

Data Analysis With The
TI-83/84
Graphing Calculator

Third Edition

Brian Jean

Copyright © 2000-2008 by 3RingPublishing.com

Printed in the United States of America

All rights reserved. No part of this work may be reproduced or used in any form or by any means - graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems - without the written permission of the publisher.

ISBN 0-9716560-8-8

Preface

The purpose of this manual is to provide students with step-by-step instructions for their TI-83/84 graphing calculators in an introductory level statistics course. This manual covers most of the statistical functions native to the TI-83/84 that are commonly used in an introductory statistics course. In addition, the manual covers step-by-step instructions for using various statistical programs that have been written for the TI-83/84. All programs are available free of charge; however, you will need the appropriate cable and software from Texas Instruments (or a third party cable) to transfer the programs from your computer to your calculator. We will do our best to keep up to date file transfer instructions available on line at <http://www.3RingPublishing.com>.

For additional information/comments, contact the publisher at:

<http://www.3RingPublishing.com>

CHAPTER 1	<i>Introduction To The TI-83/84 Graphing Calculator</i>	9
	Setting The Contrast	9
	Setting The StatEditor	9
	<i>Going To The StatEditor / Entering Data</i>	10
	<i>Entering a named data list</i>	11
	Order of Operations	11
	Transferring Data Lists and Programs Between Calculators	12
	Deleting Lists and Programs From The TI-83/84	14
	Deleting Lists and Programs From The TI-83/84 Plus	15
	Basic Descriptive Statistics	15
	Transferring Programs to Your Calculator	17
CHAPTER 2	<i>TI-83/84 Graphical Displays</i>	19
	Bar Graphs and Histograms	19
	Using the Trace Command	21
	Using Freq	22
	Box Plots	22
	Scatter Plots	23
	xyLine Plot	24
	Normal Plots	25
CHAPTER 3	<i>Generating Random Numbers / Simulations</i>	27
	Generating Integer Valued Random Data	27
	Generating Random Data from a Normal Distribution	27
	Generating Random Data from a Binomial Distribution	28
	Monte Carlo Method of Simulation	29
	<i>Monte Carlo Simulation Example</i>	29
CHAPTER 4	<i>The TI-83/84 and Binomial Probabilities</i>	33
	The Binomial Probability Distribution	33
	The Binomial Cumulative Probability Distribution	34
CHAPTER 5	<i>The TI-83/84 and Normal Probabilities</i>	37
	Probabilities Based on the Standard Normal	37
	Non-standard Normal Probabilities	41

CHAPTER 6

Univariate Hypothesis Testing and Confidence Intervals **43**

Estimation of the mean when the standard deviation is known **43**

Hypotheses test of the mean when the standard deviation is known **45**

EPA Example **45**

Inferences about the mean when the standard deviation is unknown (Student - t) **46**

Hospital Example **46**

Hypotheses test of the mean when the standard deviation is unknown **48**

Return to the EPA Problem **49**

Inferences Regarding Proportions **49**

Campus Safety Example **50**

Hypothesis Testing for Proportions **51**

Campus Safety Example **51**

CHAPTER 7

Comparing Two Parameters **53**

Inferences Concerning The Mean Difference Using Two Dependent Samples **53**

The Difference Between Means Using Two Independent Samples **54**

Inferences Concerning The Ratio of Variances Using Two Independent Samples (F-test) **54**

2-Sample t-test for Independent Data **55**

The Difference of Two Proportions **56**

Hypothesis Test **56**

Confidence Interval **57**

CHAPTER 8

Correlation and Simple Linear Regression **59**

Finding The Regression Line / Pearson's Correlation **59**

A Nonparametric Alternative to Pearson's Correlation **61**

Checking Assumptions **61**

Inferences **62**

CHAPTER 9

Analysis of Variance **63**

One-Way ANOVA Example **63**

Calculating One-Way ANOVA on the TI-83/84 **65**

Nonparametric Alternative: The Kruskal-Wallis Test **65**

Multiple Comparisons - Fisher's LSD **66**

CHAPTER 10

Categorical Data Analysis **67**

Chi-Squared Goodness of Fit **67**

Chi-Squared Test of Independence **67**

CHAPTER 11

Programs **69**

CLEAN **69**

FLSD (Fisher's Least Significant Differences) **69**

FREQDIST (Frequency Distribution Table) **70**

GOODFIT **71**

Stored In List **71**

Need To Input **72**

KWTEST (Kruskal-Wallis Test) **72**

NORMPLOT (Normal Probability Plot) **73**

NRMHST (Histogram with Normal Curve) **74**

PREDCI (Regression - Confidence Interval for a Predicted Value) **75**

ODDSCI (Odds Ratio Confidence Interval) **75**

RANKSUM (Wilcoxon Rank-Sum Test) **76**

RHO CI (Correlation Coefficient Confidence Interval) **77**

SAMPLESZ (Sample Size) **77**

SHADNORM (Shaded Normal Curve) **78**

SIGNRANK (Wilcoxon Sign-Rank Test) **79**

SIGNTEST (One Sample Sign Test) **79**

SRCORR (Spearman's Correlation) **80**

STEMPLOT (Stem-and-leaf Display) **80**

TTESTRHO **81**



Setting The Contrast

When you turn your TI-83/84 calculator on for the first time, you will find yourself on what is referred to as the *homescreen*. It is simply a blank screen. It is here that you will do typical mathematical operations such as addition, subtraction, multiplication, and division, along with the execution of many of the TI-83/84 commands.

Initially, you may find the screen contrast (darkness) to be ideal for your eyes or may you find it difficult to read. Adjusting the contrast is a simple task. Start with pressing the **STAT** button. The reason for this is simply to put something on the screen to look at while the contrast is changing. Your screen should now look like the graphic shown here.



To change the contrast, press the yellow **2nd** key. Next press *and hold down* the blue up arrow key or the blue down arrow key. Holding down the **up arrow key** will make the display darker, whereas holding down the **down arrow key** will make the display lighter.

Suppose you are making your display darker and go too far. Once you release the up arrow key, you must press the **2nd** key again prior to pressing the down arrow key to make the display lighter.

Once you have the contrast set to your liking, we can continue to set up your calculator for statistical use.

Setting The StatEditor

The TI-83/84 handles data in what it refers to as *data lists*. You can name your data lists anything you like (as long as you don't use more than five characters) or use the six default lists. The six default lists are very convenient to use and are prenamed **L1** through **L6**. Notice the yellow **L1** through **L6** above the 1 through 6 keys on your calculator.



First we will set up the statistical editor to display **L1** through **L6** every time we go to the editor. To do this, press **STAT**. Notice the bottom line reads **5: SetUpEditor**.

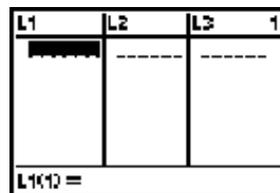
The option can be selected by using the down arrow key until that option is highlighted and then pressing **ENTER** or simply by pressing the **5** key. Once the option is selected, the command is pasted to the home screen.



Pressing **ENTER** once again will set the StatEditor in default mode, which displays all six lists. If you wanted only **L1** and **L4** to be displayed, then you could specify so by pressing **2nd 1**. The “1” key will paste L1 to the screen after the yellow 2nd key has been pressed. Next enter a comma “,” which is found directly above the “7” key, followed by **2nd 4**. The “4” will paste L4 to the homescreen. Now pressing **ENTER** will result in the editor displaying only L1 and L4.

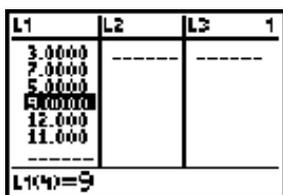
Going To The StatEditor / Entering Data

The StatEditor is accessed by pressing **STAT**, which brings you to the main Statistics Menu. Notice the first option is **1: Edit**.

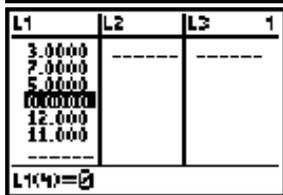


Pressing **ENTER** will bring you to the StatEditor as shown here.

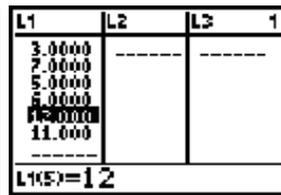
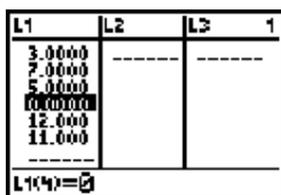
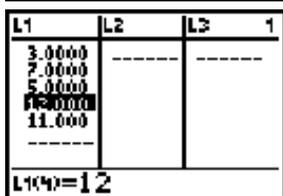
The StatEditor behaves much like a spread sheet. Suppose we wanted to enter the data values 3, 7, 5, 9, 12, 11 into list L1. Simply type each number, followed by either **ENTER** or the down arrow key. Using the up and down arrow keys, we can navigate through our data list, changing values at will. Suppose the “9” was a mistake and should have been a “0.” By going to the 9 and typing a zero, the 9 will be replaced by “0.”



Suppose the “0” was a mistake and needs to be removed from the data set. To do so, highlight the zero and press the **DEL** key. That value will then be removed from the entire list.



If we now need to add the value 6 to the data list, we could either use the blue down arrow key to highlight the space below the 11 and type in the 6, or we could insert it anywhere in the list we wanted. If we wanted it before the 12, then we would highlight the 12 and then press **2nd INS**. Notice **INS** is in yellow above the **DEL** key. This will shift the 12 down one position in the list and insert a “0” in its place. Now simply change the 0 to a 6.



Entering a named data list

Suppose you wanted to name a data set “Stuff” and wanted it to be displayed between L1 and L2. Press the blue right arrow and up arrow keys until L2 is highlighted. Next, press **2nd INS**. A new list will be added to the display that does not have a name. On the bottom of the screen, you will notice the TI-83/84 is asking you for a name. Type STUFF. Notice the alphabet is listed in green and is in alphabetical order, not the same order you would find the letters on a typical computer keyboard.

L1	DATA	L2	2
3.0000	-----	-----	
7.0000			
5.0000			
6.0000			
12.000			
11.000			

L2 =			

L1	DATA	L2	2
3.0000		-----	
7.0000			
5.0000			
6.0000			
12.000			
11.000			

Name=			

Once you type the name “STUFF,” you must press **ENTER** to finish naming the list.

To delete STUFF from the list, simply highlight the list name (as shown above to the right) and press the **DEL** key. This does not remove the data list STUFF from your calculator, it simply removes it from the StatEditor display screen. The list is still stored in your calculator.

L1	DATA	L2	2
3		-----	
7			
5			
6			
12			
11			

Name=STUFF			

L1	DATA	L2	2
3		-----	
7			
5			
6			
12			
11			

STUFF =			

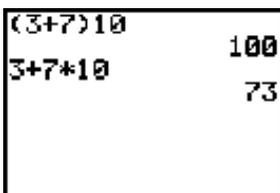
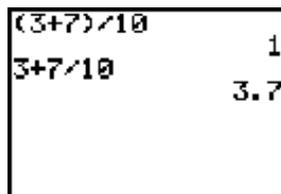
L1	DATA	L2	2
3.0000	-----	-----	
7.0000			
5.0000			
6.0000			
12.000			
11.000			

L2 =			

Order of Operations

Order of operations is important in mathematics and likewise with the TI-83/84 calculator. Consider the fraction $\frac{3+7}{10}$. This fraction simplifies to $\frac{10}{10} = 1$; however, if the fraction is not entered properly in the calculator, you could easily get an answer of 3.7.

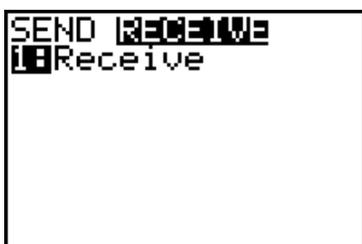
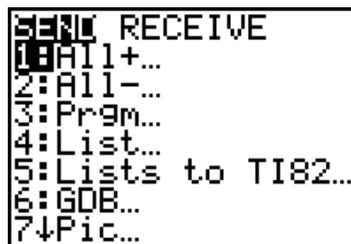
The most common mistake when entering fractions is failing to tell the calculator what the numerator and denominator is. To properly enter the above fraction, we would enter it using parentheses to encompass the numerator. This will assure us of the proper result. The following graphic shows the result of entering this fraction in the calculator, both with and without the proper “()”.



A similar situation arises with multiplication. If we want the quantity $(3 + 7)$ multiplied by 10 , then we must enter it as $(3+7)10$. Consider the following graphic with and without the proper set of “()”. The two answers clearly differ.

Transferring Data Lists and Programs Between Calculators

It is often useful to share a data list or program with another user. To do this, attach the data cable to the bottom of both calculators. Make sure the cable is secure by pressing inward firmly. Now press **2nd LINK** on both calculators. Suppose I have five named lists that I want to send to the other calculator. Those lists are BEERS, CHAL, FRMGH, PRES and STUFF. On the *receiving calculator*, press the blue right arrow one time. This will highlight **RECEIVE**, which has only one option.



Do not select **Receive** just yet. It is better to wait until you are actually ready to send.

On the *sending calculator*, either press the blue down arrow key until **4:List** is highlighted or simply press the number **4**. This will bring us to a screen that contains the names of all lists that are currently stored in our calculator. Recall earlier we said when we deleted the list STUFF from the StatEditor that it was not deleted from the calculator, only from the editor view. The data list STUFF is still in the calculator, as

shown here. Pressing the down arrow key will allow you to scroll through the list names.

Transferring Data Lists and Programs Between Calculators

Notice the arrow is pointing at the data list named BEERS. To select this list for transfer, simply press **ENTER**, and then scroll down to the next list name. The selected data list is indicated by a black square. Continuing this process, we can select all the desired data lists. Once the lists are selected, we are ready to transmit.

```
SELECT TRANSMIT
L1 LIST
L2 LIST
L3 LIST
L4 LIST
L5 LIST
L6 LIST
BEERS LIST
```

On the *receiving calculator*, press **ENTER**. The calculator will respond with “Waiting... .” On the *sending calculator*, press the blue right arrow once. The calculator will commence to transmit when you press **ENTER**.

```
SELECT TRANSMIT
L5 LIST
L6 LIST
BEERS LIST
CHAL LIST
FRMGH LIST
PRES LIST
STUFF LIST
```

The names of the lists being transmitted will be displayed on both calculator screens. If you get a transmission error, then check the following.

```
SELECT TRANSMIT
L5 LIST
L6 LIST
BEERS LIST
CHAL LIST
FRMGH LIST
PRES LIST
STUFF LIST
```

- Is the data transfer cable securely attached? Give it a good snug push to make sure.
- Did you select **Receive** on the receiving calculator *before* you selected **Transmit** on the sending calculator? Selecting the transmit command prior to the receive command will result in an error.

```
SELECT TRANSMIT
L5 LIST
L6 LIST
BEERS LIST
CHAL LIST
FRMGH LIST
PRES LIST
STUFF LIST
```

Once you have checked these two items, start the process over again.

```
SELECT TRANSMIT
Transmit
```

Transferring programs to another calculator involves the exact same steps except **3:Prgm** is selected rather than **4:List** on the initial link screen.

```
RECEIVE
1:All+...
2:All-...
3:Prgm...
4:List...
5:Lists to TI82...
6:GDB...
7↓Pic...
```

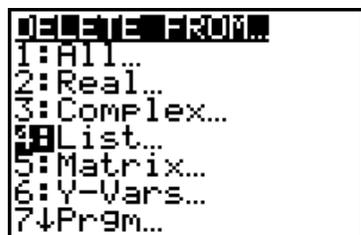
Deleting Lists and Programs From The TI-83/84

This procedure is slightly different with the TI-83/84 Plus. The TI-83/84 Plus is discussed next. To *permanently* delete lists and programs from the calculator, press **2nd** MEM. You will find MEM over the “+” key.

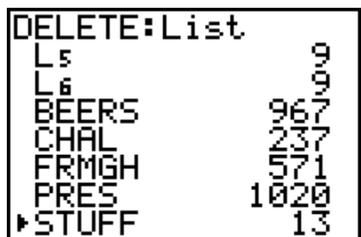


There are several options here. Option **4:ClrAllLists** will do just what the names says it will do - it will “clear” the data from all data lists. The list names will remain in your calculator; however, they will all be empty. Option **5:Reset** will delete everything you have in your calculator and restore your calculator to the original factory settings. Selecting this option, then pressing **ENTER** to execute the command, will place your calculator in the same state it was in the first time you took it out of its factory package and turned it on.

Right now we are interested in option **2: Delete**. Selecting this option brings up additional options. We want to select **4:List**. If we wanted to delete a program, we would have selected **7:Prgm**. Having selected **4:List**, you will see a screen that contains the names of all data lists in your calculator. The screen will look very much like the screen we saw when selecting lists to transmit. Pressing the blue down arrow key will allow us to scroll through the list names. Go down to STUFF.



Pressing **ENTER** at this point will *permanently delete the list STUFF from your calculator*. Once deleted with this command, it is gone forever.



The process for deleting programs is similar to that of lists.



Deleting Lists and Programs From The TI-83/84 Plus

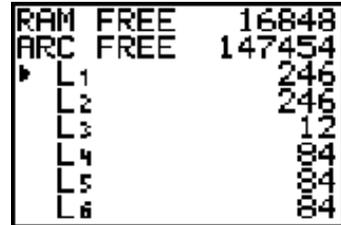
To *permanently* delete lists and programs from the calculator, press **2nd MEM**. You will find MEM over the “+” key.

Selecting option 2, **2: Mem Mgmt/Del** will allow you to permanently delete lists and programs.

After selecting **2: Mem Mgmt/Del**, select option 4, **4: List**, to permanently delete a list. Option 7, **7: Prgm**, will allow you to permanently delete programs.

Having selected **4: List**, a screen will appear showing you the names of all the lists in your calculator.

Use the blue down arrow keys to move the selector to the list STUFF. With the list STUFF selected, as indicated by the black arrow-head pointing to STUFF, press the **DEL** key on your calculator. STUFF will then be removed permanently from your calculator. Please note that pressing **ENTER** does not select the data list for deletion. Rather, pressing **ENTER** will place an asterisk (*) to the left of the list name indicating it is an *Archived* data list. An *Archived* data list is protected such that it can not be inadvertently deleted.



Basic Descriptive Statistics

Suppose you have the simple data set consisting of the values, 3,7, 5, 9, 12, 19, 31, 8 and want to find the basic sample descriptive statistics consisting of the mean, standard deviation, minimum, maximum, median along with the 1st and 3rd quartiles. If the data is stored in L₁, then the command sequence is: **STAT > CALC > 1-Var Stats** **ENTER**. This will place the command **1-Var**

Stats to the homescreen. Next tell the calculator where your data is stored. Since it is in L₁, press **2nd 1** and then **ENTER**. The results will then be displayed as shown below.

EDIT TESTS 1: 1-Var Stats 2: 2-Var Stats 3: Med-Med 4: LinReg(ax+b) 5: QuadReg 6: CubicReg 7: QuartReg	1-Var Stats L1	1-Var Stats $\bar{x}=11.75$ $\Sigma x=94$ $\Sigma x^2=1694$ $Sx=9.176834188$ $\sigma x=8.584142357$ $\downarrow n=8$
--	----------------	--

Notice the arrow pointing down along side of the last line, which reads n=8. This is telling you there is more information available. Press the down blue arrow key to scroll down to the remaining information.

1-Var Stats $\uparrow n=8$ $\text{min}X=3$ $Q1=6$ $\text{Med}=8.5$ $Q3=15.5$ $\text{max}X=31$

Suppose you had the same data stored in a list that you created and called "STUFF." To display the basic sample summary statistics for the list STUFF, start by following the same process. When the command **1-Var Stats** is displayed on the homescreen, press **2nd STAT**, which will display the names of all lists currently stored in your calculator. Use the blue down arrow key until you find the list named **STUFF** and then press **ENTER**. This will paste the proper list name to your homescreen. Press **ENTER** again to display the descriptive statistics.

2nd OPS MATH ↑ LEAF1 : MCGWR : NCOMB : RESID : SOSA : STEM1 : STUFF	1-Var Stats LSTU FF	1-Var Stats $\bar{x}=11.75$ $\Sigma x=94$ $\Sigma x^2=1694$ $Sx=9.176834188$ $\sigma x=8.584142357$ $\downarrow n=8$
---	------------------------	--

Transferring Programs to Your Calculator

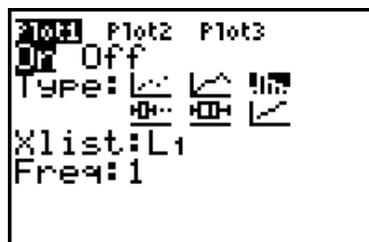
Transferring Programs to Your Calculator

There are a wide range of useful programs available to you at no fee from the book website (<http://www.3ringpublishing.com>).

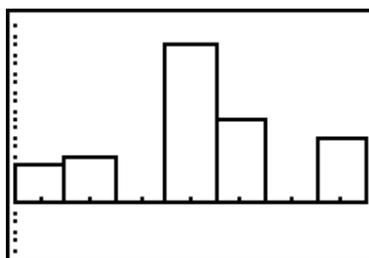
The software needed to transfer the programs from your computer to your calculator is also available at no fee from Texas Instruments. Texas Instruments routinely releases updates to their software, making it impossible to provide up to date instructions in this text. Because of this, we will do our best to provide up to date instructions on the 3RingPublishing.com website.

data in L1; if you put your data in L2 then enter L2). The entry in **Xlist** tells the calculator where you have stored the data.

Press **GRAPH**. Unfortunately, you may or may not see your bar graph. Adjusting the screen to see all of your graph is an easy task.

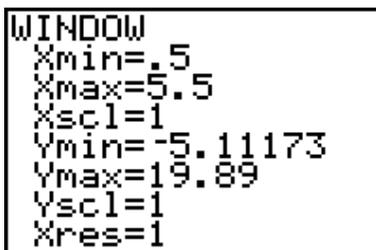


Whenever you do a stat-plot, always select **ZOOM 9**, which automatically adjusts your screen to best fit your stat-plot. Zoom 9 is the Zoom Stat command, meaning it will provide what the calculator believes to be the best picture of your statistical data. The graph you get should look like the graph to the right.



The problem with using the **GRAPH** command is your histogram may or may not be visible on the screen. More often than not, you will need to manually adjust the screen location and dimensions to see your graph when you use the **GRAPH** command. Using **ZOOM 9** will automatically set the window, insuring your graph is visible. Once visible, you may want to make manual adjustments to get a better picture.

Notice there are a total of 7 categories listed here even though we only have 5 categories in our data. The reason is because the **ZOOM 9** command chose the best fit for the screen, which is not necessarily the best display for our data. To make appropriate adjustments, select **WINDOW**. We know our data is coded from 1 to 5, so we will set our **Xmin** to 0.5, **Xmax** to 5.5, and **Xscl** to 1. We will leave the other settings alone.



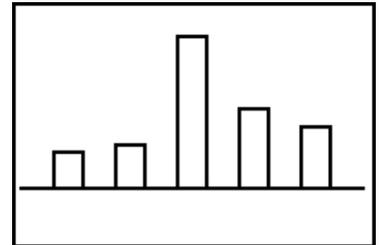
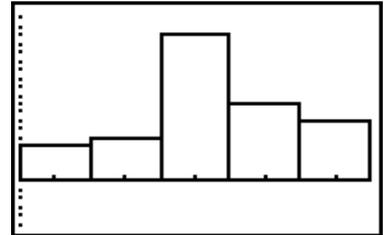
Setting **Xscl** to 1 tells the calculator to make the widths of the bars all equal to one. By setting **Xmin** to 0.5, our first category will start at $x=0.5$ with a width of 1, so the first category will include all values in our data set from 0.5 to 1.5, which will be all of the 1's. The next group will go from 1.5 to 2.5, which will include all of the 2's, and so on. By starting at 0.5 and ending at 5.5, the entire screen will be filled with our graph from left to right.

Using the Trace Command

Traditionally, bar graphs have a space between each bar. This can be accomplished by adjusting the window settings.

There are many possible settings that will produce the desired effect, and one is shown here.

Once the new window settings have been entered, press the **GRAPH** key. Using **Zoom 9** will reset the window to the previous “best fitting” values as defined by the calculator.

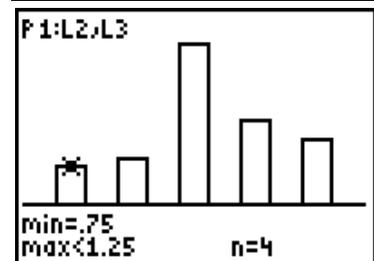
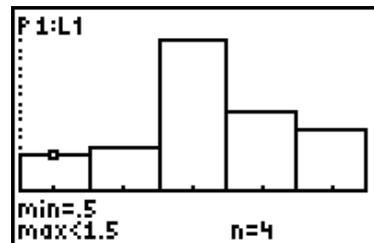


```
WINDOW
Xmin=.25
Xmax=5.75
Xscl=.5
Ymin=-5.11173
Ymax=19.89
Yscl=0
Xres=1
```

Using the Trace Command

The graph is informative, but much more information is available. Press **TRACE** and a cursor will jump to the middle of the top of the first bar. In the bottom of the display you will see the data range for that bar and the frequency ($n=4$). Since the data is integer valued, this is telling us there are 4 occurrences of the value “1” or 4 people who received an “F” based on the data coding we used.

You can jump to the next category by pressing the right arrow key. This is true for either graph, as shown here.



Using Freq

We could have obtained this same picture by using a different technique to enter the data. Rather than entering all 42 values, we could have simply entered 1 through 5 in L1, and then the frequencies in L2. Remember, we coded an “F” as a 1.

L1	L2	L3	3
1	1	██████	
2	2		
3	21		
4	6		
5	1		
L3(1)=			

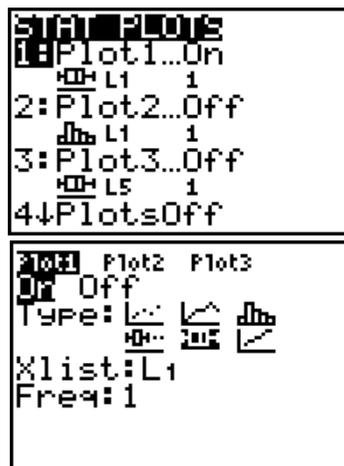
Select **STAT PLOT** and then **PLOT 1** just as we did before. The bottom entry lists **Freq:1**. Change the 1 to L2. This is telling the calculator the data is in L1, and the frequency counts are in L2. The graph that you get will be identical to the previous graph.

Box Plots

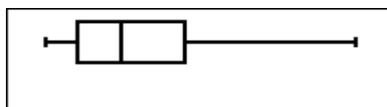
Consider the following data set.

3	7	5	12	8
2	2	21	6	5
4	11	10	9	1

A box plot is a picture of the “five-number-summary.” It contains the minimum value, maximum value, first and third quartiles, along with the median. To create a box plot, go to **STAT PLOT** and select **Stat Plot 1** (or other Stat Plot of your choice). Insure Plot 1 is turned on. Next go to **Type** and select the desired box plot icon. Notice there are two different icons for box plots. The first is a picture of a box plot with two dots on the right side. The second is a picture of a box plot without the dots. The first option will identify *outliers* with dots outside the box plot. The second will not identify outliers and will simply extend the “whiskers” out to the minimum and maximum values in the data set. An outlier on the lower end of the box plot is defined as any data value less than $Q1 - 1.5(\text{Inner Quartile Range})$ which is $Q1 - 1.5(Q3-Q1)$. Likewise, an outlier on the upper end is defined as any data value greater than $Q3 + 1.5(Q3-Q1)$.



Let's first select the box plot that does not identify outliers. The data is in L1 (or wherever you happen to have put it). Select **GRAPH** followed by the *Zoom-Stat* command as previously discussed. You should get the following picture.

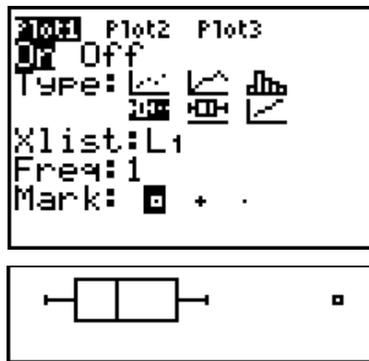
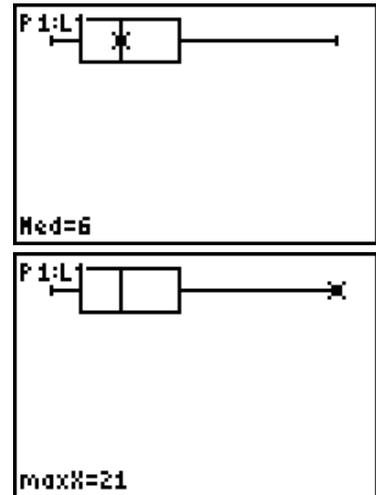


Just as with the histogram, we can use the trace command to identify data values.

Scatter Plots

Here, the trace command is identifying the median as being 6. Pressing the right and/or left arrow keys will identify the minimum value, maximum value, along with Q1 and Q3.

Now consider a box plot of the same data using the box plot option that identifies outliers. To do this, go back to Statplot-1 and select the appropriate icon. Notice when this icon is selected, an additional option called **Mark** is displayed at the bottom of the screen. You can select the type of *mark* you want the calculator to use in identifying an outlier. You can select a small box, a “+” sign, or a dot. Once this is done, press **GRAPH**. You may or may not need to use zoom-stat again.



Note the small box to the right of the box plot. This is indicating there is one outlier based on how we previously defined an outlier. We can still use the trace command to identify values.

Multiple box plots can be displayed by turning on more than one StatPlot and setting each to display a box plot of data stored in different data lists. An example is shown in the Analysis of Variance Chapter.

Scatter Plots

Another graphical display available is that of a scatter plot. Scatter plots are used for bivariate data. Consider the following bivariate data consisting of the number of hours students studied for an exam and the score received on the exam (score is out of 100 possible points).

Hours	.5	6	2.75	3	4.25	1
Score	43	97	72	81	88	52

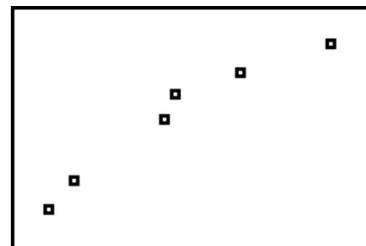
We will enter the Hours Studied in L1 and the Scores in L2.

L1	L2	L3	1
5	43	-----	
6	97		
2.75	72		
3	81		
4.25	88		
1	52		
-----	-----		
L1(1) = .5			

To obtain a scatter plot, return to **STAT PLOT** and select StatPlot-1. For all of the examples thus far, we have been returning to StatPlot-1. Keep in mind you could easily do these in StatPlot-2 or StatPlot-3. The key is to remember to turn the other stat plots off. The first plot type is the icon for the scatter plot. Select this option and fill in the rest of the required information. We will put the hours studied (L1) on the x-axis and score (L2) on the y-axis. Once this is completed, use the **ZOOM 9** command once again to view the plot.

```

ZOOM Plot2 Plot3
Off Off
Type: [Scatter] [Line] [Bar]
      [MH] [MH] [MH]
Xlist:L1
Ylist:L2
Mark: [ ] + .
    
```



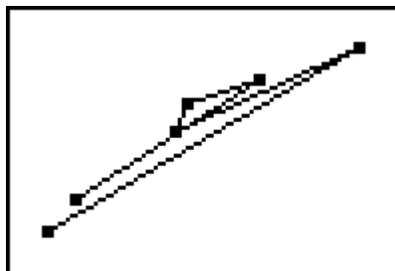
As usual, we can use the trace command to identify the coordinates of any of the data points. We can move through the data points displayed on the screen using the right and left arrow keys. By pressing the right arrow key, the cursor will jump to the next data point *in the order they appear in the data lists*. The order they appear in the data list is seldom the same order they appear in the scatter pot, hence your cursor may “jump” around from data point to data point.

```

P1:L1,L2
X=5 Y=43
    
```

xyLine Plot

The next plot option available is referred to as an xyLine plot. This is the same as a scatter plot except a line is drawn connecting the points. Lines are drawn, connecting the points, in the order you entered the data in the editor. Using the above data and selecting an xyLine plot, you would get the following picture. The **xyLine Plot** is of little use with the example data used here; however, there are circumstances where the **xyLine Plot** is quite useful.



Normal Plots

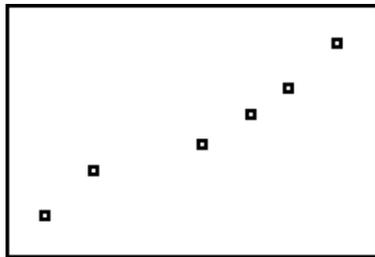
Normal Plots

The last plot available under the stat plot menu is the *Normal Probability Plot*. The normal probability plot is a plot used to assess whether or not a particular set of data is likely to have come from a normally distributed population. This plot is used for univariate data. We will not get into how the calculator creates this particular plot. Rather, it is more important that we know how to tell the calculator to produce the plot and be able to interpret the results.

The normal plot is the last icon under **Type**. Select the normal plot and tell your calculator to examine the data in L2, the scores from the exam. The **Data Axis** defaults to X; leave it that way.



Use **ZOOM 9** to view your normal plot.



Deciding if the data is normally distributed or not strictly from a plot is more of an art than an exact science. In general, the closer the normal plot is to a 45 degree line, the more likely the data follows a normal distribution. There are tests for normality available that produce p-value, but none of those tests are built into the TI-83/84 calculator. As such, we must rely on a “visual inspection” of the plot to base our decision. In the Programs chapter, we will look at a program that produces a normal plot with a reference line. The reference line is useful when making a decision regarding normality based on visual inspection of a normal plot.

The commands to generate random data are all located under the MATH menu. Go to **MATH > PRB**. It is from here you will access the commands to generate random data.

```
MATH NUM CPX PRB
1:rand
2:nPr
3:nCr
4:!
5:randInt(
6:randNorm(
7:randBin(
```

Generating Integer Valued Random Data

From the **Math > PRB** menu, select option **5:randInt** by pressing the “5” key or using the down arrow until that option is highlighted and then pressing **ENTER**. This will paste the command to your home screen as shown here.

```
randInt(
```

The syntax for generating integer random data is:

randInt(lower bound, upper bound, number of values desired)

We can optionally store the data in a particular list, which is wise in most situations.

Suppose we wanted to generate 25 random integers from 1 to 10 and store those values in L1. The command is shown here.

```
randInt(1,10,25)
→L1
```

The arrow is obtained by selecting the **STO** key which stands for “store.” Pressing **ENTER** at this point will generate the random data and store it in L1, which can then be viewed in the editor.

Generating Random Data from a Normal Distribution

Notice option 6 under the **MATH > PRB** menu. This is the option used for generating random data from a normal distribution.

The syntax for generating normal random data is simply:

randNorm(μ , σ , number of desired values)

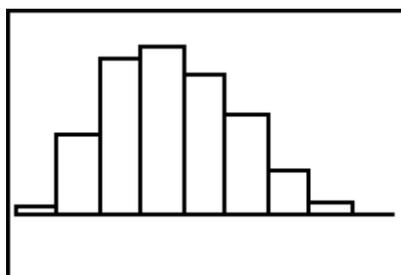
```
MATH NUM CPX PRB
1:rand
2:nPr
3:nCr
4:!
5:randInt(
6:randNorm(
7:randBin(
```

Generating Random Numbers / Simulations

Suppose we wanted to generate 150 random numbers from the standard normal distribution (mean of zero and standard deviation of 1) and produce a histogram of the data. A logical place to store the data would be L1, although any available list can be used.

```
randNorm(0,1,150)
)→L1
```

Press **ENTER**, and the data will be generated and stored in L1. The resulting histogram from the random data that we just generated looks like the graph to the right. Remember, this is *random* data, so your histogram may not look exactly like this histogram. Default values were used to display this histogram (i.e. ZoomStat was used).



If we wanted to generate data from a normal distribution with a mean of 100 and a standard deviation of 16 and place the data in L2, the command syntax would be:

```
randNorm(100,16,
150)→L2
```

Generating Random Data from a Binomial Distribution

The last option under **MATH > PRB** is option 7. This option will generate random data from a binomial distribution. The syntax is:

randBin(number of trials, probability of success, number of simulations)

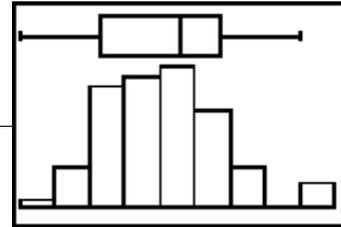
Suppose we wanted to generate 150 values from a binomial distribution based on 10 trials, a probability of success of 0.35 for each trial, and store the data in L3. The command would be:

```
randBin(10,.35,1
50)→L3
```

The graphic to the right shows the random data in L1 and L2 generated from a Normal Distribution and the random data in L3 generated from a Binomial Distribution. Remember, the data is **random**. This means your values will most certainly be different from the values shown here.

L1	L2	PRB	3
-1.043	.45323	4	
-.0322	-.8819	4	
-1.507	-.0939	6	
.72031	.5	4	
-.155	-.4687	5	
-.8628	1.8159	5	
-1.289	-.8463	1	
L3 = {4,4,6,4,5,5...			

A histogram of the binomial data with the associated box plot is shown here.



Monte Carlo Method of Simulation

Instead of studying an actual situation, which can be extremely expensive, time intensive, or dangerous, researchers often create a similar situation that is not dangerous, less time intensive, and less costly. As an example, airline companies use flight simulators to train their pilots rather than taking them up 30,000 feet and actually starting an engine on fire.

The idea of simulation has been around for a long time. They go back to ancient times when chess was used to simulate warfare. Mathematical simulation techniques use random numbers and probability to mimic real life situations.

The Monte Carlo method is a common technique used for simulation and is based on random numbers. The steps for a basic Monte Carlo are:

- List all possible outcomes of your experiment.
- Determine the probability of each outcome.
- Set up a correspondence between each outcome and random numbers.
- Generate random numbers.
- Repeat the process n times.
- Compute statistics of interest and outline your conclusions.

Monte Carlo Simulation Example

Craps is a game of chance where players roll two dice and bet on the outcome. The term “rolling a 6” then refers to rolling the two dice simultaneously such that the two values that show face up sum to 6. Using simulation, set up a method of determining the probabilities of rolling a 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12 while playing craps (this actually enumerates all possible outcomes).

- **List all possible outcomes of your experiment.** *The possible outcomes are 1, 2, 3, 4, 5, or 6 on each of two die. They combine to produce the possible results of 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12.*
- **Determine the probability of each outcome.** *Assuming the die are fair, then the probability of observing a 1, 2, 3, 4, 5, or 6 are all $1/6$.*
- **Set up a correspondence between each outcome and random numbers.** *The correspondence here is simplistic in that we can simply assign random numbers 1 through 6 to outcomes 1 through 6.*
- **Generate random numbers.** *We will use the TI-83/84 to generate random integer valued data from 1 to 6 and store them in L1 and L2. Each list will represent the*

result of one die. By summing L1 and L2, we will have a simulated result of rolling two dice once each as in playing the game of craps.

- **Repeat the process n times.** The process can be repeated by generating more than 1 random number for L1 and L2. If we generate a total of 250 random numbers into each list, then we have simulated tossing the dice 250 times.
- **Compute statistics of interest and outline your conclusions.** The statistic of interest will be the proportion of times each of the values 2 through 12 show up in our simulation. We can obtain this by summing the two columns by row and place the results in L3. A histogram of the results should provide us with the information desired.

Let's start with generating 250 random integers from 1 to 6 and placing them in L1. Once this is done, we can generate another 250 random integers from 1 to 6 and place them in L2.

```
randInt(1,6,250)
→L1
```

2nd ENTER will bring up your last command. Do so and replace L1 with L2. This will act as a short cut so you will not have to go back to MATH > PRB to select the command again and fill in all the needed parameters.

```
randInt(1,6,250)
→L1
(1 1 1 3 4 5 6 ...
randInt(1,6,250)
→L2
```

Now go to the editor and place the sum of L1 and L2 in L3.

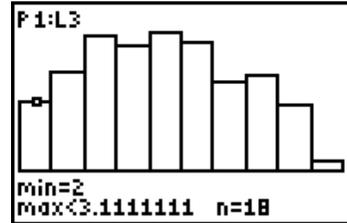
L1	L2	3	L1	L2	L3	3
1	4	-----	1	4	5	
1	3		1	3	4	
1	2		1	2	3	
3	2		3	2	5	
4	1		4	1	5	
5	1		5	1	6	
6	1		6	1	7	
L3 = L1 + L2			L3(1)=5			

To place the sum of L1 and L2 in L3, be certain to go to L3 and press the up arrow until L3 is highlighted as shown above. A typical mistake is to attempt to add L1 and L2 but put it in the first cell of L3. If you make this mistake, you will get an error message.

Now, obtain a histogram of L3. Once the histogram is obtained, it would be wise to adjust the default values such that each group contains only one value, such as all 2's, all 3's and so on.

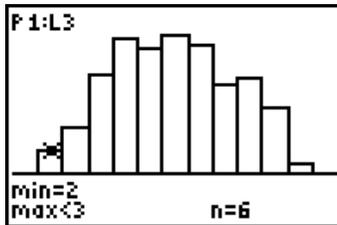
Monte Carlo Method of Simulation

This histogram is based on the default values. Notice the first group contains both the 2's and the 3's. The histogram, in this form, is of no use.



By making the following adjustments to the window, each bar in the histogram will represent the frequency of each possible result (i.e. 2's, 3's, 4's, ...).

```
WINDOW
Xmin=1
Xmax=14
Xscl=1
Ymin=-10.82484
Ymax=42.12
Yscl=0
Xres=1
```



Now, the first group contains only 2's, the second group only 3's, and so on.

Calculating binomial probabilities with the TI-83/84 is really quite simple. The trickiest part is understanding how to transform the question being asked to a logical mathematical expression. Once the expression is determined, the needed TI-83/84 commands follow naturally. You will find the binomial distribution commands under **2nd > DISTR**.

<pre> 0: 2nd DRAW 1: normalpdf(2: normalcdf(3: invNorm(4: tpdf(5: tcdf(6: X²pdf(7: X²cdf(</pre>	<pre> 0: 2nd DRAW 9: Fcdf(A: binompdf(B: binomcdf(C: Poissonpdf(D: Poissoncdf(E: geometpdf(F: geometcdf(</pre>
--	---

The binomial distribution commands are options **0:** and **A:**. The two commands are **binompdf** and **binomcdf**.

The Binomial Probability Distribution

The **pdf** ending stands for *probability distribution function*. This command will find the value of the probability distribution function, which is defined as:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

This is the probability of observing exactly x successes in n trials given the probability for success is p . The syntax for the TI-83/84 command is: **binompdf**(n , p , x).

Example: Suppose a basketball player has been hitting 73% of her field goals all season. What is the probability that the player hits 8 of the next 10 attempts?

We need to identify the values of n , p , and x . In this example, we have $n = 10$, $p = 0.73$, $x = 8$. The TI-83/84 command is:

<pre> 0: 2nd DRAW 9: Fcdf(A: binompdf(B: binomcdf(C: Poissonpdf(D: Poissoncdf(E: geometpdf(F: geometcdf(</pre>	<pre>binompdf(10,.73, 8)</pre>	<pre>binompdf(10,.73, 8) .2645592331</pre>
---	--------------------------------	--

So there is an approximate 26.45% chance that the player will hit 8 of her next 10 field goal attempts.

The equivalent hand computations are:

$$\begin{aligned}P(X = x) &= \binom{n}{x} p^x (1-p)^{n-x} \\P(X = 8) &= \binom{10}{8} (0.73)^8 (0.27)^{10-8} \\&= \left(\frac{10!}{8!(10-8)!} \right) (0.73)^8 (0.27)^2 \\&= \left(\frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 (2 \cdot 1)} \right) (0.73)^8 (0.27)^2 \\&= \left(\frac{10 \cdot 9}{2} \right) (0.73)^8 (0.27)^2 \\&= 45(0.73)^8 (0.27)^2 = 0.2645\end{aligned}$$

The Binomial Cumulative Probability Distribution

The **cdf** at the end of the other binomial probability option stands for **cumulative distribution function**. If we asked the question “What is the probability she will hit no more than 8 of her next 10 attempts,” then we are asking the probability she will hit 0, or 1, or 2, or 3, or 4, or 5, or 6, or 7, or 8. We would represent this as:

$$\begin{aligned}P(X \leq x) \text{ or } P(X \leq 8) \\&= \sum_{i=0}^8 \binom{10}{i} p^i (1-p)^{10-i} \\&= 1 - \sum_{i=9}^{10} \binom{10}{i} p^i (1-p)^{10-i}\end{aligned}$$

The calculations by hand can be time consuming. The TI-83/84 syntax for the cumulative probability command is the same as the probability distribution command, except that **binomcdf** is used rather than **binompdf**.

$$\mathbf{binomcdf}(n, p, x)$$

<pre> 0:QUIT DRAW 1:Fcdf(2:binompdf(3:binomcdf(4:poissonpdf(5:poissoncdf(6:geometpdf(7:geometcdf(</pre>	<pre> binomcdf(10,.73, 8) </pre>
--	----------------------------------

Now press **ENTER**, and we get the probability of observing *no more than* 8 successes in 10 trials. The *no more than* statement is also often worded as “8 or less” or “8 or fewer” successes. You must pay close attention to the phrasing. If a problem was phrased as “fewer than 8 successes,” then the problem is actually asking for the probability of observing 7 or fewer successes. If this seems confusing, then look closely at the wording.

```

binomcdf(10,.73,
8)
.7980704576

```

Recall we said this same problem can be represented with appropriate notation as:

$$\begin{aligned}
 &P(X \leq x) \text{ or } P(X \leq 8) \\
 &= 1 - \sum_{i=9}^{10} \binom{10}{i} p^i (1-p)^{10-i}
 \end{aligned}$$

In English, this mathematical statement is saying the probability of observing 8 or fewer successes in 10 trials is equal to 1 minus the probability of observing at least 9 success in 10 trials (this is using the idea of compliments).

The TI-83/84 can find the probability of observing at least 9 successes in 10 trials, but it needs to do so by finding the probability of 8 or fewer and then subtracting it from 1.

```

1-binomcdf(10,.7
3,8)

```

Now press **ENTER**, and we get the probability of observing *at least* 9 successes in 10 trials.

```
1-binomcdf(10,.7
3,8)
      .2019295424
```

Notice if we add the probability of observing no more than 8 successes to the probability of observing at least 9 successes, the total is 1.

Now suppose we were interested in finding the probability she will hit between 6 and 8 (inclusive) of her next 10 free throws. This is telling you to find:

$$\begin{aligned} &P(6 \leq X \leq 8) \\ &= \sum_{i=6}^8 \binom{10}{i} p^i (1-p)^{10-i} \\ &= P(X=6) + P(X=7) + P(X=8) \end{aligned}$$

Once again, calculating this by hand can be time consuming. The TI-83/84 can find the desired probability by using the cumulative probability distribution command. First we will find the probability of observing 8 or fewer successes, and then subtract from it the probability of observing 5 or fewer successes. What remains is the probability of 6 successes, 7 successes, and 8 successes.

```
binomcdf(10,.73,
8)-binomcdf(10,.
73,5)
```

Now press **ENTER**, and we get the probability of observing 6 to 8 successes (inclusive) in 10 trials.

```
binomcdf(10,.73,
8)-binomcdf(10,.
73,5)
      .6943873136
```

The TI-83/84 does a nice job of calculating probabilities based on the normal distribution. You can transform a non-standard normal distribution to a standard normal or calculate the needed probabilities directly with the parameters of the non-standard normal distribution. The biggest advantage to using a standard normal distribution is the fact that a *z-score* represents the number of standard deviations a particular observation resides above or below the mean. This information can be useful to someone with just a little experience. The draw back is the need to physically calculate the value of z . The calculation is not difficult, but it can become tedious if you have a need to transform several values.

Probabilities Based on the Standard Normal

There are typically three different situations in which we may be interested in finding probabilities. Those situations are:

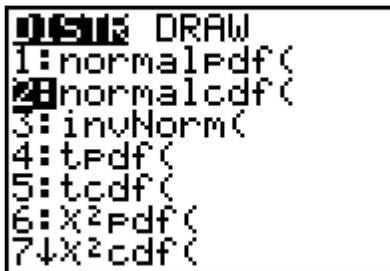
1. $P(Z \leq z)$ where Z represents the random variable and z represents the value of interest. As an example, if we were interested in the probability a standard normal random variable took on a value of less than or equal to 1.2, then we would write it as $P(Z \leq 1.2)$.

The general command format for the TI-83/84 is:

normalcdf(lower bound, upper bound, mean, sd)

The **normalcdf** command can be found by pressing **2nd DISTR**. Option **2** is the option we want.

Once the command is selected, it will be pasted on the home screen where you can provide the remaining parameters. The nice thing about the standard normal distribution on the TI-83/84 is if the mean and standard deviations are omitted, the calculator understands a standard normal distribution is being used.



If we were looking for $P(Z \leq 1.2)$, then we want the cumulative probability from $-\infty$ to 1.2. The TI-83/84 does not have a $-\infty$ option as such. Rather, what we do is select an extremely large number (positive or negative) to simulate positive or negative infinity. The proper key strokes on the TI-83/84 are **2nd EE 99**. It will show up on your screen as: **-E99**.

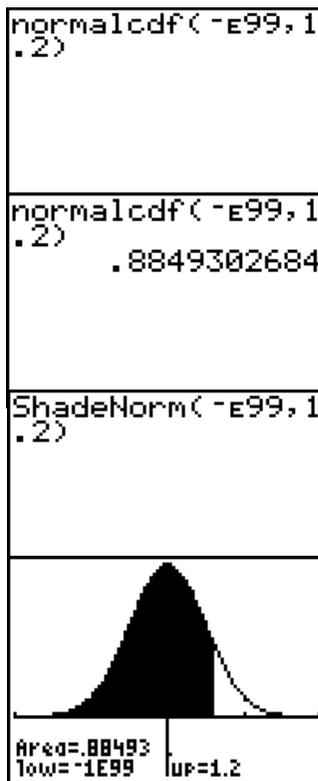
The quantity **EE99** is not obtained by typing in two E's and then 99. Rather, **EE** is located over the “,” button. The quantity is obtained by **2nd** , . This is the way the TI-83/84 represents scientific notation. The value is actually -10^{99} which is a HUGE negative number. The complete command in the TI-83/84 window will look like this.

Notice that I did not enter the mean and standard deviation. When you press **ENTER**, the TI-83/84 will display the desired probability. When the mean and standard deviation are not entered, the calculator assumes you are working with a standard normal distribution and uses a mean of 0 and a standard deviation of 1. If I were working with a non-standard normal, then the mean and standard deviation would need to be entered.

If I wanted a picture, I could select **ShadeNorm**(under the **DISTR** > **DRAW** menu. The parameters are supplied in the same fashion.

The difference is that this time when we press **ENTER**, we get a picture along with the probability. Notice that the upper and lower bounds are also displayed.

The biggest inconvenience when using the **ShadeNorm** command is the window is not automatically adjusted to provide the best possible picture. In fact, you may not see a picture at all. You must go in and adjust the window parameters in order to see the graph. The previous picture was obtained with $X_{min} = -4$, $X_{max} = 4$, $Y_{min} = -0.25$, $Y_{max} = 0.4$. There is a program called **shadnorm** that will be discussed in the *Programs* chapter that will handle the window display settings for you.



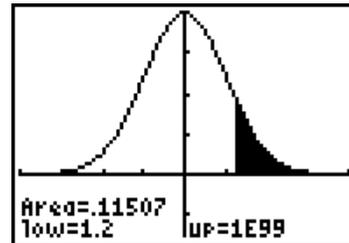
2. $P(Z \geq z)$. As an example, if we were interested in the probability a standard normal random variable took on a value greater than or equal to 1.2, then we would write it as $P(Z \geq 1.2)$. The TI-83/84 commands are almost identical to the previous example. Rather than going from $-\infty$ to 1.2, we are going from 1.2 to ∞ . The first example looked at the left tail of the distribution; this example looks at the right tail. The TI-83/84 command is shown here. Since we are going to ∞ , the lower bound became 1.2 and the upper bound became **E99**.

```
normalcdf(1.2, E99
9)
```

```
normalcdf(1.2, E99
9)
.1150697316
```

Notice that if you were to add up the probabilities from both of these examples, they will sum to 1. This should not be surprising since between the two examples, we have accounted for all of the area under the curve, which must be 1.

We can produce a picture in a similar manner as example #1.



3. $P(z_1 \leq Z \leq z_2)$. As an example, if we were interested in the probability a standard normal random variable took on a value between 1.2 and 2.3, then we would write it as $P(1.2 \leq Z \leq 2.3)$.

As you may have guessed, the TI-83/84 command is simply:

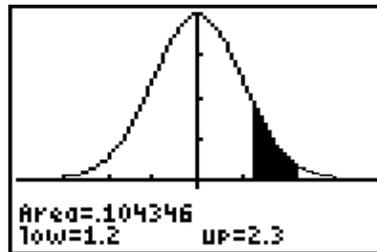
```
normalcdf(1.2,2.  
3)
```

Pressing **ENTER** gives us:

```
normalcdf(1.2,2.  
3)  
    .1043456505
```

Just as we did previously, we can obtain a picture.

```
ShadeNorm(1.2,2.  
3)
```



Non-standard Normal Probabilities

Non-standard Normal Probabilities

Finding probabilities for non-standard normal distributions is nearly identical to standard normal distributions. The only exception is that we must specify the mean and standard deviation.

Suppose we wanted to find the probability the value of a random variable from a normal distribution, that has a mean of 5.1 and a standard deviation of 2.6, falls on the interval 5.3 to 6.9. To do this we would use the following TI-83/84 command.

Pressing **ENTER** produces the result of 22.49%.

```
normalcdf(5.3,6.9,5.1,2.6)
```

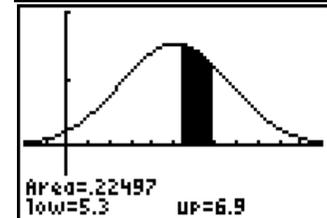
```
normalcdf(5.3,6.9,5.1,2.6)
.2249703137
```

Likewise, the command to produce a picture of the situation is:

```
ShadeNorm(5.3,6.9,5.1,2.6)
```

Which results in this picture:

The vertical bar on the left side of the graph is the y-axis.



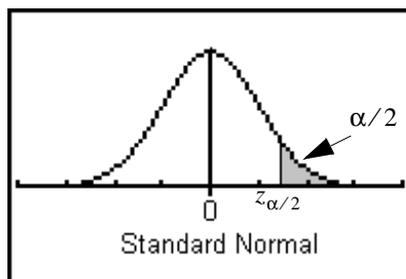
The window dimensions used to make the graph visible are shown here.

```
WINDOW
Xmin=-2
Xmax=12
Xscl=1
Ymin=-.1
Ymax=.2
Yscl=.1
Xres=1
```


Estimation of the mean when the standard deviation is known

To construct a $(1 - \alpha)100\%$ confidence interval (CI) for μ with known σ .

- Collect your data and calculate the sample mean, and call it \bar{x} .
- Determine α , the reasonable risk for a Type I error.
- Find $z_{(\alpha/2)}$ where $z_{(\alpha/2)}$ is the value of z such that the area under the curve in the right tail is equal to $\alpha/2$. If α is 0.05, then $\alpha/2 = 0.025$.
- The $(1 - \alpha)100\%$ CI for μ is then $\bar{x} \pm z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}$

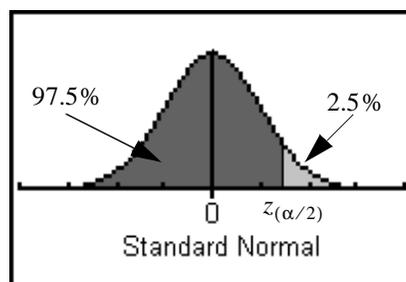


Example: Suppose you want to know how far the average student commutes (one way) to school. A random sample of 100 commuting students was identified. The mean distance was 8.7 miles (assume $\sigma = 5$). Find a 95% CI for μ , the true population mean.

Solution: The form of the confidence interval is

$$\bar{x} \pm z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}$$

confidence interval, $z_{(\alpha/2)}$ is the value of z that gives us 2.5% in the right tail. On the TI-83/84, we can use the **invNorm** (Inverse Normal) command. Keep in mind that the value of z that gives us 2.5% in the right tail is the same value that gives us 97.5% in the left tail.



The **invNorm** command will give you the value of z such that the specified area under the curve is to the left of z . In other words, the **invNorm** command gives you the value of z that has the desired area in the left tail.

Just as with the other commands involving the normal distribution, if we do not specify a mean and standard deviation, the calculator assumes we are working with the standard normal. The command will be **invNorm(.975)**. This will give us the value of z that provides 97.5% of the area under the curve to the left of z , which means 2.5% is to the right of z .

From the **DISTR** menu, select **3:invNorm**.

The following command is then pasted to your home screen:

```
invNorm(■
```

```
DISTR DRAW
1:normalPdf(
2:normalcdf(
3:invNorm(
4:tpdf(
5:tcdf(
6:X²Pdf(
7↓X²cdf(
```

Provide the amount of area we need in the left tail (.975) and press **ENTER**.

```
invNorm(.975)      invNorm(.975)
                    1.959963986
```

Thus the value of z that we need for our confidence interval is approximately 1.96. This results in the following confidence interval calculation.

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 8.7 \pm 1.96 \frac{5}{\sqrt{100}} = 8.7 \pm .98$$

(7.72 , 9.68)

If needed, the **invNorm** command can be used with a non-standard normal distribution simply by inputting the desired mean and standard deviation. The basic syntax is:

invNorm(area, mean, standard deviation)

When the mean and standard deviation are omitted, it is assumed that you are working with a standard normal distribution and the calculator automatically uses 0 for the mean and 1 for the standard deviation. If you wanted to know what value from a normal distribution with a mean of 13 and a standard deviation of 7 would give you 97.5%, the command is simply: **invNorm(.975, 13, 7)** as shown here. The answer, in this case, would be approximately 26.72.

```
invNorm(.975,13,
7)
26.7197479
```

Hypotheses test of the mean when the standard deviation is known

Find a 99% CI for μ , the true population mean with this same information. You can work out the details. The actual confidence interval is (7.41, 9.99). Here's a TI-83/84 screen to help you through it. The confidence interval can be constructed directly on the TI-83/84 without the need to first find the appropriate value of z using **invNorm**. Go to **Stat > TESTS > 7: ZInterval** and fill in the needed information.

```
invNorm(.995)
2.575829303
```

<pre>EDIT CALC TESTS 1:Z-Test... 2:T-Test... 3:2-SampZTest... 4:2-SampTTest... 5:1-PropZTest... 6:2-PropZTest... 7:ZInterval...</pre>	<pre>ZInterval Inpt:Data STATS σ:5 \bar{x}:8.7 n:100 C-Level: .95 Calculate</pre>	<pre>ZInterval (7.72,9.68) \bar{x}=8.7 n=100</pre>
---	---	---

The 95% confidence interval is 7.72 to 9.68, just as we calculated by hand using the TI-83/84 for only the value of $z_{(\alpha/2)}$. Notice we selected **Stats**. This is the option we use when we are given the mean and standard deviation. If we have data to work with, then we would have selected **Data** and filled in the needed information.

Hypotheses test of the mean when the standard deviation is known

Typically we have a value which we believe may represent the population mean and we want to determine if the population mean may be greater than, less than, or simply different from the value we are considering.

If we are interested in $<$ then we have a One Tailed Test (left tail)

If we are interested in $>$ then we have a One Tailed Test (right tail)

If we are interested in \neq , then we have a Two Tailed Test

EPA Example

Suppose the EPA is suing a city for noncompliance of carbon monoxide standards. The EPA wants to show the mean level is greater than 4.9 parts per million, the allowed level. Suppose we take 30 samples at different locations and find the sample mean to be 5.3 parts per million. Further suppose the population standard deviation is known to be 2.1 parts per million. The alternative hypothesis is that the average carbon monoxide level is greater than 4.9 parts per million. We can find the p -value on the TI-83/84 with two different methods.

$$H_0: \mu = 4.9 \quad H_a: \mu > 4.9$$

Method #1: The first method would be to convert the observed mean to a z-score by hand and use the `normalcdf` command found under **Math > PROB**. The z-score is 1.04.

```
normalcdf(1.04, E
99)
.1491699711
```

Method #2: The second method is to go to **STAT > TESTS > Z-Test**.

<pre>EDIT CALC TESTS 1:Z-Test... 2:T-Test... 3:2-SampZTest... 4:2-SampTTest... 5:1-PropZTest... 6:2-PropZTest... 7:ZInterval...</pre>	<pre>Z-Test Inpt:Data [DATA] μ₀:4.9 σ:2.1 x̄:5.3 n:30 μ:≠μ₀ <μ₀ >μ₀ Calculate Draw</pre>
---	--

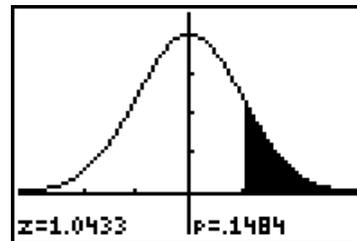
Once you select **Z-test**, you will be presented with a screen that will ask you for the necessary information. The first option is: **Data Stats**. If the information was contained in a list (you were working with actual data), then you would select **DATA**. In this example we are given the summary statistics without the data, so we will select **Stats**. We then simply fill in the rest of the information.

Since our alternative hypothesis is greater than (>), we select the > option on the second line from the bottom. The last line gives two options. The first is **Calculate**. If you select **Calculate**, then you will get a p-value. If you select **Draw**, you will get a picture along with a p-value.

Output from CALCULATE command.

```
Z-Test
μ>4.9
z=1.043281062
P=.1484090917
x̄=5.3
n=30
```

Output from Draw command.



Inferences about the mean when the standard deviation is unknown
(Student - t)

Hospital Example

A random sample of 29 babies was taken of the weights of babies at a hospital in 1995. A mean of 6.53 pounds and a standard deviation of 1.76 pounds was calculated from the sample data. Esti-

Inferences about the mean when the standard deviation is unknown (Student - t)

mate, with 95% confidence, the mean weight of all babies born in 1995 at this particular hospital. Assume the distribution of baby weights is normal.

Since we do not have the value of σ and we do have the value of s , we will use a *t-distribution* for our confidence interval.

On your TI-83/84, select **STAT > TESTS**.

Now press the down arrow until the number 8, **TInterval** is selected or simply press **8**, which will automatically select **TInterval**.

```
EDIT CALC TESTS
1:Z-Test...
2:T-Test...
3:2-SampZTest...
4:2-SampTTest...
5:1-PropZTest...
6:2-PropZTest...
7:ZInterval...
8:TInterval...
```

```
EDIT CALC TESTS
2:T-Test...
3:2-SampZTest...
4:2-SampTTest...
5:1-PropZTest...
6:2-PropZTest...
7:ZInterval...
8:TInterval...
```

Selecting **TInterval** will give you the following dialog box. All you need to do is fill in the needed information, and the calculator will produce the appropriate confidence interval.

The first option is **Inpt:** which stands for input. If you have data in a list, then select **Data**. In this example we have the statistics, so select **Stats**.

Fill in the sample mean, the sample standard deviation, the sample size and the **C-Level**, which is the level of confidence we want for our interval. Notice that the confidence interval is entered in decimal format as .95, not as 95%.

```
TInterval
Inpt:Data STATS
x:6.53
Sx:1.76
n:29
C-Level:.95
Calculate
```

Select **Calculate** and then press **ENTER**, and you will get the following display.

```
TInterval
(5.8605,7.1995)
x=6.53
Sx=1.76
n=29
```

So our confidence interval is (5.86, 7.20).

Hypotheses test of the mean when the standard deviation is unknown

Consider the same situation, but someone else believes the true mean weight of the babies to be 5.02 pounds. This information would most likely have come from a previous study or some other form of historical data. We will complete a hypotheses test as we did using z ; however, we will make the appropriate substitutions and base our decision on the t -distribution rather than the z -distribution.

$$H_0: \mu = 5.02$$

$$H_A: \mu > 5.02$$

The actual test statistic has the form:

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}$$

The process is very similar to what we did using the z -distribution. From **STAT > TESTS**, select option **2:T-Test**.

```

EDIT CALC TESTS
1:Z-Test...
2:T-Test...
3:2-SampZTest...
4:2-SampTTest...
5:1-PropZTest...
6:2-PropZTest...
7↓ZInterval...
    
```

```

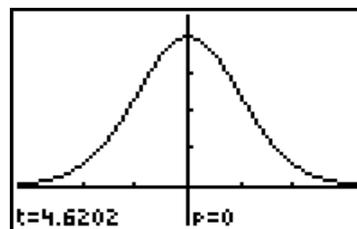
T-Test
Inpt:Data Stats
μ₀:5.02
x̄:6.53
Sx:1.76
n:29
μ:≠μ₀ <μ₀ >μ₀
Calculate Draw
    
```

Output from CALCULATE command.

Output from Draw command.

```

T-Test
μ>5.02
t=4.620226624
P=3.914582E-5
x̄=6.53
Sx=1.76
n=29
    
```



The p-value is reported as 3.914582E-5, which is 0.00003914582 represented in scientific notation, or approximately zero as reported in the Draw graphic.

Return to the EPA Problem

Suppose we had the following information in the EPA problem:

$$\bar{x} = 5.1 \quad s = 1.71 \quad n = 22$$

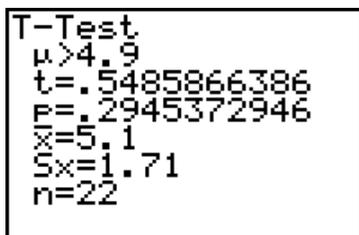
Recall the EPA wants to show the average carbon monoxide level is greater than 4.9 parts per million.

$$H_o: \mu = 4.9$$

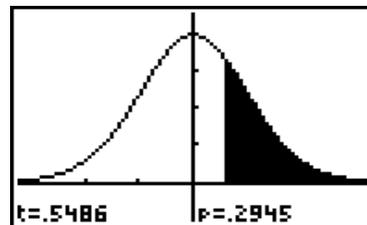
$$H_A: \mu > 4.9$$



Output from CALCULATE command.



Output from Draw command.



The p-value is reported as 0.2945.

Inferences Regarding Proportions

Let the Greek letter π refer to the population parameter for the probability of success on any given trial for a binomial distribution. The observed probability of success is defined as

$$p = \frac{x}{n}$$

where x is the observed number of successes in n trials.

A confidence interval for the true probability of success (proportion of successes) is defined

as: $p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{pq}{n}}$ where $q = 1-p$.

Campus Safety Example

A survey was conducted to estimate the proportion of students who did not feel safe on campus. A total of 76 students, from a sample of 267, said they did not feel safe on campus. Estimate the true proportion of students who do not feel safe on campus with a 95% confidence interval.

The observed proportion is $\frac{76}{267} = 0.2846$, so the confidence interval, calculated by hand, is:

$$\begin{aligned} &0.2846 \pm 1.96 \sqrt{\frac{(0.2846)(0.7154)}{267}} \\ &= 0.2846 \pm 1.96(0.0276) \\ &= (0.2846 \pm 0.0541) \\ &\quad (0.2304, 0.3387) \end{aligned}$$

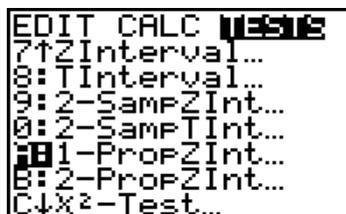
The TI-83/84 handles confidence intervals for proportions quite nicely. The confidence interval is located under **STAT > TESTS** and is listed as option **A: 1-PropZInt**. Selecting this option brings you to the next screen, where you simply fill in the needed information. Based on the previous example, we observed 76 out of 267 students who did not feel safe on campus and wanted a 95% confidence interval.

The needed information is entered as shown here:

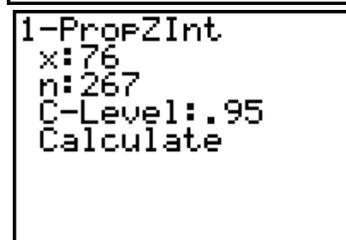
The value for x is the number of observations that has the quality you are looking for. In this example it is the number of surveyed residences with a college graduate. Of course, n is our sample size. C-Level refers to the confidence level, which is 95%, or 0.95 in this case. Now select

Calculate, and the confidence interval is provided for you along with the sample proportion and sample size.

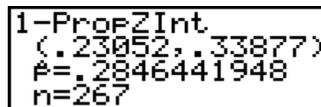
Notice how this confidence interval matches (within rounding) to the confidence interval we calculated by hand. The TI-83/84 uses the notation \hat{p} to represent the sample proportion. This helps reduce the potential for confusion between the notation



```
EDIT CALC RESUME
7↑ZInterval...
8:TInterval...
9:2-SampZInt...
0:2-SampTInt...
▣1-PropZInt...
B:2-PropZInt...
C↓X²-Test...
```



```
1-PropZInt
x:76
n:267
C-Level:.95
Calculate
```



```
1-PropZInt
(.23052, .33877)
p̂=.2846441948
n=267
```

used for a sample proportion and the notation used for a p-value.

Hypothesis Testing for Proportions

The test statistic for a proportion is: $z = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}$. Notice that we have $\pi(1 - \pi)$ in the

denominator rather than pq . The reason for this is the standard error of the mean used in a hypothesis test is the standard error under the null hypotheses. The null hypotheses is about the true proportion, not the sample proportion.

Campus Safety Example

A survey was conducted to estimate the proportion of students who did not feel safe on campus. A total of 76 students, from a sample of 267, said they did not feel safe on campus. It has been believed that 30% of all students on campus did not feel safe. Do your findings show the percent of those worried about their safety is significantly less than the 30% previously believed?

$$H_o: \pi = 0.30 \quad H_A: \pi < 0.30$$

The TI-83/84 handles hypotheses tests for proportions quite nicely. The hypotheses test is located under **STAT > TESTS** and is listed as option **5: 1-PropZTest**.

Selecting this option brings you to the following screen, where you must fill in the needed information. Based on the previous example, we have a hypothesized proportion of 0.27, which is represented by the TI-83/84 as p_o . The number of observations with the desired quality (feel unsafe) was 76 and the sample size was 267. The alternative hypotheses is less than, so we select $< p_o$.

Selecting this option brings you to the following screen, where you must fill in the needed information. Based on the previous example, we have a hypothesized proportion of 0.27, which is represented by the TI-83/84 as p_o . The number of observations with the desired quality (feel unsafe) was 76 and the sample size was 267. The alternative hypotheses is less than, so we select $< p_o$.

We now have the option to **Calculate** or **Draw**. The Calculate option is shown here. The p-value is 0.292.

```
EDIT CALC 11:31E
1:Z-Test...
2:T-Test...
3:2-SampZTest...
4:2-SampTTest...
5:1-PropZTest...
6:2-PropZTest...
7:ZInterval...
```

```
1-PropZTest
P0: .3
x: 76
n: 267
PROP#P0 (<P0)P0
Calculate Draw
```

```
1-PropZTest
PROP<.3
z= -.5475434381
P=.2920026871
P#=.2846441948
n=267
```


Inferences Concerning The Mean Difference Using Two Dependent Samples

Example: A group of women in a large city were given instructions on self-defense. Prior to the course, they were tested to determine their self-confidence. After the course they were given the same test. A high score on the test indicates a high degree of self-confidence. Do these self-confidence scores indicate that the course significantly increased their self-confidence?

Woman	1	2	3	4	5	6	7	8	9
Before Course	6	10	8	6	5	4	3	8	5
After Course	7	12	7	5	8	6	5	8	6
Differences	1	2	-1	-1	3	2	2	0	1

This is clearly paired data in that we have a test / re-test situation. The before and after data is dependent upon each other in that the measurement is taken from the same person.

We want to know if the self-defense course significantly increased the self-confidence of those women who took the course. For this reason I chose to subtract the before course scores from the after course scores, expecting mostly positive values for the differences if the after course scores are, in general, larger. This leads to an alternative hypothesis of greater than zero. Now I can set out to test if the averages of the differences are greater than zero.

If I placed the Before data in L1 and the After data in L2, then the TI-83/84 will calculate the differences.

L1	L2		3
6	7		-----
10	12		
8	7		
6	5		
5	8		
4	6		
3	5		
L3 = L2 - L1			

L1	L2	L3	3
6	7	1	
10	12	2	
8	7	-1	
6	5	-1	
5	8	3	
4	6	2	
3	5	2	
L3(1) = 1			

The null and alternative hypothesis are stated as follows:

$$H_o: \mu_d = 0 \quad (\text{There is no difference})$$

$$H_A: \mu_d > 0 \quad (\text{The course increases confidence})$$

Comparing Two Parameters

The average of the differences is 1, and the standard deviation is 1.414. This leads to the following observed value of t along with the associated p-value.

```
1-Var Stats
x̄=1
Σx=9
Σx²=25
Sx=1.414213562
σx=1.333333333
↓n=9
```

```
T-Test
Inpt:  Data  Stats
μ₀: 0
List: L₃
Freq: 1
μ: ≠μ₀ <μ₀ >μ₀
Calculate Draw
```

```
T-Test
μ>0
t=2.121320344
P=.0333439996
x̄=1
Sx=1.414213562
n=9
```

The Difference Between Means Using Two Independent Samples

Consider the same data used in the previous example and suppose the data were actually independent. We want to know if the After average is greater than the Before average. To complete a 2-sample t-test when the samples are independent of each other, we must first decide if the variances are equal or not equal. To do this, we will use an F-test.

Inferences Concerning The Ratio of Variances Using Two Independent Samples (F-test)

The formal hypothesis statement will look like this: $H_0: \frac{\sigma_B^2}{\sigma_A^2} = 1$ $H_a: \frac{\sigma_B^2}{\sigma_A^2} \neq 1$

The TI-83/84 will handle this test for us. Under the **STAT** > **TEST** menu, select **D: 2-SampFTest**.

```
EDIT CALC  TEST
0↑2-SampTInt...
A: 1-PropZInt...
B: 2-PropZInt...
C: x²-Test...
 2-SampFTest...
E: LinRegTTest...
F: ANOVA(
```

The Difference Between Means Using Two Independent Samples

Selecting the 2-Sample F-test provides the following input screen.

If we had the data entered in our calculator, then we need to enter the list name where the data is stored, along with the alternative hypothesis, and then select **Calculate** or **Draw**. In this example we have the actual data, so we will select **Data** and fill in the needed information.

```

2-SampFTest
Inpt:  Data  Stats
List1:L1
List2:L2
Freq1:1
Freq2:1
σ1:  =σ2  <σ2  >σ2
Calculate Draw
  
```

Selecting **Calculate** provides the following information:

The p-value is large, so we would fail to reject the null hypotheses. In English this means there is insufficient evidence to suggest the variances are different. This is important to know because the two sample t-test on the TI-83/84 asks if we want to pool the variances or not. If we can not show the variances are different, then we will pool them. Pooling tells the calculator to use an estimate for the common variance. If the F-test suggested the variances are different, then we would tell the calculator not to pool the sample variances. In this example, we will pool.

```

2-SampFTest
σ1≠σ2
F=1.054216867
P=.9423055963
Sx1=2.20479276
Sx2=2.14734979
↓X1=6.111111111
  
```

2-Sample t-test for Independent Data

<pre> EDIT CALC <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 1:Z-Test... 2:T-Test... 3:2-SampZTest... 4:2-SampTTest... 5:1-PropZTest... 6:2-PropZTest... 7↓ZInterval... </pre>	<pre> 2-SampTTest ↑List1:L1 List2:L2 Freq1:1 Freq2:1 μ1:≠μ2 <input type="checkbox"/> <μ2 <input type="checkbox"/> >μ2 Pooled:No <input type="checkbox"/> Yes Calculate Draw </pre>
--	--

Now **Calculate** which provides the following output.

Notice how large the p-value is when we considered the data to be independent as compared to the p-value when the data was thought of as dependent. For this reason it is important to understand the difference between dependent and independent data.

```

2-SampTTest
μ1<μ2
t=-.9747546499
P=.17209494
df=16
X1=6.111111111
↓X2=7.111111111
  
```

The before and after data are clearly dependent; however, it serves well to demonstrate the mechanics of the independent t-test on the TI-83/84 while simultaneously demonstrating the potential severity of an improper technique.

The Difference of Two Proportions

Consider the following scenario. A recent survey consisting of 125 males and 142 females showed 31 males and 45 females did not feel safe on campus. Do the data suggest the proportion of females who do not feel safe on campus is greater than the proportion of males who do not feel safe on campus?

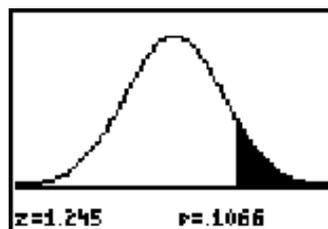
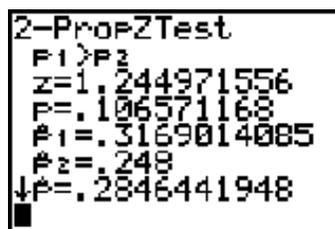
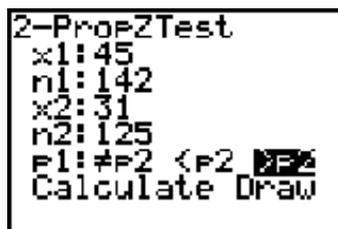
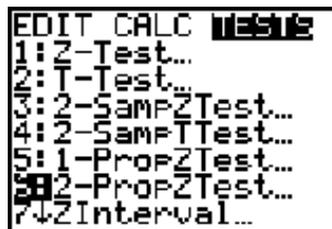
Hypothesis Test

$$H_0: \pi_F - \pi_M = 0 \quad H_A: \pi_F - \pi_M > 0$$

Here we are comparing the proportion of males and females where π_F represents the true proportion of females and π_M represents the true proportion of males. The sample statistics are

$$p_F = \frac{45}{142} \text{ and } p_M = \frac{31}{125}.$$

To complete the hypothesis test go to **Stat>Tests>6: 2-PropZ-Test** and fill in the necessary information. Both the **Calculate** and **Draw** results are shown here.



Confidence Interval

The confidence interval option is obtained by going to **Stat>Tests>B: 2-PropZInt**. The input screen looks very similar to the input screen for a hypothesis test. Fill in the required information as before, along with the confidence level in decimal format (not percent) and then **Calculate**.

<pre>2-PropZInt x1:45 n1:142 x2:31 n2:125 C-Level: .95 Calculate</pre>	<pre>2-PropZInt (-.0387, .17655) p1=.3169014085 p2=.248 n1=142 n2=125</pre>
--	---

The 95% confidence interval for the difference between the two proportions is (-0.3087, 0.17655).

The TI-83/84 reports the value of the correlation coefficient as part of the output for regression, if the calculator's *diagnostics* setting is turned on. To turn the diagnostics setting on, select the calculator's catalog by pressing **2nd** CATALOG.

Now type the letter **D**, which will cause the catalog to jump to the commands that start with the letter "D." Use the down arrow key until the arrow points to **DiagnosticOn**. Press **ENTER** and the command will paste to the home screen.

Press **ENTER** again and your calculator will be placed in the "diagnostics on" condition.

This only needs to be done once. Your calculator will remain in this mode until you repeat the process, selecting **DiagnosticOff** or you return your calculator to its original factory settings.

Finding The Regression Line / Pearson's Correlation

Typically a regression line is sought for the purpose of prediction. Consider the following example.

An article read by a physician indicated that the maximum heart rate an individual can reach during intensive exercise decreases with age. The physician decided to do his own study. Ten randomly selected men performed exercise tests, and their peak heart rates were recorded. The results are listed below.

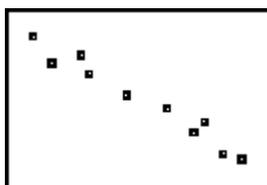
Age (years)	30	38	34	26	29	44	46	41	42	24
Heart Rate	188	183	185	190	191	176	175	179	181	194

Determine the regression equation and use the equation (if appropriate) to estimate the maximum heart rate of a man who is 33 years of age. This suggests we will be using the age in years as the predictor variable (x-variable) to predict heart rate. The heart rate is then the response variable (y-variable). I will place the x-variable data in L1 and the y-variable data in L2.

Correlation and Simple Linear Regression

Before we determine the regression equation and make a prediction, we should first look at our data. Remember, simple linear regression is using a straight line as a statistical model. If the data is not linear, then perhaps we shouldn't even be considering linear regression as a model.

TI-83/84 scatter plot of our data.

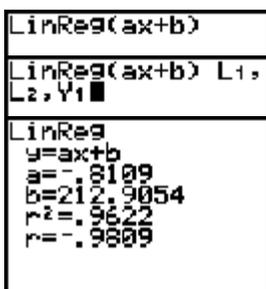


The data does appear to have a strong linear trend, suggesting linear regression may be appropriate.

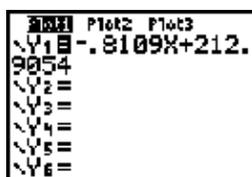
Select **LinReg** from the **STAT > CALC** menu as shown.



This will paste the command in your home screen.

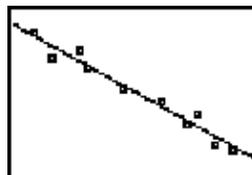


Now we will add the location of the x - and y -variables. In addition, we will tell the TI-83/84 to store the regression equation in Y_1 by selecting **VARS > Y-VARS > FUNCTIONS > Y1**. Adding the Y_1 command at the end stores the regression equation in Y_1 .



When you plot the line $\hat{y} = -0.81x + 212.87$ over your scatter plot, you will see something that looks like this:

This gives us a rough idea as to just how well our linear model fits our data.



Notice the correlation coefficient, r , is reported as $-.9809$.

A Nonparametric Alternative to Pearson's Correlation

See Spearman's Correlation in the Programs chapter.

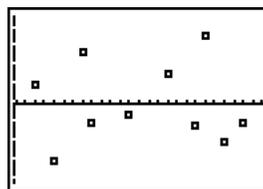
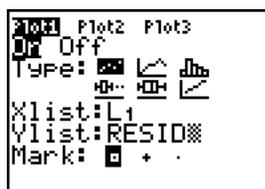
Checking Assumptions

As part of the process of checking our assumptions for linear regression, we need to examine the residual plot. Before you actually look at the residual plots, it may be a good idea to turn Y1 off, else you may get both the regression line and the residuals on the same screen, which will be confusing. Once Y1 is turned off, go to **STAT PLOTS** as shown below.



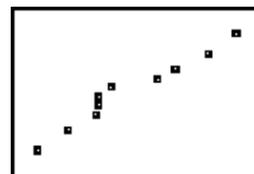
The list named **RESID** is generated automatically by the TI-83/84 when the regression line is calculated. To paste the list to the Stat Plot window, simply select **2nd List** and scroll down to **RESID**. Selecting **RESID** will cause it to paste directly to the Stat Plot window. The resulting residual plot is shown here.

A scatter plot of the residuals against the x values is obtained with the following TI-83/84 commands.



The horizontal line observed in the middle of the plot is the x -axis.

The normality of the residuals can be checked by producing a normal plot of the residuals. Normal plots were discussed in the *Graphical Displays* chapter.



Inferences

Correlation by itself is not an inferential technique. Rather, correlation is a descriptive statistic in that it describes the level of linear association between two data sets. Typically, having calculated a value for the correlation coefficient, we want to determine if it is different from zero since a correlation of zero says there is no linear relationship. Testing to determine if the correlation coefficient is different from zero is equivalent to testing if the slope of the regression line is non-zero. For this reason, we can use the same hypothesis test to address both issues.

$$H_o: \rho = 0$$

$$H_A: \rho \neq 0$$

The TI-83/84 handles this with a t-test. Select **STAT > TESTS > E: LinRegTTest**.

Enter the location for the x- and y- data, along with the appropriate alternative hypothesis and then calculate.

The p-value is 0.00000028, which is strong evidence in support of the true correlation coefficient (hence the true slope of the regression line) being less than zero.

```
EDIT CALC TESTS
0:1-2-SampTInt...
A:1-PropZInt...
B:2-PropZInt...
C:X2-Test...
D:2-SampFTest...
E:LinRegTTest...
F:ANOVA()
```

```
LinRegTTest
Xlist:L1
Ylist:L2
Freq:1
0 & ρ ≠ 0  > 0
RegEQ:
Calculate
```

```
LinRegTTest
y=a+bx
0 < 0 and ρ < 0
t=-14.2675
P=2.8387E-7
df=8.0000
↓a=212.9054
```

One Way Analysis of Variance (One Way ANOVA) is used to compare multiple independent samples.

One-Way ANOVA Example

During cooking, doughnuts absorb fat. Lowe wished to determine if the amount of fat absorbed is dependent on the type of fat used. An experiment was designed where 4 different types of fats were tested by preparing 6 different batches of doughnuts (a batch consisted of 24 doughnuts) for each of the 4 different types of fat. The following table provides the amount of fat absorbed by each batch.

Fat Type	1	2	3	4
	164	178	175	155
Amount	172	191	193	166
Absorbed	168	197	178	149
	177	182	171	164
	156	185	163	170
	195	177	176	168

*Data Source: Lowe, B. 1935. Data, Iowa Agric. Exp. Stn.
(Statistical Methods, Snedecor and Cochran, Seventh Edition)*

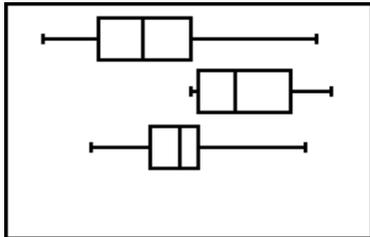
Homogeneity of variance assumption: The TI-83/84 will only allow 3 box plots to be displayed at the same time. As such, if you are using a visual method to satisfy the homogeneity of variance assumption, then we must do two separate displays.

The first will show Type 1, Type 2, and Type3 fats. This is accomplished by activating StatPlots 1 through 3. If L1 is placed in StatPlot-1, then the top box plot will be of the data in L1. The plots will be displayed in the order StatPlot-1, StatPlot-2 and StatPlot-3 from top to bottom on your calculator display.

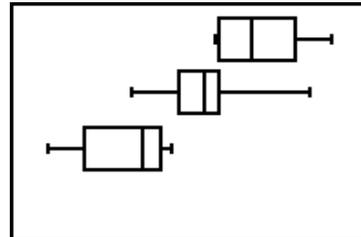
The second display will show Type 2, Type 3 and Type 4. To do this, simply assign the data in L2 to StatPlot-1, L3 to StatPlot-2 and L4 to StatPlot-4. The reason for the overlap is to make it easier for you to make a visual inspection of the variability.

From here we can make an intelligent decision as to how similar the variances are.

Type 1, Type 2, Type 3



Type 2, Type3, Type 4



By visual inspection, it would seem the variability is more similar than dissimilar. If you are using a formal F-test, then examine the 1-variable statistics for each group and compare the largest and smallest sample variance.

Fat Type 1

```
1-Var Stats
x̄=172
Σx=1032
Σx²=178394
Sx=13.34166406
σx=12.17921727
↓n=6
```

Fat Type 2

```
1-Var Stats
x̄=185
Σx=1110
Σx²=205652
Sx=7.771743691
σx=7.094598885
↓n=6
```

Fat Type 3

```
1-Var Stats
x̄=176
Σx=1056
Σx²=186344
Sx=9.879271228
σx=9.018499506
↓n=6
```

Fat Type 4

```
1-Var Stats
x̄=162
Σx=972
Σx²=157802
Sx=8.221921916
σx=7.505553499
↓n=6
```

Fat Type 1 has the largest standard deviation (hence variance) and fat Type 4 has the smallest standard deviation (hence variance). So an F-test between types 1 and 4 would be a formal method of testing the homogeneity of variance assumption.

```
2-SampFTest
Inpt: Data Stats
List1:L1
List2:L4
Freq1:1
Freq2:1
σ1:≠σ2 <σ2 >σ2
Calculate Draw
σ1≠σ2
F=2.633136095
p=.3115867835
Sx1=13.3416641
Sx2=8.22192192
↓x1=172
```

Based on the p-value, it seems very reasonable to assume the variances are more similar than dissimilar.

Calculating One-Way ANOVA on the TI-83/84

Calculating a One-Way ANOVA on your TI-83/84 is quite simple. From the **STAT > TESTS** menu, select option **F: ANOVA**.

Selecting **ANOVA** will paste the command to your home screen. Now simply tell your calculator where the data is stored. In this case it is stored in L1, L2, L3, and L4 so we will pass the calculator this information. Be certain to separate each list name with a comma.

```
EDIT CALC TESTS
0↑2-SampTInt...
A:1-PropZInt...
B:2-PropZInt...
C:χ²-Test...
D:2-SampFTest...
E:LinRegTTest...
F:ANOVA(
```

Notice the p-value, which is 0.00687. The p-value is very small in this case, so we will reject the null hypothesis which stated the means for the groups are the same. In English we are saying there is sufficient evidence to suggest the type of fat used makes a difference in the amount of fat absorbed by the doughnuts.

To determine which of the fat types are different from each other in terms of the amount of fat absorbed, requires multiple comparison techniques. The TI-83/84 does not do multiple comparisons without the aid of a separate program written specifically for that task.

```
ANOVA(L1,L2,L3,L4)
↓
One-way ANOVA
F=5.4063
P=.0069
Factor
df=3.0000
SS=1636.5000
↓ MS=545.5000
One-way ANOVA
↑ MS=545.5000
Error
df=20.0000
SS=2018.0000
MS=100.9000
SxP=10.0449
```

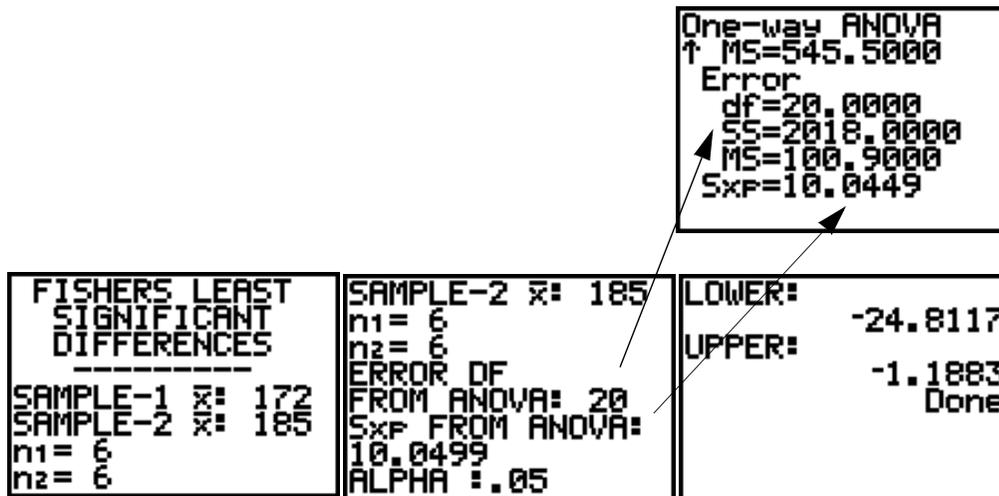
Nonparametric Alternative: The Kruskal-Wallis Test

See the Kruskal-Wallis Test in the Programs chapter.

Multiple Comparisons - Fisher's LSD

Fisher's Least Significant Difference technique can be found in the program FLSD. Fisher's LSD compares two means at a time, looking for a difference between the two by use of a confidence interval. Unfortunately, the program on the TI-83/84 must be run separately for each comparison.

Having completed the ANOVA and now believing that at least one of the averages is different from the others, we must investigate further to see which is different. Suppose we were interested in testing the 1st and 2nd means. The average from the first group is 172, and the average from the second group is 185. The sample size is 6 for each group. With this information, and additional information obtainable from the ANOVA output, we can complete Fisher's LSD for these two groups.



The confidence interval calculated by Fisher's LSD method is (-24.8117, -1.1883). Since 0 is not in the interval (both values are negative), the two averages are statistically different from each other based on Fisher's LSD. This procedure needs to be repeated for every comparison that needs to be made. If zero were in the interval, then we would not have sufficient evidence to suggest these two group means are different from each other.

Chi-Squared Goodness of Fit

The TI-83/84 is not equipped with a Chi-Squared Goodness of Fit test. This test must be completed by hand or by the use of a program written for that purpose. A program written for goodness of fit will be discussed later in the *Programs* chapter.

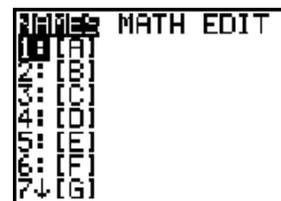
Chi-Squared Test of Independence

Consider data from a survey regarding the number of Hamburgers consumed and Cholesterol Level.

How Often/ Cholesterol Level	Low	Medium	High
Daily	8	31	49
Often	12	27	31
Never	25	10	5

The TI-83/84 will calculate the value of the χ^2 test statistic along with the associated p-value. This is completed by placing the data into a matrix. Specifically, we will be using the matrix identified as “Matrix A.” Matrix A is the default matrix used by the calculator.

Step #1: Press the **MATRIX** key on your calculator. This will result in the following screen.



Select **EDIT** matrix “A” as shown here.



The table that contains the data, in this example, is known as a 3x3 table in that there are 3 rows and 3 columns of data. Adjust the dimension of the table by entering a “3”, followed by **ENTER**,

Categorical Data Analysis

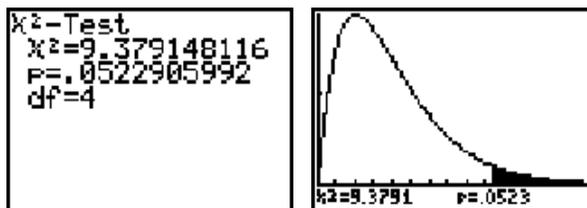
and then another “3”, followed by **ENTER** once again. The table then changes to a 3x3 table containing all zero’s. Simply input the appropriate data values and then **QUIT**.

<pre> MATRIX[A] 3 x3 [0 0 0] [0 0 0] [0 0 0] 1:1=0 </pre>	<pre> MATRIX[A] 3 x3 [12 21 28] [15 26 31] [21 15 14] </pre>
---	--

From the home screen, select **STAT > TESTS > C:χ²-TEST**.

<pre> EDIT CALC TESTS 0:2-SampTInt... A:1-PropZInt... B:2-PropZInt... C:χ²-Test... D:2-SampFTest... E:LinRegTTest... F:ANOVA(</pre>	<pre> χ²-Test Observed: [A] Expected: [B] Calculate Draw </pre>
--	---

The test uses Matrix-A as the default matrix to store the observed frequencies. The TI-83/84 will then calculate the expected frequencies and store them in Matrix-B. You may now select either **Calculate** or **Draw**. The results of each, based on the above data, is shown below.



Matrix-B can be viewed to check the assumptions needed for the χ^2 test of independence. Use the same procedure you used to construct matrix A to view the contents of matrix B.

<pre> MATRIX[B] 3 x3 [20 30.222 37.778] [15.909 24.04 30.051] [9.0909 13.737 17.172] </pre>

All programs, once transferred to your calculator, can be found under the **PRGM** menu. Once the **PRGM** key is selected, a list of all available programs is displayed. To execute a program, select it from the menu by highlighting the name and pressing **ENTER** or pressing the identifier key listed in front of the name. Based on the screen shown here, I could select the program **GOODFIT** simply by pressing the number “2.” Likewise, **FREQDIST** could be selected by pressing “1.” There is no need to highlight the *identifier key* prior to selecting a program.



```

EDIT EDIT NEW
1: ACCEL
2: CLEAN
3: FLSD
4: FREQDIST
5: GOODFIT
6: KWTEST
7: MREG

```



Once I select the program, the name is pasted to the homescreen. The program is actually executed by pressing **ENTER**.

The following is a brief explanation of the use of programs we have written or obtained for the TI-83/84.

CLEAN

The program **clean** was written to restore the basic calculator settings that are commonly changed when other programs are run. **Clean** will turn off all functions and statplots. It will also turn the Axes on, clear the draw screen along with the home screen and set your calculator decimal display accuracy to *float*. This program is of most use after running, as an example, the program **normplot**, which activates statplot1 along with function Y_0 . The next time you plot or graph something, the normal plot information may still be in the graphing window, which can be annoying. You can manually turn the statplot and function off, along with turning your axes back on, or simply run **clean** to handle it all for you. Resetting these default values could be done within each program, but doing so would create limitations to some programs, hence the reason **clean** was written.

FLSD (Fisher's Least Significant Differences)

The program **flsd** will calculate a Fisher's Least Significant Differences confidence interval needed in the post hoc comparison once a significant p-value is found with an ANOVA. To be meaningful, a CI for each pair should be computed. The following example is based on the Algebra Exam Score data for Instructors 1, 2 and 3 as found in the ANOVA chapter of the text book. Only one comparison is shown here between Instructors 1 and 2. To complete the analysis, the same process should be repeated for comparisons between Instructors 1 and 3, along with Instructors 2 and 3.

All values needed to run this program can be found from the ANOVA output and 1-Var Stats for each data list.

Programs

FISHERS LEAST SIGNIFICANT DIFFERENCES ----- SAMPLE-1 \bar{x} : 28.4 231 SAMPLE-2 \bar{x} : 29.0 769	769 n1= 26 n2= 13 ERROR DF FROM ANOVA: 54 SXP FROM ANOVA: 5.6115 ALPHA :.05	LOWER: -4.4405 UPPER: 3.1329 Done
--	--	--

The interpretation of the interval is the same as any confidence interval comparing two parameter means. Since 0 is in the interval, we conclude there is not a significant difference between Instructors 1 and 2.

FREQDIST (Frequency Distribution Table)

The program **freqdist** will display a frequency distribution of a univariate data set. Consider the following data set, which consists of the launch temperatures for the space shuttle, up to and including the Challenger launch.

31	53	57	58	63
66	67	67	67	68
69	70	70	70	70
72	73	75	75	76
76	78	79	80	81

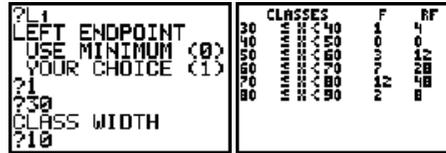
To run the program, select **freqdist** from the programs menu and press enter. The first prompt asks for the location of your data. Suppose we have the Challenger data in list L1. We would then enter L1 at this prompt as shown below. The next prompt is asking questions regarding the left endpoint of the frequency distribution table. If you want to use the smallest value in the data set, then enter 0. The program gives you the ability to specify the left endpoint (Your Choice). I know that the smallest data value is 31 and the largest is in 81. As such, I may want to start the frequency distribution at 30. To do this I will select 1, as shown below. When you select "1" and press ENTER, you will immediately be presented with another "?" which is asking you for the left endpoint you want to use. I will enter 30 for this example.

FREQ EDIT NEW 1: ACCEL 2: CLEAN 3: FLSO 4: FREQDIST 5: GOODFIT 6: KWTEST 7: MREG	DATA IS IN WHICH LIST ?L1	DATA IS IN WHICH LIST ?L1 LEFT ENDPOINT USE MINIMUM (0) YOUR CHOICE (1) ?1	DATA IS IN WHICH LIST ?L1 LEFT ENDPOINT USE MINIMUM (0) YOUR CHOICE (1) ?1 ?30
---	---------------------------------	--	---

The final prompt is for the class width. Knowing my data runs from 31 to 81, I will select a class width of 10. This simply seems reasonable based on my data. The TI-83/84 will now produce a fre-

GOODFIT

quency distribution of your data as shown below.



CLASSES	F	RF
30-40	4	4%
40-50	10	10%
50-60	12	12%
60-70	15	15%
70-80	10	10%
80-90	6	6%

The output shows the classes, frequencies (denoted by F), and relative frequencies (denoted by RF). The 4 under relative frequencies is indicating the first class makes up 4% of the entire data set.

A word of caution. The program is not able to properly display a frequency distribution with more than 9 classes. The reason is because this program does not have the ability to “scroll” down. If you had 10 or 11 classes, the last 2 classes would be off the bottom of the screen and will not be displayed.

GOODFIT

GOODFIT is a TI-83/84 program that calculates the χ^2 test statistic, along with the associated p-value, for the Goodness of Fit Test. A Goodness of Fit Test is not inherent to the TI-83/84, hence the need for this program.

Once the program has been executed, you will be presented with the following screen. The first option, which is selected by pressing the number “1,” indicates the data you will be examining is stored in a list, such as L1, or L2. The second option, selected by pressing the “2,” indicates the data needs to be input into the calculator.



Stored In List

If you have stored your data in a list, such as L1, then you must also have stored the probabilities, as defined under the null hypothesis, in another list, such as L2.

Consider the following example: It is believed that a die is “loaded.” To test this theory, the die is rolled 72 times and the results recorded. The results are shown in the table below.

Die Value	1	2	3	4	5	6
Observed Count	8	13	12	15	10	14
Expected Probability	1/6	1/6	1/6	1/6	1/6	1/6

If the die is not loaded, then the probability of getting a “1” is the same as the probability of getting any other value, which is 1/6 or approximately 0.16667.

Programs

Let's enter the observed values in L1 and the expected probabilities in L2, as shown here.

L1	L2	L3	3
8	.16667	*****	
13	.16667		
12	.16667		
15	.16667		
10	.16667		
14	.16667		
-----	-----		
L3(1)=			

Now, run the program and tell it where to find the necessary information.

<pre> PROGRAM KWTEST 1) STORED IN LIST 2) NEED TO INPUT </pre>	<pre> LIST FOR OBSERVED COUNTS ?L1 </pre>	<pre> LIST FOR OBSERVED COUNTS ?L1 LIST FOR PROBABILITIES ?L2 </pre>	<pre> χ² = 2.8333 P-VALUE = .7257 </pre>
--	---	---	--

The value of the χ^2 test statistic is reported along with p-value.

Need To Input

The program allows you to input both the observed values along with the expected probabilities under the null hypothesis, rather than storing that information in a list. If you select this option, you will be asked for the “observed list”, which is the observed values. You will then be asked for the “expected probabilities.” The information *must* be passed on to the program in a specific format. That format starts with an opening bracket “[”. Each data value is then separated by a comma “,”. The list concludes with an ending bracket “]”. The following output shows an example of the syntax.

<pre> OBSERVED LIST? ?8,13,12,15,10, 14} </pre>	<pre> CHI-SQUARE = 2.8333 </pre>
<pre> EXPECTED PROBABILITIES? ?1/6,1/6,1/6,1/6,1/ 6,1/6,1/6} </pre>	<pre> P-VALUE = .7257 Done </pre>

The probabilities could have been entered each as 0.1667 rather than 1/6.

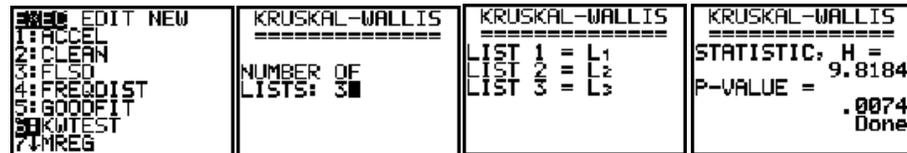
KWTEST (Kruskal-Wallis Test)

The program **kwtest** is the Kruskal-Wallis test. The Kruskal-Wallis test is a non-parametric procedure used when the assumptions needed for one-way ANOVA are violated. Consider the following data set comprised of three lists of 8 observations each.

List 1	9	2	5	7	8	5	3	4
List 2	15	5	9	7	12	15	15	11
List 3	16	7	8	17	11	13	7	18

NORMPLOT (Normal Probability Plot)

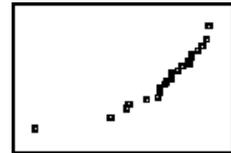
If we were interested in a comparison of the means, we would use ANOVA. However, in this example, the normality assumption is grossly violated hence we will use the Kruskal-Wallis test. The above data has been entered in lists L1, L2, and L3.



NORMPLOT (Normal Probability Plot)

The program **normplot** constructs a Normal Probability Plot. The TI-83/84 has a built in routine for a normal probability plot; however, we felt that a reference line would be useful in terms of making a decision. As such, the program **normplot** was written. The program uses the TI-83/84's normal probability plot routine with a reference line added.

Consider the following data that consists of the launch temperatures for the space shuttle, up to and including the Challenger launch. A standard TI-83/84 normal probability plot is shown here.



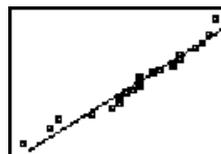
31	53	57	58	63
66	67	67	67	68
69	70	70	70	70
72	73	75	75	76
76	78	79	80	81

Now run the **normplot** program. Just as with the Goodness of Fit Test, you have the option of inputting your data directly into the program, or storing the data in a list. If you have your data stored in a list, then chose option 1. In this example, the data has been stored in L1.

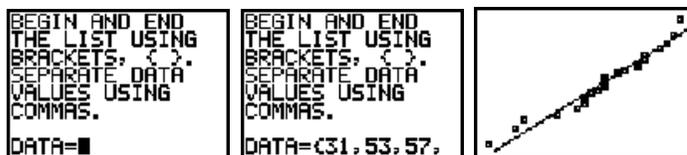


Programs

Now, for exploratory purposes, consider a normal plot of the same data set with the value 31 (the temperature at the time of the Challenger launch) removed. That normal plot is shown here.



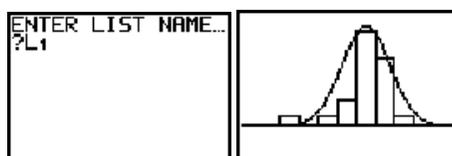
If you decide you would like to input your data directly into the program, choose option 2. You will then receive instructions as to how the data needs to be entered. The instructions, the first few values from the Challenger data set and output are shown below.



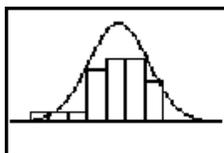
NRMHST (Histogram with Normal Curve)

The program **nrmhst** was written more as a teaching tool than a diagnostic tool. The program uses the built-in TI-83/84 histogram function and superimposes the normal probability density function based on the mean and standard deviation from the sample. The density function is then scaled to properly fit the screen at the same time the histogram is displayed; it is not scaled to match the area represented by the histogram.

Consider the Challenger data once again. Run the **nrmhst** program. The only prompt is asking for the name of the list where the data is stored. In this case, it is in L1.



Now, for exploratory purposes, consider the same data set with 31 (the temperature at the time of the Challenger launch) removed.



PREDCI (Regression - Confidence Interval for a Predicted Value)

PREDCI (Regression - Confidence Interval for a Predicted Value)

The program **predci** computes a confidence interval for a predicted value based on simple linear regression. Consider the Cigarettes vs. CHD Mortality data. Having calculated the regression equation and finding a predicted mortality rate of 195.41 for a cigarette value of 2850, we can now complete the prediction interval. All values needed come from 1-VarStats for Cigarettes and the regression output.

```
CONFIDENCE LEVEL  ? .95
INPUT Sx          ? 2807.9766
INPUT n          ? 21
INPUT x FROM THE ? 2148
REGRESSION TEST  ? 40.2077
INPUT x          ? 2148
INPUT THE       ? 195.41
PREDICTED y
INPUT THE       ? 195.41
OBSERVED x     ? 2850

LOWER BOUND      107.7370
UPPER BOUND      283.0830
Done
```

The conclusion is “We are 95% certain the true CHD Mortality rate per 100000 for a country with a value of 2850 cigarettes adult/year is (107.7370, 283.0830).”

ODDSCI (Odds Ratio Confidence Interval)

The program **odds** computes a confidence interval for the odds ratio of a 2x2 table. The odds ratio is a nice measure of independence for 2x2 tables. This is particularly true when the individual proportions are very small and you are not able to show the necessary conditions for the normality assumption regarding the sampling distribution of the sample proportions (assuming you were interested in comparing proportions, possibly with a difference of proportions hypothesis test).

The following data was obtained from a report on the relationship between aspirin use and heart attacks by the Physician’s Health Study Research Group at Harvard Medical School. The physician’s Health Study is a randomized clinical trial, testing whether aspirin taken regularly reduces mortality from cardiovascular disease. Every day, physicians participating in the study took either one aspirin tablet or a placebo. The study was blind.

	Placebo	Aspirin
Heart Attack Yes	189	10845
Heart Attack No	104	10933

The odds ratio is computed as $\frac{n_{11}n_{22}}{n_{12}n_{21}} = \frac{(189)(10933)}{(10845)(104)} = 1.832$. This is interpreted as “for every

1.832 physicians that had a heart attack in the placebo group, 1 physician had a heart attack in the

Programs

aspirin group. The program **oddscl** will calculate the sample odds ratio along with a confidence interval. Suppose we wanted a 95% confidence interval for the above example.

<pre> N11: 189 N12: 10845 N21: 184 N22: 10933 SIGNIFICANCE LEVEL: .95 </pre>	<pre> ODDS RATIO: 1.822 CONFIDENCE INTERVAL (1.44, 2.331) PRESS ENTER TO CONTINUE </pre>
--	--

RANKSUM (Wilcoxon Rank-Sum Test)

The program **ranksum** is the *Wilcoxon Rank-Sum Test*. To demonstrate its use, consider a study undertaken to determine whether the number of trials required to master a task is influenced by taking a particular drug. The dependent measurement is the number of trials required to master a given task. A group of 28 subjects are randomly assigned to the experimental and control groups. Those assigned to the experimental group are given the drug, and those assigned to the control group are given a placebo. After a given period of time, the numbers of trials to master the task are recorded as listed here:

Experimental Group	17	15	5	14	18	3	16	19
	13	15	16	17	8			
Control Group	14	10	1	12	11	14	8	10
	2	12	16	12	15	4	12	

Looking at the normal plots suggests the data may not have come from normal parent populations, hence a 2-sample t-test may not be appropriate. As an alternative, we can employ the nonparametric procedure known as the *Wilcoxon Rank Sum Test*. The name of the program in your calculator is **ranksum**. This test computes a p-value based on a two sample t-test from the *ranks* of the data, not the original data. As such, the actual hypothesis test, although commonly worded as a test for the population mean or median, is actually a test of the individual distributions. The question being answered is “are the two distributions the same?” If the distributions are the same, then the means and medians are also the same, hence the logic for often writing the formal hypothesis test in terms of the population mean (μ) or the population median (θ). The program approaches this as a difference between the two distributions, hence the reason the alternative hypothesis is stated in terms of $\neq 0$, > 0 , or < 0 . I entered the Experimental Group in L2 and the Control Group in L3. I am testing to see if the number of times it takes to master a task is higher in the experimental group, or my alternative hypothesis is greater than ($>$).

<pre> WILCOXON RANK SUM ----- <0> <-1> <1> ?1 </pre>	<pre> DATA 1 IN LIST ?L2 DATA 2 IN LIST ?L3 </pre>	<pre> 2-SAMPLE T TEST ON THE RANKS T 2.53 P-VALUE .0089 </pre>
--	---	---

RHO CI (Correlation Coefficient Confidence Interval)

RHO CI (Correlation Coefficient Confidence Interval)

```
R=.73
N=21
ALPHA=.02

LOWER: .3631
UPPER: .9009

PRESS ENTER TO CONTINUE
```

The program **rhoci** computes a confidence interval for rho (ρ), the population correlation coefficient. All that is needed is the sample correlation, the sample size, and the desired value for α . Suppose you calculated a sample correlation coefficient of 0.73 based on a sample size of 21. A 98% confidence

interval is desired, hence $\alpha = 0.02$. The 98% confidence interval is then calculated as (0.363, 0.900). The confidence interval is based on Fisher's z-transformation.

SAMPLESZ (Sample Size)

The program **samplesz** will calculate the needed sample size for collecting data with which the mean or a proportion will be calculated.

When you run the program, you are initially asked if you will be working with a mean or a proportion.

```
SAMPLE SIZE FOR
1: THE MEAN
2: A PROPORTION
```

Sample Size for the Mean

Select 1 to calculate the needed sample size for a mean. The program will then prompt you for the information it needs, starting with the believed standard deviation. Suppose, for this example, information indicates the standard deviation will be 1.3 and we need a margin of error of no more than 0.08 with 95% certainty. We would enter this information as shown below. The sample size needed is then calculated to be $n = 1014.385$ or 1015.

```
STANDARD
DEVIATION: 1.3

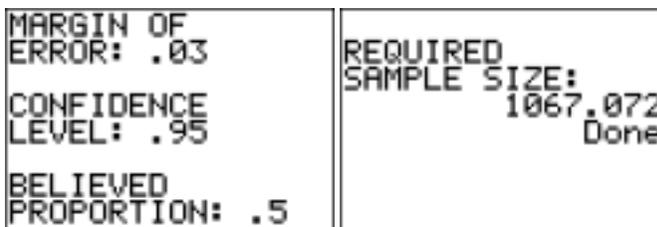
MARGIN OF
ERROR: .08

CONFIDENCE
LEVEL: .95

REQUIRED
SAMPLE SIZE:
1014.385
Done
```

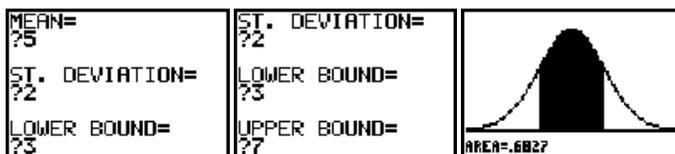
Sample Size for a Proportion

If we selected Proportion, we will need to provide the margin of error, believed proportion, and Level of Confidence. Suppose we were conducting a political poll where we had no idea what the believed proportion was (so we use 0.50 to provide a “safe” calculation), but we needed a margin of error no more than 3% with 95% certainty. The calculated sample size would then be 1067.072 or 1068.

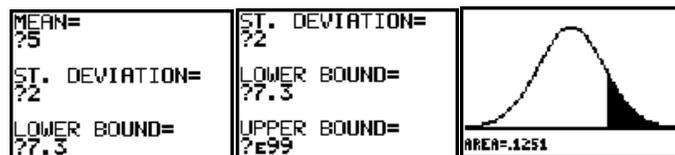


SHADNORM (Shaded Normal Curve)

The program **shadnorm** is an improvement over the built-in “shadenorm” command found under the Distr > Draw menu. Although the TI-83/84’s shadenorm command may have its pedagogical uses, day to day implementation is extremely inconvenient in that you have to figure out the window settings in order to see the graph. The **shadnorm** program will prompt you for the mean, standard deviation, lower bound and upper bound. A normal probability distribution function is then drawn to the screen with the desired area shaded in and numerically identified. The following example shows the program displaying the area between 3 and 7 from a normal distribution with a mean of 5 and a standard deviation of 2.



Suppose we wanted to know the probability of observing a value greater than 7.3 from a normal distribution with a mean of 5 and a standard deviation of 2. The upper bound in this case is infinity. The TI-83/84 does not have an infinity key, but we can do well with simply giving the calculator a very large positive value such as 9999999999. Another option is giving the calculator a value of 10^{99} , which is yet another very large number. The value 10^{99} is typically used for infinity in the calculator and is obtained by pressing 2nd EE 99. The “EE” here is found on your calculator above the “,” key. Likewise, -10^{99} is used for negative infinity.



SIGNRANK (Wilcoxon Sign-Rank Test)

SIGNRANK (Wilcoxon Sign-Rank Test)

The program **signrank** is the *Wilcoxon Sign-Rank Test*, which is used to test for a difference in the medians between two dependent data sets. This is used when the normality assumptions needed for the 2-sample t-test for dependent data are violated. Consider the following data which consists of test scores before and after a one-day remedial training course. Normal plots of the data strongly suggests the “after” data is not normal. The “before” data has been entered in L1 and the “after” data entered in L2.

Student	1	2	3	4	5	6	7	8
Before	60	51	72	43	50	51	42	70
After	73	85	71	72	84	71	71	62

The program name is **signrank**. The alternative hypothesis is set up just as the Wilcoxon Rank-Sum test. Select the appropriate alternative hypothesis and press ENTER. Next, enter the location of the data and press enter. The p-value and the value of the test statistic is then displayed.

<pre>WILCOXON SIGN RANK ----- # 0 (0) < 0 (-1) > 0 (1) ?0</pre>	<pre>DATA 1 IN LIST ?L1 DATA 2 IN LIST ?L2</pre>	<pre>TEST OF = 0 VS. ≠ 0 T -2.9499 P-VALUE .0214</pre>
--	---	--

SIGNTEST (One Sample Sign Test)

The program **signtest** computes a hypothesis test for the population median based on the nonparametric procedure known as the “sign test.” There are a few different approaches to the calculation of the test statistic and p-value. The algorithm employed in this program automatically deletes all ties (a tie is established when an observation has the same value as the median under the null hypothesis), calculating a p-value based on the binomial distribution with n (the number of trials) equal to the sample size minus the number of ties, and p , the probability of a success, equal to 0.5. Consider the following data set, which has a sample median of 6.5.

2	6	9	4	8	0	6	0
9	12	7	13	5	17	8	2
16	12	5	7	4	21	1	3

Suppose we believe this data came from a population with a median of 8. Is there sufficient evidence to suggest the true population median is actually less than 8? This scenario is a left tailed hypothesis test. The following screen shots show the sequence of commands and output from the sign test program from this scenario.

Programs



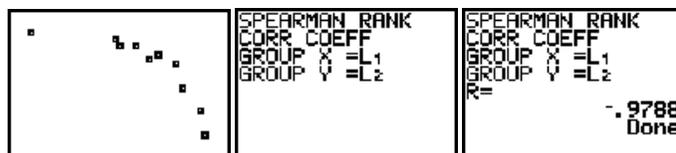
SRCORR (Spearman's Correlation)

Note: *srcorr* and *zrank* were obtained from <http://www.hsu.edu/faculty/lloyd/ti/prgmtabl.htm>.

The program **srcorr** computes Spearman's Correlation Coefficient. It uses the program **zrank** as a subroutine. Consider the following example: A math professor believes there is an association between a student's exam score and their level of test anxiety. The following data was collected.

Anxiety	61	47	82	65	72	84	55	49	75	12
Math Score	76	93	32	79	71	12	87	87	51	98

The scatter plot shows obvious monotonic curvature, making Pearson's Correlation Coefficient inappropriate. Simply run the program **srcorr**. The program will ask you where the x- and y-data are stored. Supply the names of the lists and press ENTER.



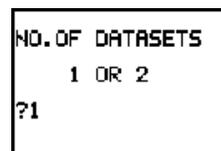
This value of Spearman's Correlation compares to a Pearson's Correlation value of -0.807.

STEMPLOT (Stem-and-leaf Display)

Note: *stemplot* was obtained from <http://www.hsu.edu/faculty/lloyd/ti/prgmtabl.htm>.

The program **stemplot** will produce a stem-and-leaf display to your data. The program does a nice job with small data sets that do not have a great deal of variability. It works well as a teaching tool, but caution must be used when using it as a general purpose tool. The primary limitation is that it does not have the ability to scroll, hence the program will only display the first 10 classes.

Consider the Challenger data set once again. **Stemplot** first asks for the number of data sets you want to display. You may select one, or two. Choosing 1 results in the typical stem-and-leaf display. Choosing 2 results in a back-to-back display.



Programs
