Section T
Confidence Intervals for the Population Mean (Continued)

Now what if, as is usually the case in real life, the population standard deviation is unknown. We need to estimate the population standard deviation, σ, with the sample standard deviation, s. Therefore, we can no longer use a z-score to construct a confidence interval for the population mean, we need to use what is called the studentized version of $\bar{x}$, which is called a *t*-distribution.

t – distribution

Suppose a variable *x* of a population is normally distributed with mean, μ. Then, for samples of size *n*, the studentized version of $\bar{x}$ is

$$t = \frac{\bar{x}-\mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{\bar{x}-\mu}{s_{\bar{x}}},$$ where $s_{\bar{x}} = \frac{s}{\sqrt{n}}$, has the *t*-distribution with *n* – 1 degrees of freedom.

Basic Properties of *t*-curves

1) Area under curve is equal to 1
2) Extends indefinitely in both directions
3) Symmetric about 0
4) As the number of degrees of freedom becomes larger *t*-curves look increasingly like the standard normal curve.

Unlike z-values which do not change for a specific confidence level, a t-value is based on the degrees of freedom and therefore has different values for a specific confidence level. To find a t-value for a given confidence interval we use **Table C with df = n – 1**

Examples:

1) Given n= 15 and CL = 95%, find $t_{\alpha/2}$= **2.145**(df = 14) 2) Given n=74 and CL = 99%, find $t_{\alpha/2}$= **2.645** (df = 73)

3) Given n= 51 and CL = 98%, find $t_{\alpha/2}$= **2.403**(df = 50) 4) Given n=38 and CL = 90%, find $t_{\alpha/2}$= **1.687** (df = 37)

Case II: Confidence Intervals for One Population Mean, μ, when σ is Unknown

Assumptions:
1) Simple random sample
2) Normal population or large sample (n ≥ 30)
3) The population standard deviation, σ, is unknown

$$\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$ To find $t_{\alpha/2}$ use Table C with *df* = *n* – 1.

Confidence Interval $\left(\bar{x} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right),\ \ \bar{x} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)\right)$

Note: The confidence interval is exact for normal populations and is approximately correct for large samples from non-normal populations

Examples:

1) A manufacturer of cookies and crackers does a small survey of the age at sale of one of its brands. Assume the ages are normally distributed. A random sample of 23 retail markets in a particular region is chosen. In each store, the number of days since manufacture of the front most box of crackers is determined by a date code on the box. The results yielded a mean of 22.64 days and standard deviation of 5.76 days. Construct a 99% confidence interval for the population mean age. Interpret your result.

CL = 99%

$\bar{x}$ **= 22.64**

n = 23

s = 5.76

$t_{\alpha/2}$ **= 2.819**
df = 22

$$\mathbf{22.64 \pm 2.819\left(\frac{5.76}{\sqrt{23}}\right)}$$

(19.25, 26.03)

I am 99% confident the population mean age at sale is between 19.25 and 26.03 days.

b) A claim is made that the population mean age is 30.5 days. Using the results from part (a) of this problem, would this claim be valid or invalid?

Invalid

2) A consumer group wants to estimate the average price of a certain model of microwave in the NY metropolitan area. Assume prices are normally distributed. Prices are determined by comparison shoppers at 14 randomly selected stores in the area. The dollar price (including taxes) are
341 347 319 331 326 298 335 351 316 307 335 320 329 346 $\bar{x}$ =328.64 s = 15.49
Construct a 98% confidence interval for the population mean. Interpret your result.

CL = 98%

$\bar{x}$ **= 328.64**

n = 14

s = 15.49

$t_{\alpha/2}$ **= 2.650**
df = 13

$$\mathbf{328.64 \pm 2.650\left(\frac{15.49}{\sqrt{14}}\right)}$$

(317.67, 339.61)

I am 98% confident the population mean price is between $317.67 and $339.61.

b) A claim is made that the population mean price (including taxes) is $326.76. Using the results from part (a) of this problem, would this claim be valid or invalid?

Valid

3) A lawyer must drive from a northern suburb to Chicago's Loop every day to get to work. In order to decide whether he should take the train instead, he computes his gas mileage each day for 34 days. He finds that he has a mean mileage of 12.2 miles per gallon, with an estimated standard deviation of 2.1 miles per gallon. Find a 95% confidence interval for his mileage, assuming that gas mileage is normally distributed. Interpret your result.

CL = 95%

$$12.2 \pm 2.035\left(\frac{2.1}{\sqrt{34}}\right)$$

$\bar{x}$ = 12.2

(11.47, 12.93)

n = 34

s = 2.1

I am 95% confident the population mean mileage is between 11.47 and 12.93mpg.

$t_{\alpha/2}$ = 2.035
df = 33

b) A claim is made that the population mean is 11.25 miles per gallon. Using the results from part (a) of this problem, would this claim be valid or invalid?

Invalid

4) A sample of 57 women ages 18-24 yields a mean systolic blood pressure of 117.2mm Hg and a standard deviation is 10.3mm Hg, find a 90% confidence interval for the population mean systolic blood pressure for all women aged 18-24. Interpret your result.

CL = 90%

$$117.2 \pm 1.673\left(\frac{10.3}{\sqrt{57}}\right)$$

$\bar{x}$ = 117.2

(114.92, 119.48)

n = 57

s = 10.3

I am 90% confident the population mean systolic blood pressure is between 114.92 and 119.48mmHg.

$t_{\alpha/2}$ = 1.673
df = 56

b) A claim is made that the population mean is 124.65mmHg. Using the results from part (a) of this problem, would this claim be valid or invalid?

Invalid

5) The Physician's Handbook provides statistics on heights and weights of children by age. The heights, in inches, of 67 randomly selected 6 year-old girls were obtained. The mean of the sample is 43.75 inches, and the standard deviation is 3.295 inches. Obtain a 95% confidence interval for the mean height of all 6 year-old girls. Interpret your result.

CL = 95%

$\bar{x}$ = 43.75

n = 67

s = 3.295

$t_{\alpha/2}$ = 1.997
df = 66

$$\mathbf{43.75 \pm 1.997\left(\frac{3.295}{\sqrt{67}}\right)}$$

(42.95, 44.55)

I am 95% confident the population mean height is between 42.95 and 44.55 inches.

b) A claim is made that the population mean height is 43.25 inches. Using the results from part (a) of this problem, would this claim be valid or invalid?

Valid

6) A real estate agent wants to estimate the population mean price of houses for sale in her area. A random sample of 46 houses had a mean sale price of $142,500 with a standard deviation of $2,350. Find a 98% confidence interval for the population mean price of houses. Interpret your result.

CL = 98%

$\bar{x}$ = 142,500

n = 46

s = 2350

$t_{\alpha/2}$ = 2.412
df = 45

$$\mathbf{142500 \pm 2.412\left(\frac{2350}{\sqrt{46}}\right)}$$

(141,664.27, 143,335.73)

I am 95% confident the population mean price of houses is between $141,664.27 and $143,335.73.

b) A claim is made that the population mean sale price is $142,950.36. Using the results from part (a) of this problem, would this claim be valid or invalid?

Valid