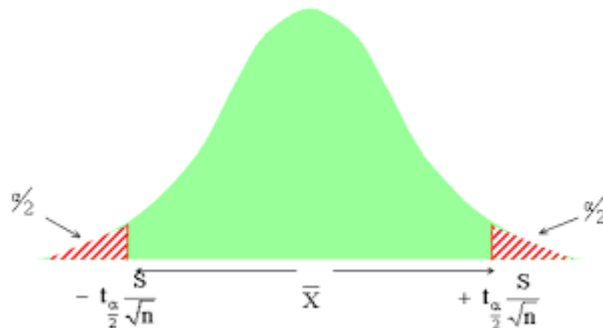
Finally we are ready to use the calculator to build the $(1 - \alpha)\%$ confidence interval using the sampling distribution that \bar{X} has a t-distribution with $n-1$ degrees; i.e. $\bar{X} \sim t_{n-1}$.

Case 3: If $n < 30$ and the data comes from a Normal distribution, then the sampling distribution of \bar{X} is t_{n-1} ; i.e. $\bar{X} \sim t_{n-1}$.

Case 4: If $n \geq 30$ (the sample size is large, $n \geq 30$). Then the sampling distribution of \bar{X} is t_{n-1} ; i.e. $\bar{X} \sim t_{n-1}$.

A C.I. for μ is $\bar{X} \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$ where $1 - \alpha$ is the confidence coefficient and $t_{\frac{\alpha}{2}}$ is the t-value providing an area of $\frac{\alpha}{2}$ in the upper tail of a t-distribution with $n - 1$ degrees of freedom.

Note: $t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$ is the sampling error or (margin of error).

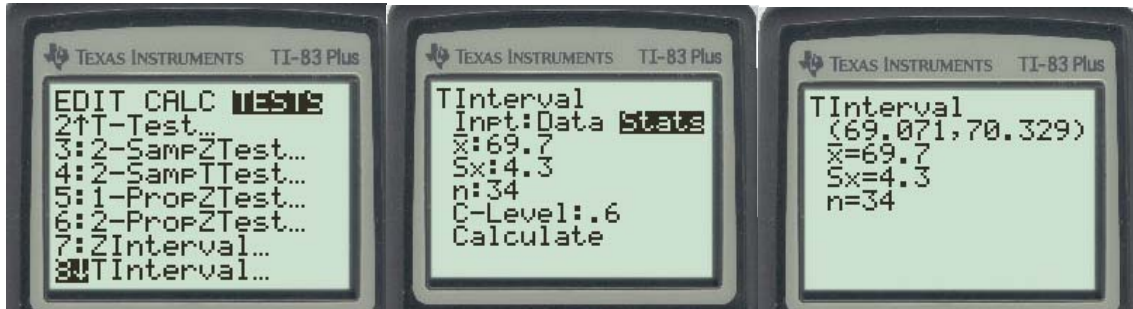


Example 1: The U.S. Department of Health, Education, and Welfare collected data for 34 males between the ages of 18 and 24. The sample of 34 males yield a mean, $\bar{X} = 69.7$ inches and a standard deviation, $S=4.3$. **Note:** This is case 4 since ; i.e. $\bar{X} \sim t_{33}$. $n \geq 30$.

a) Compute the 60% confidence interval for μ and interpret it.

$$\bar{X} \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \Rightarrow (69.071, 70.329) \quad (\text{Using TI-83/84})$$

Using TI-83/84: stat, TESTS, TINTERVAL (is #8), then choose “stats”. Enter \bar{X} : 69.7, s_x : 4.3, n: 34, C-Level: 60 or .60, put your cursor on “calculate” and hit enter. You should get the range (69.071, 70.329).



To find the sampling error or margin of error take the upper limit of the confidence interval and subtract the sample mean \bar{X} ; i.e. sampling error(margin of error) = 70.329 - 69.7 = 0.629.

Note: the sampling error(margin of error); $t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = 0.629$. (Using TI-83/84)

Interpretation: We are 60% confident that the mean height of males between the ages of 18 and 24 is somewhere between 69.071 and 70.329 inches.

b) Compute the 99% confidence interval for μ and interpret it.

$$\bar{X} \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \Rightarrow (67.684, 71.716) \quad (\text{Using TI-83/84})$$

Using TI-83/84: stat, TESTS, TINTERVAL (is #8), then choose “stats”. Enter \bar{X} : 69.7, s_x : 4.3, n: 34, C-Level: 99 or .99, put your cursor on “calculate” and hit enter. You should get the range (67.684, 71.716).

To find the sampling error or margin of error take the upper limit of the confidence interval and subtract the sample mean \bar{X} ; i.e. sampling error(margin of error) = 71.716 - 69.7 = 2.016.

Note: the sampling error ; $t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = 2.016$. (Using TI-83/84)

Interpretation: We are 99% confident that the mean height of males between the ages of 18 and 24 is somewhere between 67.684 and 71.716 inches.

c) In **Example 1** compare parts **a** and **b** and draw a conclusion

The 99% confidence interval in b is wider than the 60% confidence interval in a. We can conclude that as the confidence level increases the interval gets wider.

Example 2: A manufacturer wants to estimate the speed of its new printer. A random sample of 24 printers yielded a mean of 111.875 characters per second with a sd of 8 characters per second. Assume that speed of these printers is normally distributed. **Note: This is case 3 since $n < 30$ and the sample data came from a normal distribution. ; i.e. $\bar{X} \sim t_{23}$.**

Using TI-83/84: stat, TESTS, TINTERVAL (is #8), then choose “stats”. Enter \bar{X} : 111.875, S_x : 8, n: 24, C-Level: 95 or .95, put your cursor on “calculate” and hit enter. You should get the range (108.50, 115.25). Repeat for part b.

- Construct a 95% confidence interval for the population mean.
- Construct a 99% confidence interval for the population mean.
- Is your 99% confidence interval wider or narrower than 95% confidence interval? Explain your answer carefully.

Soln. (a) A 95% C.I. is $\bar{X} \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = (108.50, 115.25)$ Using TI-83/84

(b) A 99% C.I. is $\bar{X} \pm t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = (107.29, 116.46)$ Using TI-83/84

(c) Wider than 95% since the confidence level was increased to 99%.

Example 3: An economist wants to estimate the average annual income of widows living in a certain city. A random sample of 60 widows gives a sample mean of \$11,700 with a sample standard deviation of \$2,400. Calculate a 90% confidence interval for the average income of the population. **Note: This is case 4 since $n \geq 30$; i.e. $\bar{X} \sim t_{59}$.** Follow the calculator sequence of commands above.

Soln. A 90% CI is $\bar{X} \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = (11182, 12218)$ Using TI-83/84

Example 4: A sample of 10 cracker boxes yielded the following net weights (in ounces): 16.1, 16.01, 15.82, 15.93, 16.05, 16.05, 15.86, 15.90, 16.05, 16.13. Construct a 95% confidence interval for the population mean. You may assume that "net weight" is a normal random variable. **Note: This is case 3 since $n < 30$ and the sample data came from a normal distribution; i.e. $\bar{X} \sim t_9$.**

In this example we must enter the data in a list first. **Stat, 1:Edit** and then enter the data in L1. Now follow the sequence of commands below.

Using TI-83/84: stat, TESTS, TINTERVAL (is #8), then choose "data". List: L1 (enter the data in L1), Freq: 1, C-Level: 95 or .95, put your cursor on "calculate" and hit enter. You should get the range (15.915, 16.065). Note: Freq is always 1. Repeat for part b.



A 95% CI is $\bar{X} \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = (15.915, 16.065)$ Using TI-83/84

Determining the Required Sample Size

If we try to achieve a certain sampling error; i.e $E = Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$ we can solve this

formula for $n = \left(\frac{Z_{\frac{\alpha}{2}} S}{E} \right)^2$ and call it the sample size formula. This formula

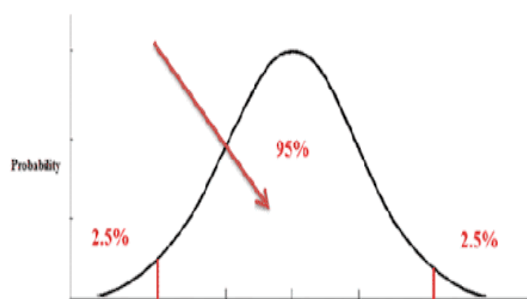
finds the right sample size n in order to achieve a sampling error of E . To use this formula we need S , E , and $Z_{\frac{\alpha}{2}}$. S and E are always given in the word problem. $Z_{\frac{\alpha}{2}}$ needs to be determined using the confidence level given in the word problem. The TI83/84 command is `Invnormal(percent, μ , σ)`. We practice the `Invnormal` in Chapter 7.

Example 5: We wish to be 95% sure that a random sample of IQ scores yields a sample mean that is within 2 points of its true mean. How large should the sample be if the standard deviation is 15.

Since the confidence level is 95% then

$$1 - \alpha = 0.95 \Rightarrow \alpha = 1 - 0.95 \Rightarrow \alpha = 0.05 \text{ and } \frac{\alpha}{2} = \frac{0.05}{2} = 0.025 .$$

$$\alpha = 0.05$$



We need to find the value of $Z_{\frac{\alpha}{2}=0.025}$ that has an area to the right of 0.025 and an area to the left of $1 - 0.025 = 0.975$. The calculator can only find Z scores using the area to the left. Hence;

$$\text{Invnormal}\left(1 - \frac{\alpha}{2}\right) = \text{Invnormal}(1 - 0.025) = \text{Invnormal}(0.975) = 1.959963986.$$

You can also find the Z-value by doing $\text{Invnormal}\left(\frac{\alpha}{2}\right) = \text{Invnormal}(0.025)$

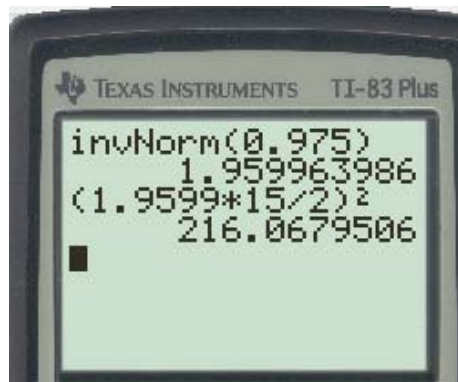
$= -1.959963986$. It's the same Z-value but negative. Just ignore the negative

Using TI-83/84: 2nd and vars, invnorm(is #3). Invnormal(0.975)=1.959963986 or Invnormal(0.025)= -1.959963986(Ignore the negative).

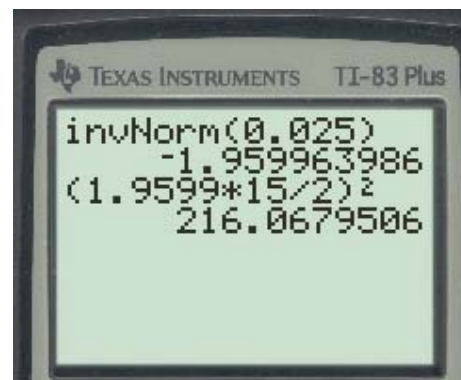
To apply the sample size formula, take 4 decimals without rounding; i.e. 1.9599.

$$\text{Soln. } n = \left(\frac{Z_{\frac{\alpha}{2}} S}{E} \right)^2 = \left(\frac{(1.9599)(15)}{2} \right)^2 = 216.06795 = 217 \quad \text{Ans. 217 scores.}$$

Use your calculator in one step to evaluate the formula. This way you will be avoiding rounding problems.



or



Please note: For the final answer for **n** we always round up regardless of the decimal. It does not follow the rounding rules.

Example 6: Find the sample size necessary to estimate the mean life of a brand of tires if we want the estimate to be within 500 miles of the true mean with 99% confidence if $S=1,200$ miles.

Since the confidence level is 99% then

$$1 - \alpha = 0.99 \Rightarrow \alpha = 1 - 0.99 \Rightarrow \alpha = 0.01 \text{ and } \frac{\alpha}{2} = \frac{0.01}{2} = 0.005$$

We need to find the value of $Z_{\frac{\alpha}{2}=0.005}$ that has an area to the right of 0.005 and an area to the left of $1 - 0.005 = 0.995$. The calculator can only find Z scores using the area to the left. Hence;

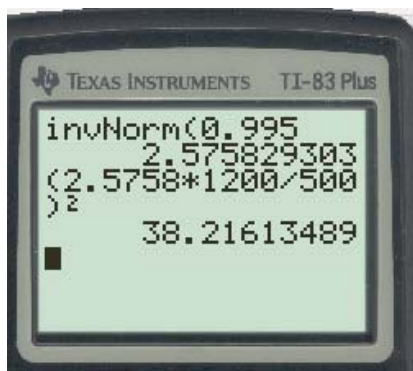
$$\text{Invnormal}(1 - \frac{\alpha}{2}) = \text{Invnormal}(1 - 0.005) = \text{Invnormal}(0.995) = 2.575829303.$$

You can also find the Z-value by doing $\text{Invnormal}(\frac{\alpha}{2}) = \text{Invnormal}(0.005) = -2.575829303$. It's the same Z-value but negative (ignore the negative).

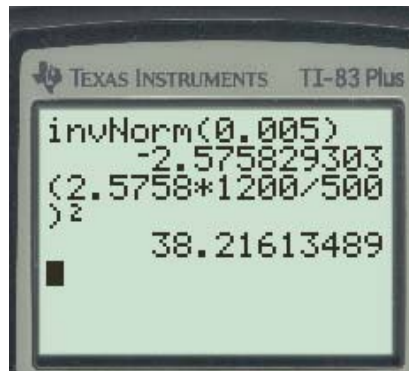
Using TI-83/84: 2nd and vars, invnorm(is #3). $\text{Invnormal}(0.995) = 2.575829303$ or $\text{Invnormal}(0.005) = -2.575829303$ (Ignore the negative).

To apply the sample size formula, take 4 decimals without rounding; i.e. 2.5758.

$$\text{Soln. } n = \left(\frac{Z_{\frac{\alpha}{2}} S}{E} \right)^2 = \left(\frac{(2.5758)(1,200)}{500} \right)^2 = 38.21613489 = 39 \quad \text{Ans. 39 tires.}$$



OR



Please note: For the final answer for **n** we always round up regardless of the decimal. It does not follow the rounding rules.

Homework-Section 9.2 Online - MyStatLab .