

## Activity 7

### Estimating and Finding Confidence Intervals

This activity begins with estimating the mean of a population with a normal distribution using a simple random sample from that population.

 It is important that you read Topic 33 *before* the other topics in Activity 7 because it explains the functions under the **STAT TESTS** menu, the notation, and the meaning of a “confidence interval.” You will be using these things in making other estimates.

The last topic in this activity, Topic 39, covers the idea of biased and unbiased estimators with a simulation that explains why we divide by **(n-1)** when calculating the sample standard deviation.

 The topic number of the related hypothesis test topic is given in parenthesis after the estimation topic number. For example, the (40) after Topic 33 below indicates that Topic 40 is the related hypothesis test topic for Topic 33.

### Topic 33 (40)—Estimating A Normal Population Mean $\mu$ ( $\sigma$ Known)

A random sample of size 10 from a population of heights that has a normal distribution (with  $\sigma = 2.5$  inches) is given below (with the sample mean).

Store this data in **L1**.

{66.71, 66.27, 62.81, 66.92, 62.91,  
71.42, 67.39, 63.79, 65.81, 62.81} = **L1**

$\bar{x} = 65.68$        $n = 10$

What is the 80 percent confidence interval for the population mean?

## Activity 7, Estimations and Confidence Intervals (cont.)

There are two possible methods of input, depending on how the data is presented to you in a problem. Sometimes all the data is given and it can be stored in a list. Sometimes only the summary stats are given, such as  $\bar{x}$  and  $n$ , and you must work with these. This activity shows you how to work with both.

1. Press **[STAT]** **<TESTS>** **7:ZInterval** for one of the two possible screens (see screen 1 or 3), depending on what input (**Inpt:**) is highlighted.
2. Enter the correct values from above, highlight **Calculate** in the last row, and then press **[ENTER]** for output screen 2 or 4.

We are 80 percent confident that the population mean lies between **64.67** and **66.7** inches.

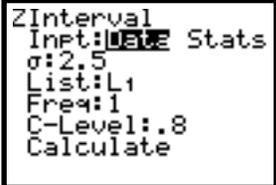
### Home Screen Calculations

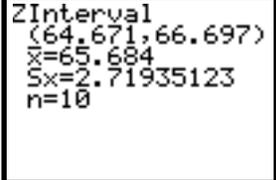
What formula was used by the above procedure?

1. With a confidence level of **C = 0.80**, there is  $(1 - 0.80)/2 = 0.10$ , or 10 percent of the area in each tail of the probability distribution of sample means.
2. To find the “critical  $z$  value” that divides the upper 10 percent from the lower 90 percent, use **[2nd]** **[DISTR]** **3:invNorm(0 [ ] 90 [ ] [ENTER])** for **1.28** as shown in screen 5.
3. The “margin of error” **E** is 1.28 times the standard deviation of the sample means, or  $E = 1.28 * \sigma/\sqrt{n} = 1.28 * 2.5/\sqrt{10} = 1.01$  as shown in screen 6.

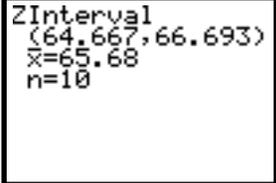
Note that the margin of error can be calculated from the confidence interval above by subtracting the mean from the larger value of the interval, or  $66.69 - 65.68 = 1.01$ .

The confidence interval becomes  $\bar{x} \pm E$ , or  $65.68 \pm 1.01$ , or  $65.68 - 1.01$  to  $65.68 + 1.01$ , or 64.67 to 66.69 as before.

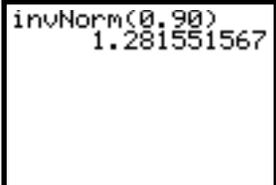
(1)  ZInterval  
Inpt: Stats  
 $\sigma$ : 2.5  
List: L1  
Freq: 1  
C-Level: .8  
Calculate

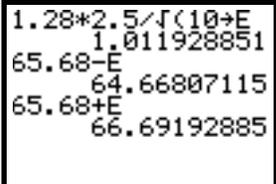
(2)  ZInterval  
(64.671, 66.697)  
 $\bar{x}$ : 65.684  
 $S_x$ : 2.71935123  
 $n$ : 10

(3)  ZInterval  
Inpt: Data  
 $\sigma$ : 2.5  
 $\bar{x}$ : 65.68  
 $n$ : 10  
C-Level: .8  
Calculate

(4)  ZInterval  
(64.667, 66.693)  
 $\bar{x}$ : 65.68  
 $n$ : 10

**Note:** The slight differences shown in the intervals in screen 2 and 4 are because the sample mean calculated from the Data List had three decimal places ( $\bar{x} = 65.684$ ) while the stats version was rounded to  $\bar{x} = 65.68$ .

(5)  invNorm(0.90)  
1.281551567

(6)  1.28\*2.5/sqrt(10)+E  
1.011928851  
65.68-E  
64.66807115  
65.68+E  
66.69192885

## The Meaning of Confidence Interval (Simulation)

The above sample was randomly generated from a normal distribution with a mean of **65** inches and a standard deviation of **2.5** inches (the first interval in screen 7), so the confidence interval of 64.67 to 66.69 did, in fact, contain the population mean. However, we take a sample to estimate the population mean because we do not know it. How confident can we be of our answer?

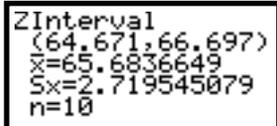
In the above case, we can be 80 percent confident because if we took all possible samples of size 10 and calculated the confidence intervals as above, 80 percent of these intervals calculated would contain the correct population mean. In screens 7-17 we generate just ten samples, and seven of the ten contain the true mean. (We did not expect exactly eight of ten for so few tries.) The intervals that miss the true mean (screens 11, 13, and 17) miss by 1.01, 0.37 and 0.57 inches.

The sample means are shown as gaps in the middle of the confidence interval lines drawn beneath the normal plot in screen 18 for the ten samples generated. Only when the sample means are in the shaded regions (10 percent in each tail) will the margin of error not be long enough to reach the population mean (the vertical line at **65** inches). Thus, 80 percent of the sample means in this distribution of sample means will be close enough to the population mean for its confidence interval to contain it.

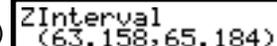
To generate the random samples and calculate the confidence intervals as was done in screens 7-17, proceed as follows.

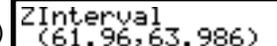
1. Set the random seed as explained in Topic 21 and shown in the first two lines of screen 7.
2. Complete screen 7 by pressing **[MATH] <PRB> 6:randNorm( 65  $\square$  2  $\square$  5  $\square$  10) [STO] L1 [ALPHA] [:]** (above the  $\square$ ); ties generating the sample and calculating the confidence together in one statement).
3. Press **[2nd] [CATALOG]**. Notice the **A** (in the upper right corner of the screen) indicating you are in alpha mode.
4. Press **Z  $\square$   $\square$   $\square$**  to point out **ZInterval**, and then press **[ENTER]**. This pastes it to the home screen. Finish the line by typing **2  $\square$  5  $\square$  L1  $\square$  0  $\square$  80**.
5. Press **[ENTER]** to generate the first sample and calculate the confidence interval (**64.671, 66.697**) as shown in screen 8.
6. Press **[ENTER]** to take the next sample and calculate the confidence interval (**63.237, 65.263**) as in screen 9. Continue pressing **[ENTER]** for the rest of the 10 intervals as in screens 10-17.

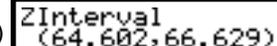
(7) 

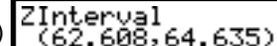
(8) 

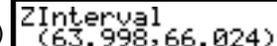
(9) 

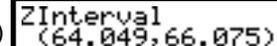
(10) 

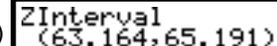
(11) 

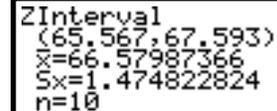
(12) 

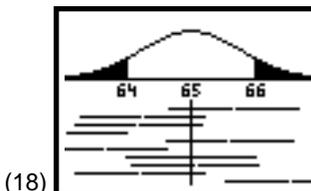
(13) 

(14) 

(15) 

(16) 

(17) 



**Activity 7, Estimations and Confidence Intervals (cont.)**

**Topic 34 (41)—Estimating a Population Mean  $\mu$  ( $\sigma$  Unknown)**

A random sample of size 10 from a population of heights that has a normal distribution is given below (with the sample mean and the standard deviation).

Store this data in **L1**.

{66.71, 66.27, 62.81, 66.92, 62.91, 71.42, 67.39, 63.79, 65.81, 62.81}  
= **L1**

$\bar{x}$  = 65.68, **Sx** = 2.72, **n** = 10

What is the 90 percent confidence interval for the population mean?

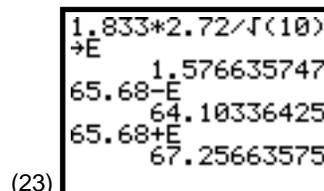
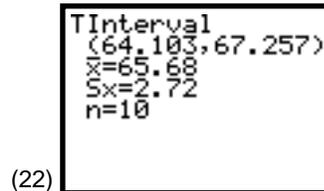
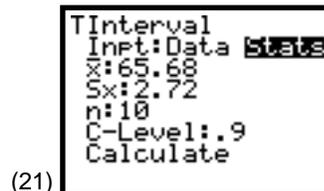
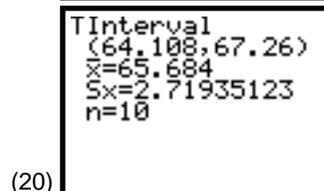
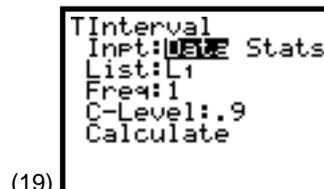
1. Press **[STAT]** <**TESTS**> **8:TInterval** for one of the two possible screens (see screen 19 or 21), depending on what input (**Inpt:**) is highlighted.
2. Enter the correct values as above, highlight **Calculate** in the last row, and then press **[ENTER]** for output screen 20 or 22.

We are 90 percent confident that the population mean lies between **64.1** and **67.26** inches.

**Home Screen Calculations**

To find the “critical *t* value” with the TI-83, there is no **invT** like the **invNorm**, so we will use the equation solver at the end of this topic to show a value of **1.833** (or you can look it up in a table).

The margin of error is calculated very much like in Topic 33, but now **Sx** is used instead of  $\sigma$  for a value of 1.58. Or, from the previous interval, subtract the mean from the upper interval value (67.26 - 65.68 = 1.58). From screen 23, we also see that the confidence interval is **64.10** to **67.26** as above, with the width of the interval 67.26 - 64.10 = 3.16, or twice the margin of error (2 \* 1.58).



Notice the interval above is wider than the one shown in screen 24 for an 80 percent confidence interval ( $66.87 - 64.50 = 2.37$ ). (The **Inpt** screen is not shown, but is like screen 19 with **C-Level: .8**.) The wider the interval for a given sample size, the more confident we are of the interval reaching the population mean.

(24)

Topic 33 used the same data but used a known  $\sigma$  and obtained an 80 percent confidence interval of 64.67 to 66.69 for a width of 2.02 inches ( $66.69 - 64.67 = 2.02$  inches). The interval above is wider for two reasons: the critical  $t$  is bigger than the critical  $z$  (1.833 compared to 1.28) and **Sx** is bigger than  $\sigma$  (2.72 compared to 2.50). The **Sx** will change from sample to sample, so it is possible that the interval width could sometimes be smaller than the  $\sigma$  case. If we consider all the possible variable width intervals, 80 percent will contain the true population mean.

### Using the Equation Solver for Critical t Values

For a 90 percent confidence interval, there is an area of  $(1 - 0.90)/2 = 0.05$  in each tail. We want to solve **tcdf(X, E99, 9) = 0.05** for  $X$ , the  $X$  that gives the area of 0.05 in the right tail (from  $X$  to **E99**) of a  $t$  distribution with nine degrees of freedom ( $n - 1 = 10 - 1 = 9$ ).

1. Press **[MATH] 0:Solver** for screen 25. If your screen does not start with **EQUATION SOLVER**, you must first press **[ $\Delta$ ]**.
2. Enter your equation using **[2nd] [DISTR] 5:tcdf(X [ ] E99 [ ] 9 [ ] - 0 [ ] 05**.
3. Press **[ENTER]**, and modify the screen to look like screen 26.

We pick **X = 2** for  $t$  as a reasonable first guess in the right tail. The bound will probably show as **{-1E99, 1E99}**, which works fine; although you could restrict it to positive values by making it **{0, 1E99}**.

4. With the cursor blinking on or after the **2**, press **[ALPHA] [SOLVE]**.

Notice the busy signal while the equation is being solved numerically. Then screen 27 appears showing a critical  $t$  value of **1.833**.

(25)

**Note:** We write the equation to be solved as  $0 = \text{tcdf}(X, E99, 9) - 0.05$  to meet the requirements of the Solver.

(26)

(27)

**Activity 7, Estimations and Confidence Intervals (cont.)**

**Topic 35 (42)—Estimating a Population Proportion**

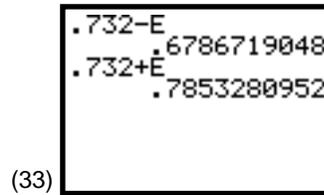
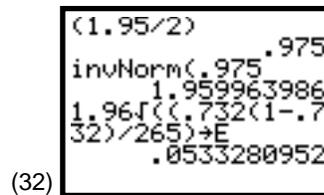
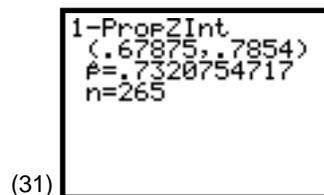
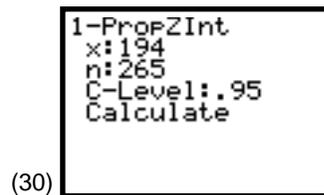
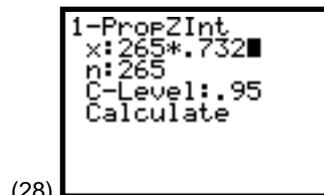
A random poll of 265 people from a population of interest found 73.2 percent agreement with a public policy. What is the 95 percent confidence interval for the proportion in the whole population who would agree? (Note that  $265 * .732 \geq 10$  and  $265(1-.732) \geq 10$ , so a normal distribution can be used to approximate a binomial.)

1. Press **[STAT]** <TESTS> **A:1-PropZInt** for screen 28.  
The input screen is looking for the number who agreed ( $x$ ), and this was not given in this particular problem.
2. Let the TI-83 do the calculation for you ( $265 * .732$ ). However, when you leave this line, the calculation reveals a noninteger, **193.98** (see screen 29).
3. Highlight **Calculate** in the last row, and press **[ENTER]**.  
A **DOMAIN** error results.
4. As shown in screen 30, round the value to the nearest integer (**194**), highlight **Calculate** in the last row, and press **[ENTER]** for screen 31.

We are 95 percent confident that between 67.88 to 78.54 percent of the population agree with the public policy.

**Home Screen Calculations**

The margin of error is  $.7854 - .7321 = .0533$  as shown in screen 32. The confidence interval is calculated in screen 33.



### Finding a Sample Size

What size sample would be needed to reduce the margin of error in the above problem from 0.0533 to 0.02, or 2 percent?

Solving for  $n$  in our margin of error equation, we have  $n = 1.96^2 * 0.732(1 - 0.732)/0.02^2 = 1885$  (rounding up). (See screen 34.)

The original sample was generated (as described in Topic 25 and shown in screen 35) from a population with a population proportion of 0.70. Thus the interval obtained (67.88 to 78.54 percent) does contain the 70 percent. (See screen 35.)

For the larger sample ( $n = 1885$ ) shown in screen 36, the resulting 0.713 is within 0.02 of 0.70.

```
1.96^2*0.732(1-0.732)/0.02^2
1884.074304
```

(34)

```
321→rand
randBin(265,0.7)
/265
.7320754717
```

(35)

```
321→rand
randBin(1885,0.7)
/1885
.7129973475
```

(36)

**Activity 7, Estimations and Confidence Intervals (cont.)**

**Topic 36 (43)—Estimating a Normal Population Standard Deviation  $\sigma$**

A random sample of size 4 from a population that is normally distributed is as follows.

{54.54, 43.44, 54.11, 46.88} = L1.

Find the 90 percent confidence interval for the population standard deviation.

The confidence interval is based on the sample variance and the  $\chi^2$  distribution (the similarity between the distribution of sample variance and the  $\chi^2$  distribution can be seen in Topic 39).

A. Variance = **30.00**.

With the data in L1, calculate the variance of the data by pressing  $\boxed{2\text{nd}} \boxed{[LIST]} \boxed{<MATH>8:variance(}$  L1  $\boxed{ENTER}$  for **30.00** as in screen 37.

With  $n = 4$  pieces of data, we will be using  $n - 1 = 4 - 1 = 3$  degrees of freedom for the  $\chi^2$  distribution. The plot is given in screen 40 (with setup as in screens 38 and 39).

As you can see, the plot is skewed to the right. Because it is not symmetrical, two critical values are needed,  $\chi^2_L$  and  $\chi^2_R$ , for left and right. (Shading was done with  $\boxed{2\text{nd}} \boxed{[DRAW]} \boxed{7:Shade(}$  0  $\boxed{,}$  Y1  $\boxed{,}$  0  $\boxed{,}$  0  $\boxed{,}$  35  $\boxed{,}$  and  $\boxed{Shade(}$  0  $\boxed{,}$  Y1  $\boxed{,}$  7  $\boxed{,}$  8  $\boxed{,}$   $\boxed{E99.}$ ) Thus, 5 percent of each tail is shaded. See statement 2.)

B. Critical values are  $\chi^2_L = \mathbf{0.352}$  and  $\chi^2_R = \mathbf{7.814}$ .

You can look up the critical values in a table (with 0.05 in each tail of the distribution), or you can obtain them using the equation solver as explained at the end of this topic.

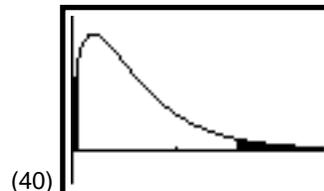
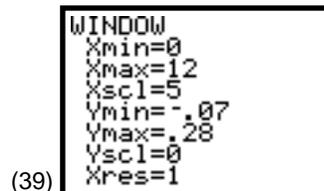
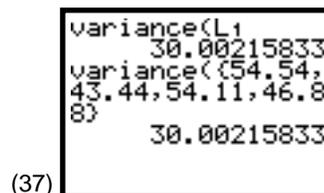
C. Confidence intervals for variance = 11.52, 255.68.

Lower value =  $(n - 1)Sx^2/\chi^2_R = 3 * 30.00/7.814 = \mathbf{11.52}$   
Upper value =  $(n - 1)Sx^2/\chi^2_L = 3 * 30.00/0.352 = \mathbf{255.68}$

D. Confidence interval for standard deviation = 3.39, 15.99.

Lower value =  $\sqrt{((n - 1)Sx^2/\chi^2_R)} = \sqrt{(11.52)} = \mathbf{3.39}$   
Upper value =  $\sqrt{((n - 1)Sx^2/\chi^2_L)} = \sqrt{(255.68)} = \mathbf{15.99}$

The above sample was the first of 100 randomly generated samples in Topic 39 from a normal distribution with a variance  $\sigma^2 = 100$  and a standard deviation  $\sigma = 10$ . These, in fact, do lie in the intervals calculated. The intervals are quite wide, but they are based on a very small sample.



## Using the Equation Solver for Critical Values $\chi^2L$ and $\chi^2R$

We want to solve the following equations for  $X$  such that the area (from 0 to  $X$  and from  $X$  to the very large number **E99**) is **0.05** under a  $\chi^2$  distribution with  $n - 1 = 4 - 1 = 3$  degrees of freedom.

(i)  $\chi^2\text{cdf}(0, X, 3) = 0.05$

(ii)  $\chi^2\text{cdf}(X, \text{E99}, 3) = 0.05$

For  $\chi^2L$ :

1. Press **[MATH] 0:Solver** for screen 42. If your screen does not start with **EQUATION SOLVER**, you must first press **[ $\Delta$ ]**.
2. To meet the requirements of the solver, write equation (i) as in screen 42 using **[2nd] [DISTR] 7:  $\chi^2\text{cdf}$** .
3. Press **[ENTER]** and modify the screens to look like screen 43.

We pick **X = 1** for our first guess for the left tail. The bounds will probably show as **{-1E99, 1E99}**, which works fine, although we could restrict them to positive values **{0, 1E99}** because  $\chi^2$  cannot be negative.

4. With the cursor blinking on or after **X = 1**, press **[ALPHA] [SOLVE]**.

The critical value **X = 0.352** =  $\chi^2L$  is calculated. See screen 44. (During calculations, a busy signal shows in the upper-right corner of the screen.)

For  $\chi^2R$ :

Repeat the above procedure for the second equation (ii) as in screen 45. Use a first guess of **10**, as in screen 46, for the critical value  $\chi^2R = 7.815$  as in screen 47.

(42) EQUATION SOLVER  
eqn:0= $\chi^2\text{cdf}(0,X,3)-0.05$

(43)  $\chi^2\text{cdf}(0,X,3)-...=0$   
X=1  
bound={-1E99,1E99}

(44)  $\chi^2\text{cdf}(0,X,3)-...=0$   
X=.35184631776...  
bound={-1E99,1E99}  
left-rt=0

(45) EQUATION SOLVER  
eqn:0= $\chi^2\text{cdf}(X,E99,3)-0.05$

(46)  $\chi^2\text{cdf}(X,E99,3)-...=0$   
X=10  
bound={-1E99,1E99}

(47)  $\chi^2\text{cdf}(X,E99,3)-...=0$   
X=7.8147279080...  
bound={-1E99,1E99}  
left-rt=0

## Activity 7, Estimations and Confidence Intervals (cont.)

### Topic 37 (44)—Estimating the Difference in Two Population Means

#### Estimating Independent Samples

A study designed to estimate the difference in the mean test scores that result from using two different teaching methods obtained the following data from two random samples of students taught with the two methods.

						mean	Sx	n	
Method A (L1)	40	38	39	32	35	38	37	2.966	6
Method B (L2)	34	36	31	37	29	33.4	3.362	5	

Store the above data in L1 and L2, and calculate the 95 percent confidence interval for the difference in the population means.

#### TInterval ( $\sigma_1$ and $\sigma_2$ Unknown)

Assume the above data is from normal populations.

1. Press **[STAT]** <TESTS> **0:2-SampTInt** for screens 48 and 50 (48↓ and 50↓ are the second page of the input screens obtained by using **[↓]**).

Screens 48 and 50 do not pool the sample variances as does screen 52. (For the pooled case, see “To Pool or Not to Pool” on the next page.) Screen 48 uses the **Data List**. However, for screen 50, you must enter the summary **Stats** of means, standard deviation, and sample sizes.

2. Make the entries as shown in screen 48 or 50, highlight **Calculate**, and press **[ENTER]** for screens 49 or 51.

The 95 percent confidence interval for the difference of the means is **-0.839 to 8.039** for the non-pooled case. Because zero is in the interval, there could be no difference in the mean test scores for the two groups, or Method A could have a higher mean than Method B (a positive difference), or Method B could have a higher mean than Method A (a negative difference). Perhaps larger samples would clarify the situation.

(48) 2-SampTInt  
Inpt: **DATA** Stats  
List1:L1  
List2:L2  
Freq1:1  
Freq2:1  
C-Level:.95  
↓Pooled:**NO** Yes  
(48)↓ Calculate

(49) 2-SampTInt  
(-.8388,8.0388)  
df=8.135354495  
x1=37  
x2=33.4  
Sx1=2.96647939  
↓Sx2=3.36154726  
n1=6  
n2=5  
(49)↓

(50) 2-SampTInt  
Inpt:Data **stats**  
x1:37  
Sx1:2.966  
n1:6  
x2:33.4  
Sx2:3.362  
↓n2:5  
C-Level:.95  
Pooled:**NO** Yes  
(50)↓ Calculate

(51) 2-SampTInt  
(-.839,8.039)  
df=8.133989496  
x1=37  
x2=33.4  
Sx1=2.966  
↓Sx2=3.362  
Sx2=3.362  
n1=6  
n2=5  
(51)↓

### To Pool or Not to Pool

Screens 52 and 53 give the input and output for the pooled procedure with a 95 percent confidence interval of **-0.7124** to **7.9124**.

The pooled procedure is appropriate if  $\sigma_1 = \sigma_2$ . It is not appropriate to pool if  $\sigma_1 \neq \sigma_2$ .

Some statistics textbooks only give the pooled version because before the TI-83, the noninteger value for degrees of freedom that results from the non-pooled procedure was difficult to handle without a computer.

If you are not sure that  $\sigma_1 = \sigma_2$ , then it is safe to use the non-pooled procedure because it will always give a conservative answer. (The confidence interval will be a little wider than in the pooled case (-0.839 to 8.039 compared to -0.7124 to 7.9124.))

### ZInterval ( $\sigma_1$ and $\sigma_2$ Known)

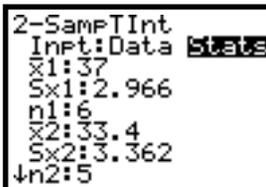
Assume the above data is from normal populations but that  $\sigma_1 = 2.966$  and  $\sigma_2 = 3.362$ .

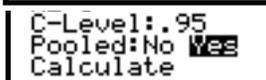
1. Press **[STAT]** **<TESTS>** **9:2-SampZInt** for screen 54.

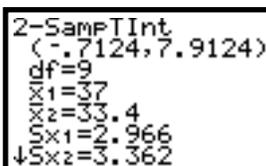
The **Data Lists** input is also possible, but only the **Stats** version is shown. Some texts use this procedure for large samples when  $\sigma_1$  and  $\sigma_2$  can be approximated by **Sx1** and **Sx2**.

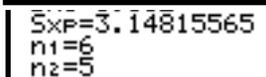
2. Make the entries as shown in screen 54, highlight **Calculate**, and then press **[ENTER]** for the output shown in screen 55.

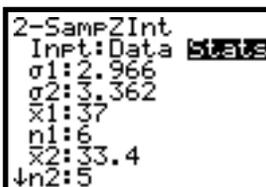
The 95 percent confidence interval for the difference of the means is **-0.184** to **7.384**. This interval is similar to, but not as wide as, the **Intervals** above.

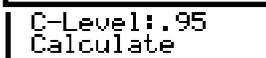
(52) 

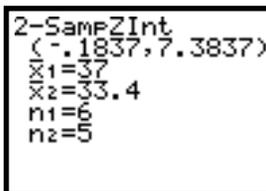
52↓ 

(53) 

53↓ 

(54) 

54↓ 

(55) 

## Activity 7, Estimations and Confidence Intervals (cont.)

### Home Screen Calculations

Screens 56, 57, and 58 show the non-pooled **TInterval**. First, the degrees of freedom (**8.134**) are calculated; second, with  $(1 - .95)/2 = 0.025$  in each tail, the critical  $t$  value of **2.299** is checked. (The critical  $t$  value is calculated by using the equation solver to solve  $\text{tcdf}(X, \text{E99}, 8.134) = 0.025$  for  $X = 8.134$ , as explained in "Home Screen Calculations" in Topic 34.) Next, the margin of error and confidence interval are calculated.

Screens 59, 60, and 61 are for the pooled case with degrees of freedom  $= n_1 + n_2 - 2 = 6 + 5 - 2 = 9$  and a critical value of **2.262** (from a table or using the procedure explained in Topic 34).

Screens 62 and 63 are for the **ZInterval**. Note that the calculation for the **ZInterval** confidence interval is the same as for the non-pooled **TInterval** except for the differences in the critical  $z$  and  $t$  values.

(56) 
$$\begin{array}{r} 2.966^2/6 \rightarrow A \\ 1.466192667 \\ 3.362^2/5 \rightarrow B \\ 2.2606088 \\ (A+B)^2 / (A^2/(6-1) \\ + B^2/(5-1)) \\ 8.133989496 \end{array}$$

(57) 
$$\begin{array}{r} \text{tcdf}(2.299, \text{E99}, 8 \\ .134 \\ .0250160486 \end{array}$$

(58) 
$$\begin{array}{r} 2.299 \sqrt{(2.966^2/6 + \\ 3.362^2/5) \rightarrow E \\ 4.438202361 \\ (37 - 33.4) - E \\ -.8382023612 \\ (37 - 33.4) + E \\ 8.038202361 \end{array}$$

(59) 
$$\begin{array}{r} 6+5-2 \\ \text{tcdf}(2.262, \text{E99}, 9 \\ .0250064226 \end{array}$$

(60) 
$$\begin{array}{r} \sqrt{((6-1)2.966^2 + ( \\ 5-1)3.362^2) / 9} \rightarrow S \\ 3.148155651 \end{array}$$

(61) 
$$\begin{array}{r} 2.262 * S \sqrt{(1/6 + 1/5 \\ ) \rightarrow E \\ 4.312057192 \\ (37 - 33.4) - E \\ -.7120571915 \\ (37 - 33.4) + E \\ 7.912057192 \end{array}$$

(62) 
$$\begin{array}{r} \text{invNorm}(.975 \\ 1.959963986 \end{array}$$

(63) 
$$\begin{array}{r} 1.96 \sqrt{(2.966^2/6 + 3 \\ .362^2/5) \rightarrow E \\ 3.783765388 \\ (37 - 33.4) - E \\ -.1837653884 \\ (37 - 33.4) + E \\ 7.383765388 \end{array}$$

## Estimating Dependent Samples

To test a blood pressure medication, the diastolic blood pressure readings of a random sample of ten people with high blood pressure were recorded. After a few weeks on the medication, their pressures were recorded again.

See the data recorded in the table below. The mean and the standard deviation of the differences in  $L_3$  were calculated with **1-Var Stats** as explained in Topic 4 and shown in screens 64 and 65. Note that the mean of the differences is **5.9** and the standard deviation of the differences is **5.666**.

Subject	1	2	3	4	5	6	7	8	9	10
<b>Before (L1):</b>	94	87	105	92	102	85	110	95	92	93
<b>After (L2):</b>	87	88	93	87	92	88	96	87	92	86
<b>L1 - L2 = L3</b>	7	-1	12	5	10	-3	14	8	0	7

Assume the diastolic blood pressure readings are normally distributed, and calculate the 95 percent confidence interval for the mean difference in pressure.

1. Press **[STAT]** **<TESTS>** **8:TInterval** for one of the two possible screens (see screen 66 or 68), depending on what input (**Inpt:**) is highlighted.
2. Put in the correct values as above, highlight **Calculate** in the last row, and then press **[ENTER]** for the output in screens 67 or 69.

We are 95 percent confident that the difference in the population mean lies between 1.85 and 9.95 units of pressure, and this indicates that the medication seems to have some effect.

Home screen calculations are similar to those done in Topic 34.

(64) **1-Var Stats L3**

```
1-Var Stats
x̄=5.9
Σx=59
Σx²=637
Sx=5.66568619
σx=5.37494186
↓n=10
```

(65)

```
TInterval
Inpt: DATA Stats
List:L3
Freq:1
C-Level:.95
Calculate
```

(66)

```
TInterval
(1.847,9.953)
x̄=5.9
Sx=5.66568619
n=10
```

(67)

```
TInterval
Inpt:Data STATS
x̄:5.9
Sx:5.666
n:10
C-Level:.95
Calculate
```

(68)

```
TInterval
(1.8468,9.9532)
x̄=5.9
Sx=5.666
n=10
```

(69)

Activity 7, Estimations and Confidence Intervals (cont.)

Topic 38 (45)—Estimating the Difference in Two Population Proportions

A polling organization found 694 of 936 ( $694/936 = 74.15$  percent) women sampled agreed with a public policy while 645 of 941 (68.54 percent) men agreed. Find the 95 percent confidence interval for the difference in the proportions between the two populations sampled.

Women:  $x_1 = 694$   $n_1 = 936$   
Men:  $x_2 = 645$   $n_2 = 941$

1. Press **[STAT]** **<TESTS>** **B:2-PropZInt** for screen 70.
2. Enter the correct values (as in screen 70), highlight **Calculate** in the last row, and then press **[ENTER]** for the output in screen 71.

We are 95 percent confident that the difference in population proportions lies between 1.52 and 9.68 percent. This indicates that a larger proportion of women are in agreement with the policy than are men.

Home screen calculations are given in screens 72 and 73 with the critical  $z$  value of **1.96** and a margin of error of **4.08** percent. These calculations also could have been obtained from screen 71 by taking  $0.0968 - (0.7415 - 0.6854) = .0407$ . The confidence interval from 1.53 to 9.69 percent differs slightly from the above because of rounding.

**Note:**  $n_1 = 936 = 694 + 242$  (both values greater than 5).  $n_2 = 941 = 645 + 296$  (both values greater than 5). Therefore, the normal approximation to the binomial can be used.

```
(70) 2-PropZInt
      x1:694
      n1:936
      x2:645
      n2:941
      C-Level:.95
      Calculate
```

```
(71) 2-PropZInt
      (.01518,.09684)
      p1=.7414529915
      p2=.6854410202
      n1=936
      n2=941
```

```
(72) invNorm(.975
      1.959963986
      1.96√(.7415(1-.7
      415)/936+.6854(1
      -.6854)/941)→E
      .0408287526
```

```
(73) (.7415-.6854)-E
      .0152712474
      (.7415-.6854)+E
      .0969287526
```

## Topic 39—Unbiased and Biased Estimators (Simulation)

 For this topic, set the mode for two decimal places. (Refer to “Setting Modes” under Do This First.)

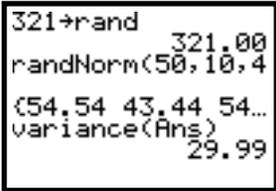
A sample statistic used to estimate a population parameter is *unbiased* if the mean of the sampling distribution of the statistic is equal to the true value of the parameter being estimated.

You saw in Topic 24 that the sample proportion is an unbiased estimator of the population proportion. In Topic 25, you saw that the sample mean is an unbiased estimator of the population mean. In this topic, you will see that  $\Sigma(x - \bar{x})^2/n$  is a *biased* estimator of the population variance while  $\Sigma(x - \bar{x})^2/(n - 1)$  is unbiased.

Because  $Sx^2$ , the “variance” on the TI-83, is unbiased (divides by  $n - 1$ ), you will need to change to biased (maximum likelihood estimator) by multiplying by  $(n - 1)/n$  as  $((n - 1)/n) * \Sigma(x - \bar{x})^2/(n - 1) = \Sigma(x - \bar{x})^2/n$ . For  $n = 4$ , you would multiply by  $3/4$  or  $0.75$ .

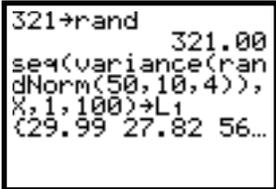
After making the above change, proceed as follows.

1. Set the random seed as explained in Topic 21 and shown in the first two lines in screen 74.
2. Press **[MATH] <PRB>6:randNorm( 50 **[.]** 10 **[.]** 4 **[ENTER]**** to simulate picking four values from a normal distribution with a mean of 50 and a standard deviation of 10 or a variance of  $10^2 = 100$ . The values are **54.54, 43.44, 54.11,** and **46.88** (middle values in screen 74).
3. Press **[2nd] [LIST] <MATH> 8:variance( [2nd] [ANS] **[ENTER]**** for **29.99**, an unbiased variance of these values (last value in screen 74).
4. Reset the seed and generate a sequence of 100 sample variances. Store these in **L1**, as in screen 75, with **seq** pasted from **[2nd] [LIST] <OPS> 5**. Notice that the first variance is **29.99** as above.
5. Obtain the mean of **L1** by pressing **[2nd] [LIST] <MATH> 3:mean( L1 **[ENTER]****, as in screen 76. The **99.32** is very close to the population value of 100.
6. Set up **Plot1** for a **Histogram** of the data in **L1** (as explained in Topic 1) with the **WINDOW** shown in screen 77 for the results in screen 78.

(74) 

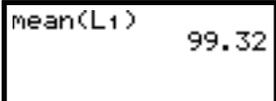
```

321→rand      321.00
randNorm(50,10,4
(54.54 43.44 54...
variance(Ans)
29.99
    
```

(75) 

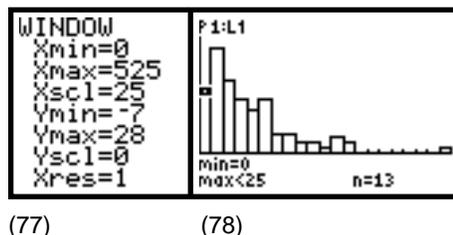
```

321→rand      321.00
seq(variance(ran
dNorm(50,10,4)),
X,1,100)→L1
(29.99 27.82 56...
    
```

(76) 

```

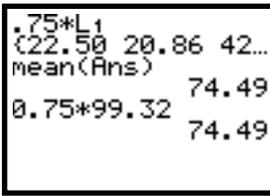
mean(L1)      99.32
    
```



**Activity 7, Estimations and Confidence Intervals (cont.)**

For the biased estimator, you can multiply each of the unbiased calculations for the variance by **0.75** as explained earlier and as shown in screen 79. These have a mean of **74.49**, which is not near the population value of 100.

The remaining screens (80, 81, and 82) give the procedure and results if you had started from scratch in calculating the biased estimator with the mean and a **Histogram** of your sample distribution. Note that the sample distribution is skewed to the right. This is not surprising when you realize that the left is bounded by zero because the variance is a sum of squares and, thus, can never be negative. The  $\chi^2$  distribution is related to this distribution and is discussed in Topic 36.

(79) 

(80) 

(81) 