TI 83/84 MANUAL
for Moore's

The Basic Practice of Statistics
Fifth Edition

Patricia Humphrey
Georgia Southern University

W.H. Freeman and Company
New York

ISBN-10: 1-4292-2786-9
ISBN-13: 978-1-4292-2786-5

Contents

Preface

The study of statistics has become commonplace in a variety of disciplines, and the practice of statistics is no longer limited to specially trained statisticians. The work of agriculturists, biologists, economists, psychologists, sociologists, and many others now quite often relies on the proper use of statistical methods. However, it is probably safe to say that most practitioners have neither the time nor the inclination to perform the long, tedious calculations that are often necessary in statistical inference. Fortunately there are now software packages and calculators that can perform many of these calculations in an instant, thus freeing the user to spend valuable time on methods and conclusions rather than on computation.

With their built-in statistical features, Texas Instruments' graphing calculators have revolutionized the teaching of statistics. Students and teachers have instant access to most commonly used statistical procedures. Advanced techniques can be programmed into the calculator which then make it (almost) as powerful as, but much more convenient than, common statistical software packages.

This manual serves as a companion to your W. H. Freeman Introductory Statistics text. Examples either taken from the text, or similar to those in the text, are worked using either the built-in TI calculator functions or programs specially written for the calculator. The tremendous capabilities and usefulness of TI calculators are demonstrated throughout. It is hoped that students, teachers, and practitioners of statistics will continue to make use of these capabilities, and that readers will find this manual helpful.

Programs

All codes and instructions for the programs are provided in the manual; however, they can be downloaded directly from the author's website at **http://math.georgiasouthern.edu/~phumphre/TIpgms/** or from your text's portal.

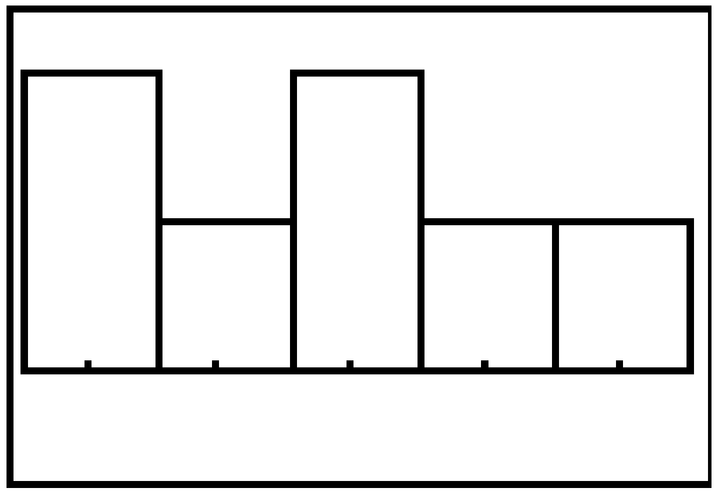
Acknowledgments

My thanks go to W. H. Freeman and Company for giving me the opportunity to revise the manual to accompany their various texts. Special thanks go to Ruth Baruth and editorial assistant Jennifer Albanese for her organization and help in keeping me on schedule. As always, my sincere gratitude goes to Professor Moore and his coauthors for providing educators and students with an excellent text for studying the practice of statistics.

Patricia B. Humphrey
Department of Mathematical Sciences
Georgia Southern University
Statesboro, GA 30460-8093

CHAPTER

0



Introduction to TI Calculators

Introduction

In this chapter we introduce our calculator companion by giving an overview of Texas Instruments' graphing calculators: the TI-83+, -84+, and -89. Read this chapter carefully in order to familiarize yourself with the keys and menus most utilized in this manual. These skills include Home screen calculations and saving and editing lists of data in the `STAT`(istics) editor.

0.1 Key Differences Between Models

All calculators in the TI-83/84/89 series have built-in statistical capabilities. Although a few statistical functions are "native" on the TI-89, most of the topics covered in a normal Statistics course require downloading the Texas Instruments Statistics with List Editor application which is free. Download requires the TI-Connect cable. See the web page http://education.ti.com/us/product/tech/89/apps/appslist.html for more information. This manual assumes the statistics application has been loaded on the calculator. If you have the newer TI-89 Titanium edition, the statistics application comes pre-loaded, and the TI-Connect cable is included with the calculator.

The TI-83 and -84 series calculators are essentially keystroke-for-keystroke compatible; however, the 84 does have some additional capabilities (some additional statistical distributions and tests, for example) with the latest version of the operating system, version 2.41 which is also available for download at education.ti.com.

There are some major differences in calculator operation and menu systems that will in some cases necessitate separate discussions of procedures for the TI-83/84 and TI-89 calculators. Some of these will become apparent in the next section. Not only are there differences between the three series, but there is also a difference in operation between the TI-89 and the TI-89 Titanium edition. When a regular TI-89 is turned on, the user is on the "home screen" similar to that for the TI-83 and -84. On the Titanium, all "functions" on the calculator are essentially applications; when the Titanium edition is first turned on, one must scroll using the arrow keys to locate the desired application, but we'll say more about this later.

0.2 Keyboard and Notation

All TI keyboards have 5 columns and 10 rows of keys. This may seem like a lot, but the best way to familiarize yourself with the keyboard is to actually work with the calculator and learn out of necessity. The keyboard layout is identical on the 83+ and 84+. The layout of the 89 (and 89 Titanium) keyboard is similar, but some functions have been relocated. You will find the following keys among the most useful and thus they are found in prominent positions on the keyboard.

- The cursor control keys |, ~, } and are located toward the upper right of your keyboard. These keys allow you to move the cursor on your screen in the direction the arrow indicates.
- The o key is the leftmost key on the top row of 83 and 84 keyboards. It is utilized more in other types of mathematics courses (such as algebra) than in a statistics course; however you will use the 2 function above it quite often. This is the `STAT PLOT` menu on the -83/84 series. We will discuss 2 functions shortly. On TI-89 calculators, the `Y=` application is accessed by pressing ∞. The `STAT PLOT` menu on the TI-89 series is found inside the Statistics application.
- The ⊃ key is in the bottom left of the keyboard. Its function is self-explanatory. To turn the calculator off, press 2⊃.
- The ⊆ key is in the bottom right of the keyboard. You will usually need to press this key in order to have the calculator do what you have instructed it to do with your preceding keystrokes.

- The σ key is in the upper right of the keyboard. On the TI-89, GRAPH is ∞□.

As mentioned briefly above, most keys on the keyboard have more than one function. The primary function is marked on the key itself and the alternative functions are marked in color above the key. Actual color depends on the calculator model. Next we describe how to engage the functions that appear in color.

The ψ key

The color of this key varies with calculator model, but in all models it is the leftmost key on the second row. If you wish to engage a function that appears in the corresponding color above a key, you must first press the ψ key. You will know the second key is engaged when the cursor on your screen changes to a blinking ⇒. As an example, on a TI-83 or -84 if you wish to call the STAT PLOTS menu, which is in color above the o key, you will press ψ o.

The □ key

You will also see characters appearing in a second color above keys that are mostly letters of the alphabet. There are some situations in which you will wish to name variables or lists and in doing so you will need to type the names. If you wish to type a letter on the screen you must first press the □ key. The color of this key depends on the model of calculator and corresponds to the color of the letters above the keys. You will know the □ key has been engaged when the cursor on the screen turns into a blinking ¬. After pressing the □ key you should press the under where the letter appears. As an example if you wish to type the letter E on an 83 or 84, press □ □ (because E is above □. To get the same letter E on an 89, press □ε.

Note: If you have a sequence of letters to type, you will want to press ψ □. This will engage the colored function above the □ key which is the A-LOCK function. It locks the calculator into the Alpha mode, so that you can repeatedly press keys and get the alpha character for each. Otherwise, you would have to press □ before each letter. Press □ again to release the calculator from the A-LOCK mode.

Some general keyboard patterns and important keys

1. The top row on 83's and 84's is for plotting and graphing. On 89's these functions are also on the top row, but are accessed by preceding the desired function with ∞.

2. The second row from the top has the important QUIT function (ψ ζ on 83's and 84's, ψN on 89's). On 83's and 84's it also contains the keys useful for editing ({, ψ { (INS), |, ~, } and □). INS and DEL on 89's are both combination commands: INS is 20 and DEL is ∞0.

3. The □ key in the first column on 83's and 84's leads to a set of menus of mathematical functions. Several other mathematical functions (like ϒ) have keys in the first column. On a TI-89, 2ζ leads to the Math menu.

4. The keys for arithmetic operations are in the rightmost column (∞ ↓ ≠ ℘).

Note: When displaying input, the ∞ shows as /, and the ↓ shows as *. On both 89 models, when the command is transferred to the display area the * is replaced with a · and division looks like a fraction.

5. The □ key will be basic to this course. Submenus from this allow editing of lists, computation of statistics, and calculations for confidence intervals and statistical tests. On 89's with the statistics application, one starts the application using the key sequences ∞O and selecting the Statistics application. On the 89 Titanium, quit the current application (ψN) and locate the `Stats/ListEd` application, and press ⊆ to start the application. On 83's and 84's the second function of the □ key is `LIST`. This key and its submenus allow one to access named lists and perform list operations and mathematics.

6. The □ key on 83's and 84's allows one to access named variables. On TI-89's this key is 2| which is named °, and is used to access names of lists and variables.

7. 2□ calls the distributions (`Distr`) menu. This is used for many probability calculations. To get this menu on a TI-89, press □ from within the `Stats/ListEd` application.

8. The ′ key is located directly above the ← key on 83's and 84's; on 89's it is above the o key. It is used quite often for grouping and separating parameters of commands.

9. The ↵ key is used for storing values. It is located near the bottom left of the keyboard directly above the ⊃ key on all the calculators. It appears as a ! on the display screen.

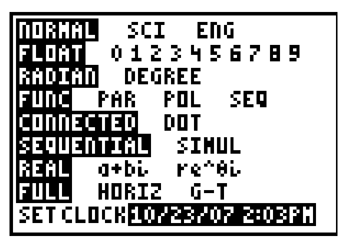
10. The ⊂ key on the bottom row (to the left of ⊆) is the key used to denote *negative* numbers. It differs from the subtraction key ≠.

Note: The ⊂ shows as ⁻ on the screen, smaller and higher than the subtraction sign.

11. On TI-89 models, switching back and forth between two apps is easiest done by pressing 2O.

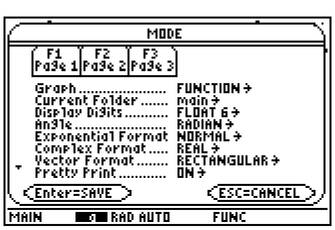
12. The ∀ key on TI-89 models switches the application from wherever you are to the home screen immediately.

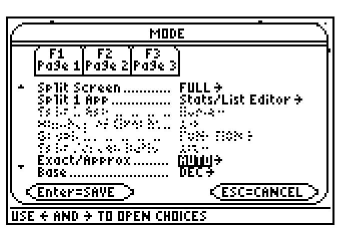
0.3 Setting the Mode

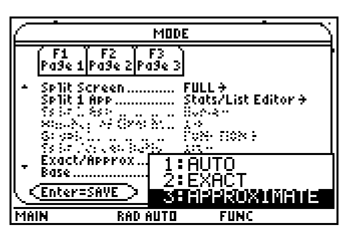


If your answers do not show as many decimal places as the ones shown, or if you have difficulty matching any other output, check your `MODE` settings. On an 83 or 84, press the ζ key (second row, second column). If your calculator has been used previously by you or someone else the highlighted choices may differ. If your screen has different highlighted choices use the } and □ keys to go to each row with a different choice and press ⊆ when the blinking cursor is on the first choice in each row. This will move the highlight to the first choice in each row. Continue until your screen looks like the one at right. Press ψ ζ (`QUIT`) to return to the Home Screen.

On TI-89's, the default is to give "exact" answers. For statistical calculations, we will want decimal approximations. To set this, press 3. Press □ to proceed to the second page of settings, then arrow to Exact/Approx and use the right and down arrows to change the setting to `3:Approximate`. Press ÷ to complete the set-up. The sequence is shown below.







0.4 Screen Contrast and Battery Check

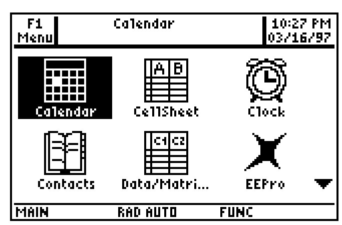
To increase the contrast, press and release the ψ key and hold down the } key. You will see the contrast increasing. There will be a number in the upper-left corner of the screen that increases from 0 (lightest) to 9 (darkest).

To decrease the contrast, press and release the ψ key and hold down the □ key. You will see the contrast decreasing. The number in the upper-left corner of the screen will decrease as you hold. The lightest setting may appear as a blank screen. If this occurs, simply follow the instructions for increasing the contrast, and your display will reappear.

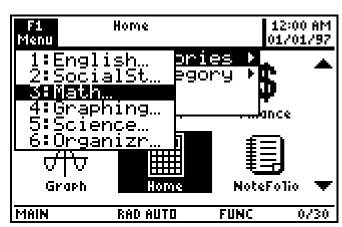
When the batteries are low, the display begins to dim (especially during calculations) and you must adjust to a higher contrast setting than you normally use. If you have to set the contrast setting to 9, you will soon need to replace the four AAA batteries. With newer versions of the operating system, your calculator will display a low-battery message to warn you when it is time to change the batteries. After you change batteries, you will need to readjust your contrast as explained above.

Note: It is important to turn off your calculator and change the batteries as soon as you see the "low battery" message in order to avoid loss of your data or corruption of calculator memory. Change batteries as quickly as possible. Failure to do so may result in the calculator resetting memory to factory defaults (losing any data or options which have been set).

0.5 The TI-89 Titanium

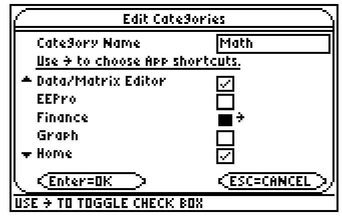


On the TI-89 Titanium, most important functions that on other calculators are accessed by keystrokes, are applications (`Apps`). When the calculator is first turned on, you will be presented with a graphical menu of these applications, as at right. Paging through the screen to find the one you want can be tiresome and time consuming. There is a way to customize this screen so that you only see those applications you want to see.



Press □. Press the right arrow key to expand menu selection `1:Edit Categories`. You will be presented with a list of possible categories. Press ♠ to select option `3:Math`.

On this screen, use the down arrow to page through the list of applications. When you find one you want to be displayed, press the right arrow key to place a checkmark in the box. The screen at right shows that the `Data/Matrix Editor` and the `Home` screen have been selected. For this statistics course, you will want these applications, along with the `Stats/List Editor` and `Y=` applications. Press ÷ when you have finished making your selections.



On this calculator, pressing 2N (`Quit`) will return you to the applications selection screen.

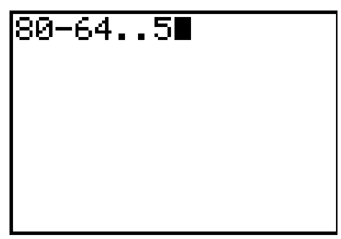
0.6 Home Screen Calculations

The following example illustrates some techniques which will be useful in performing home screen computations. We also point out the importance of correctly using parentheses in calculations.

Example 0.1: $\dfrac{70-64.5}{2.7}$

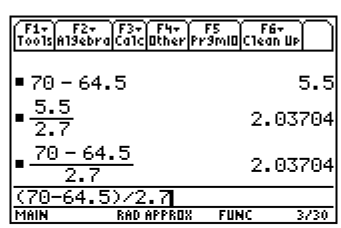
We will calculate the value in two ways. In doing so, we will intentionally make a mistake to show you how to correct errors using the { key. We also discuss the `Ans` and `Last Entry` features.

Type 80-64..5 (two intentional mistakes). To correct these, use the | cursor key to move backward until your cursor is blinking on one of the double decimal points. Press { (on an 89, either ∞0 or position the cursor to the right of the character to be deleted and press 0) and the duplicate decimal point will be deleted. Now press | until the cursor is blinking on the 8. Type a 7, and it will replace the incorrect 8. On an 89, move the cursor to the right of the error, press 0 and then type the correct 8. Press ⊆ for the numerator difference of 5.5 as shown in the top of the screen below.



Press ∞. (Note that "`Ans/`" appears on the screen). Type 2.7 and press ⊆ for the result of ˉ2.037.
Note: `Ans` represents the last result of a calculation that was displayed alone and right-justified on the Home screen. Pressing ∞ without first typing a value called for something to be divided, so `Ans` was supplied.

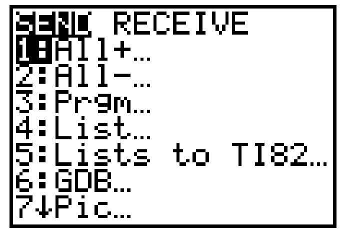
To do the calculation in one step, press ψ ⊆. This calls the "last entry" to the screen (in this case Ans/2.7). Press ψ ⊆ again to get back to 70-64.5. Press the } key to move to the front of the line. On an 89, press 2A.



On an 83 or 84, you will need to press ψ { (for INS); 89's are always in insert mode. You will see a blinking underline cursor. Type ≤ to insert a left parenthesis before the 7. Press □ (2B on an 89) to jump to the end of the line. Type ⁄ ∞ 2.7 to see the result. Press ⊆ for the same result as before.

0.7 Sharing Data

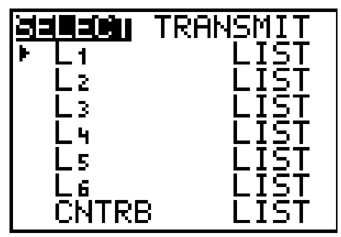
Sharing data between calculators (TI-83/84)



Data and programs may be shared between calculators using the communications cable which is supplied. The TI-83 and 84 series can share any TI-83/84 information with the exception of flash applications and their associated variables.

On the TI-83+, the I/O port is at the base of the calculator. On the TI-84+, you can use either the USB port or the I/O port on the top to link to another 84 series calculator. To link to an 83 series, you must use the I/O port. Connect the appropriate cable to the ports. On both calculators, press ψ□ to activate the LINK menu.

On the receiving calculator, press ~ to highlight RECEIVE, then press ⊆. The calculator will display the message "Waiting..." The rolling cursor on the upper right indicates the calculator is working.

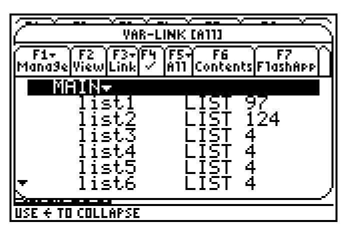


On the sending calculator, use the arrow keys to select the type of information to send. For example sending lists, either arrow to 4:List and press ⊆ or press ∂. The screen at right will be shown.

To select items to send, move the cursor to the item, press ⊆ to select it. After selecting all items to send, press ~ to highlight TRANSMIT, then press ⊆.

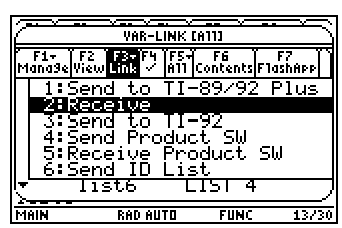
Sorry, TI-83's and 84's cannot communicate with TI-89's.

Sharing data between calculators (TI-89 series)



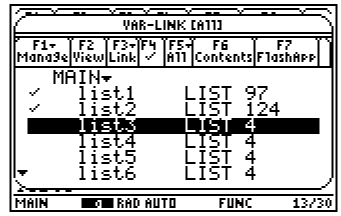
Data and programs may be shared between calculators using the communications cable which is supplied. The TI-89 can only communicate with the other TI-89's and TI-92s.

Connect the supplied cable to the port at the base of each calculator. On both calculators press ψ| to activate the VAR-LINK menu.

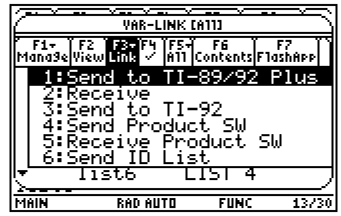


On the receiving calculator, press □ to select Link, Δ to highlight Receive, then press ⊆. The screen reverts to main VAR-LINK Menu with a "Waiting to receive" message at the bottom.

On the sending calculator, use the arrow keys to select the item to send, then press □ to "check" the item. The example at right will send `list1` and `list2`.



Now press □ to select Link, then ⊆ to select menu choice `1:Send to TI-89/92Plus` which is highlighted by default.



An analogous procedure can be used to send applications between calculators. Applications (such as the Statistics with List Editor) are selected from the □ `FlashApp` menu (press ψ□ for □.)

Sharing data between the calculator and a computer

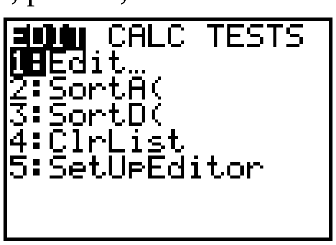
Data lists, screen shots, and programs may be shared between the calculator and either Microsoft Windows or Macintosh computers using a special cable and TI-Connect or software. This can be helpful since all data sets in the text are included on the CD and companion website for your text. Cables for the TI-83+ are available for either serial or USB ports; they can be found through many outlets such as OfficeMax, or Amazon.com. The software can be downloaded free through the Texas Instruments website at education.ti.com. The needed USB cable and computer software are included with the TI-84 and -89 models. See the section on using supplied data sets for more information.

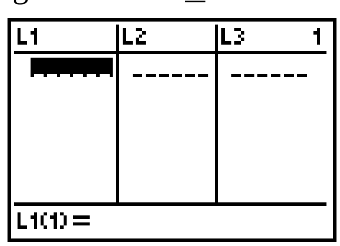
0.8 Working with Lists

The basic building blocks of any statistical analysis are lists of data. Before doing any statistics plot or analysis the data must be entered into the calculator. The calculator has six lists available in the statistics editor; these are `L1` through `L6` (`list1` through `list6` on an 89). Other lists can be added if desired. The number of lists is only limited by the memory size.

The statistics editor

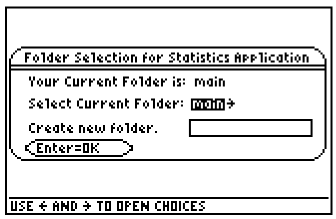
On a TI-83/84, press □, `1:Edit...` will be highlighted. Press ⊆ to select this function.

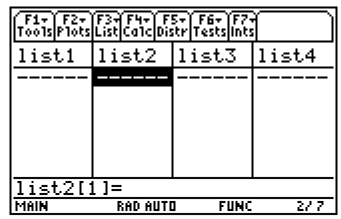




To enter data, simply use the right or left arrows to select a list, then type the entries in the list, following each value with ⊆. Note that it's not necessary to type any trailing zeros. They won't even be seen unless decimal places (found with the ζ key) have been set to some fixed number.

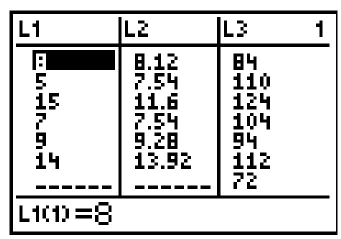
On a TI-89, press ∞O followed by the selection of the Statistics/List Editor Flash Application followed by ÷. On the Titanium edition, select the Statistics/List Editor from the main application menu screen. If this is the first time the editor has been accessed since the calculator has been turned on, you will be prompted for a data folder as in the middle screen. The default folder is `main`. Press ⊆ to select main as the current folder, or press B to allow a new folder to be created. To enter a new folder name, arrow to the entry block and type the name of the new folder. Your lists will be stored in the new folder and it will be set to default. To change folders, press B to select a folder. Then press ⊆ to proceed to the list editor. If the editor has been used since the calculator has been turned on, pressing ⊆ to select the application will automatically open the editor.





One word of advice: Most lists of data in texts are entered across the page in order to save space. Don't think that just because there are four (or more!) columns of data they belong in four (or more) lists. Data that belongs to a single variable always belongs in a single list.

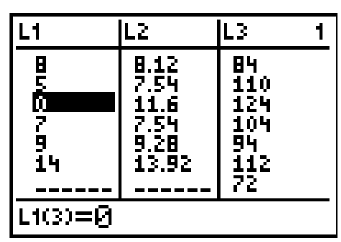
Entering data into the STAT editor



With the cursor at the first row of L1, type 8 and press ⊆. The cursor moves down one row. Type 5 followed by ⊆ and the 5 will be pasted into the second row of the list. Continue with 15, 7, 9, and 14 as seen at right.

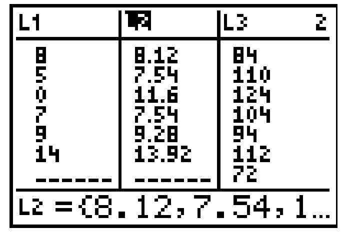
Correcting mistakes with DEL and INS

In the screen above, we can delete the 15 by using the } key until it is highlighted and then pressing { (∞0 on an 89).



To insert a 12 above the 7 move the cursor to the 7 then press ψ { (20 on an 89) (to choose INS or insert mode). Note a 0 was inserted where you wanted the 12 to go. Just type over the place-holding 0 with the value you want.

Clearing lists without leaving the STAT editor



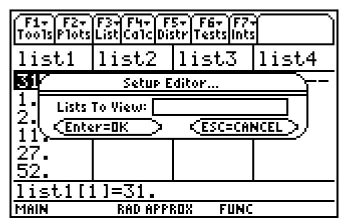
Suppose you wish to clear a list, say L2, while you are still in the STAT Editor. You should use the cursor to highlight the name of the list at the top. With the name highlighted, press □ and you will see this. Press ⊆ and the contents of the list will be cleared. *Make sure not to press* { or the list will be deleted entirely and you will have to use SetUpEditor as described below to retrieve it.

Deleting a list from the STAT editor

If you wish to delete a list from your STAT Editor, simply highlight the list name and press {. The name and the data are gone from the Editor but not from the memory. To recover a list inadvertently deleted, use SetUpEditor as described below.

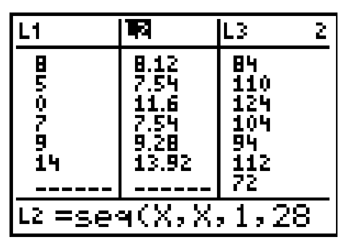
SetUpEditor

Setting up the editor will remove unwanted lists from view. It also will recover lists that have inadvertently been deleted. On an 83 or 84, if you want the STAT Editor to be restored to its original condition (with lists L1 to L6 only), press □ • ⊆. Often students find this necessary because they have inadvertently deleted one of the original lists.

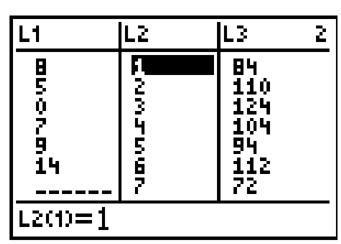


On an 89, in the Statistics Editor, press □ (Tools), then select option 3:Setup Editor. You will see the screen at right. Leave the box empty and press ⊆ to return to the six default lists.

Generating a sequence of numbers in a list



From time to time one may want to enter a list of sequenced values (years for example in making a time-series plot). It is certainly possible (but tedious) to enter the entire sequence just as one would enter normal data. There is an easier option, however. Use cursor control keys to highlight L1 in the top line. Press ψ □ ~ •. You are choosing the LIST menu and then choosing the OPS submenu. From the OPS submenu you choose option 5 which is seq(.



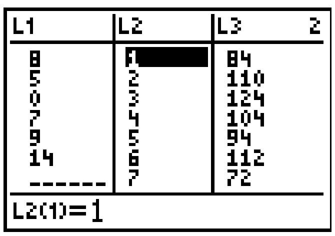
This has been pasted onto the bottom line of the screen. Type in the rest so that you have seq(X,X,1,28. Press ⊆ and the sequence of integers from 1 to 28 will be pasted into L2 as in my screen. Find the X on the □ key on an 83 or 84, on 89's it has its own key. Notice that it was not necessary to clear the list first.

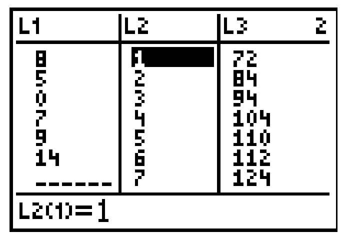
On an 89, the procedure is analogous, but access the `LIST OPS` menu by pressing □♥, then select option 5.

Note: To quickly check the values on a multi-screen list you can press the green □ key followed by either the } or □ key. This will allow you to jump up or down from one page (screen) to another. The green arrows on the keyboard near the } and □ keys are there to remind you of this capability. On a TI-89, instead of the □ key, press ∞.

Sorting lists (TI-83/84)

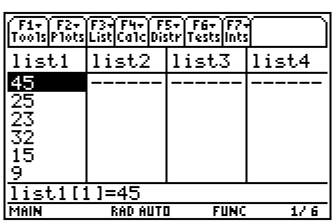
Lists may be sorted in either ascending (smallest to largest value) or descending order. The resulting list will replace the original list. On the main □ menu select either 2:SortA(or 3:SortD(. The command will be transferred to the home screen. Enter the name of the list to be sorted (ψn where n is the number of the list). Execute the command by pressing ⊆. The example below sorts list `L3`.

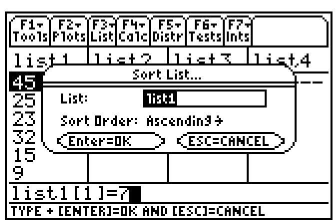


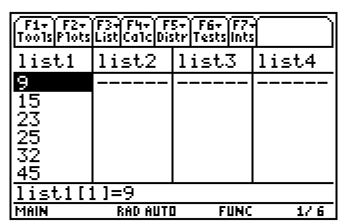


Sorting lists (TI-89)

While in the List Editor, press □ for the `List` menu, press Δ followed by ⊆ or ♥ to select List Ops, then press ⊆ to Select 1:Sort List, ψ≠ (`Var-Link`) and use the arrows to select the list to be sorted. Press the right arrow if necessary to change the sort order (use B to activate the menu choices). Press ⊆ to carry out the command.

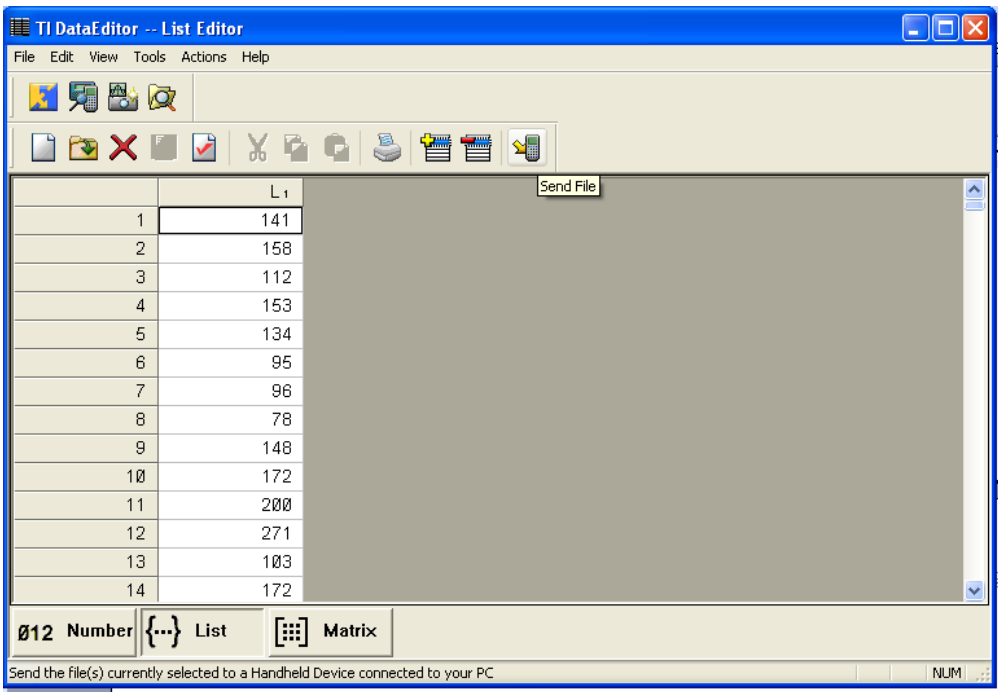




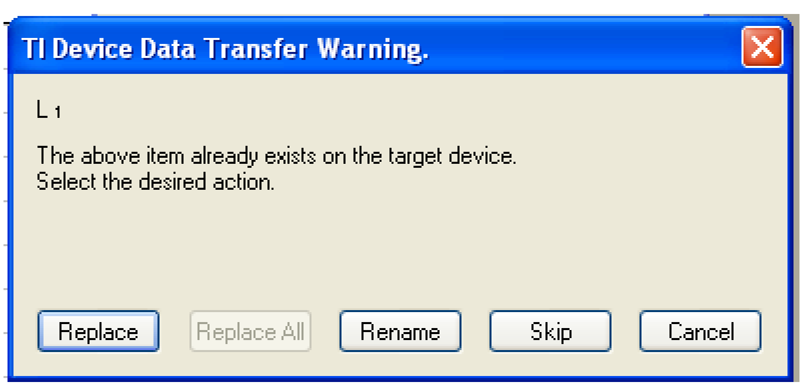


0.9 Using the Supplied Datasets

If you have the additional Connect cable and TI-Connect software, you can load the datasets for all examples, exercises, and tables in the text easily without needing to retype all the numbers. Datasets are also on the text's accompanying web site. If you're using the CD, insert it in the drive — it should automatically open (you may get a warning about "dangerous content" — and simply click to continue. Click on the datasets link, then select PC or Mac TI-83 format. You may at this point want to copy the folder to your desktop for easier future use. Click on the folder to open it. Select the file you wish to use — names beginning with eg are examples, ex are exercises, and ta are tables. In the example below, I have opened a file from Chapter 1 of the *Basic Practice of Statistics* for exercise 11 about blood glucose levels. Notice that its heading is `L1`.



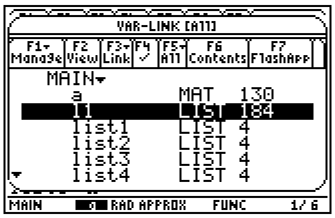
To transfer the file to the calculator, press the Send File icon (at top right). The transfer should start automatically, after the software locates your calculator. You may see a warning like the one below:

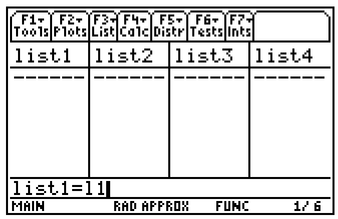


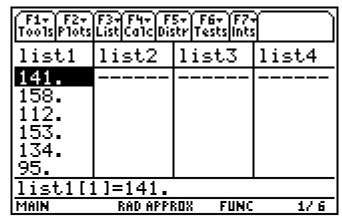
Press ÷ to replace the current contents of list `L1`. If you want to save the current contents of `L1`, click to rename and send the data to another list, or select Cancel to abort the transfer.

Additional TI-89 step

With these calculators, the lists will default to being named `L1`, `L2`, etc in the `main` folder. They can be used (and accessed) with ° just like the default lists for graphs and other calculations. If you want to be able to look at them in the statistics editor, highlight the desired statistics editor list name, then use 2| (°) and move the cursor to the desired list name. Press ÷ to select the name, then ÷ again to fill the list.







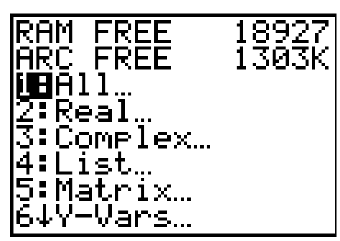
0.10 Memory Management

Too many applications loaded or lists in active memory can overload the calculator. Just as a computer disk can be filled up, so can memory on the calculator. The TI calculators have two types of memory – RAM and archival. If you use the applications supplied on the CD (on a TI-83/84), these are loaded into archival memory. Active lists are in RAM.

TI-83/84 procedure

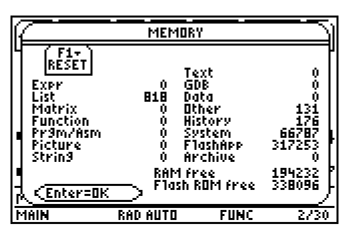
To find out the current free memory status, press ψ℘ (MEM) and select option 2:Mem Mgmt/Del.

The screen at right shows my calculator currently has 18,927 bytes (characters) of free RAM and 1303Kilobytes of free archival memory.

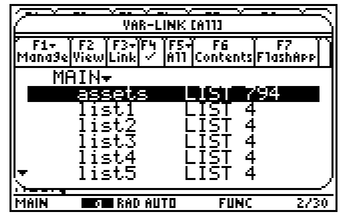


If you need to free some memory, decide the type. If you want to delete some lists, for example, select 4:List. Move the cursor to the lists you wish to delete and press { for each one. This can also be done for any applications you no longer need from previous chapters, but use choice 5:Archive to access the list of archived applications.

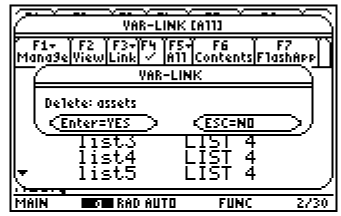
TI-89 series procedure



To find out the current free memory status, press 2{. The screen at right shows I currently have 194,232 free bytes of RAM and 338096 free bytes of Flash memory free. Pressing □ here will reset RAM, Flash, or all memory to either totally blank or factory default settings. I do not recommend either of these options under normal circumstances. Resetting Flash, for example, would erase the Statistics Flash application, which would then need to be reloaded.



To delete lists that are no longer needed, press 2| (VAR-LINK). Move the cursor to highlight the list to be deleted, then press □. Press ÷ to select option 1:Delete.

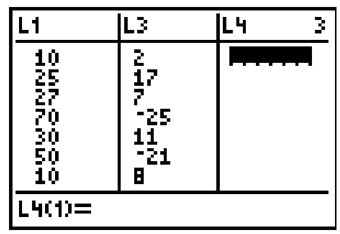


You will be prompted to verify that the selected item is to be deleted. Press ÷ to confirm the deletion, or N to cancel. You can continue this process to delete all unneeded lists.

0.11 Common Errors

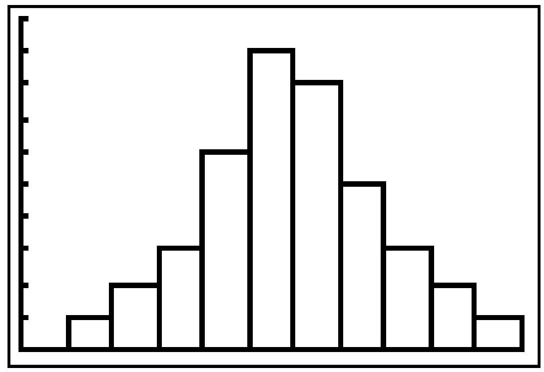
Why is my list missing?

By far the most common error, aside from typographical errors is improper deletion of lists. When lists seem to be "missing" the user has pressed { rather than □ in attempting to erase a list. Believe it or not, the data and the list are still in memory. To reclaim the missing list press □ and select choice 5:SetUpEditor followed by ⊆ to execute the command. Upon return to the Editor, the missing list will be displayed.



CHAPTER

1



Looking at Data—Distributions

Introduction

In this chapter, we use the TI calculators to view data sets. We first show how to make bar graphs, histograms, and time plots. Then we use the calculator to compute basic statistics, such as the mean, median, and standard deviation, and show how to view data further with boxplots. Lastly, we use the TI-83 Plus for calculations involving normal distributions.

1.1 Displaying Distributions with Graphs

We start by using the calculator to graph data sets. In this section, we will use the `STAT EDIT` screen to enter data into lists and use the `STAT PLOT` menu to create bar charts, histograms, and time plots.

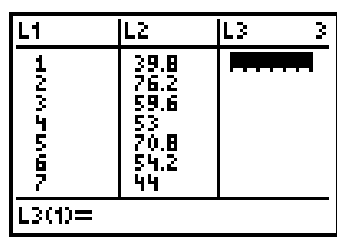
Throughout the manual, we will be working with data that is entered into lists `L1` through `L6` on the TI-83/84 (`list1` through `list6` on an 89 model). These lists can be found in the `STAT EDIT` screen. A list should be cleared before entering new data into it.

Example 1.1 Women's Degrees: A Bar Graph of Categorical Data. TI calculators cannot make true bar graphs, since all quantities entered on the `STAT EDIT` screen must be numeric. Also, proper bar graphs should have the bars separated. These calculators can however, guide you in creating a bar graph. Here are the percents of women among students seeking various graduate and professional degrees during the 1999–2000 academic year.

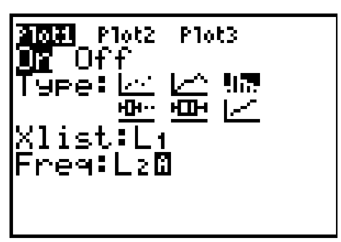
Degree	**Percent female**
MBA	39.8
MAE	76.2
Other MA	59.6
Other MS	53.0
Ed.D.	70.8
Other Ph.D.	54.2
MD	44.0
Law	50.2
Theology	20.2

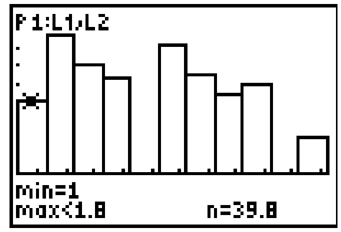
We would like first to make a bar graph of the data.

TI-83/84 Solution. First, label the nine categories as 1–9 and enter these values into list `L1`, then enter the percents into list `L2`.

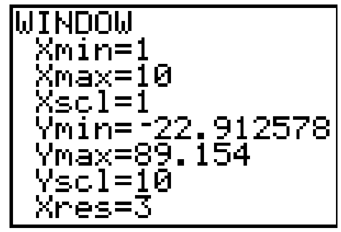


Press 2o (`STAT PLOT`). Press ⊆ to select `Plot1`. Our plot will use the histogram plot type, and we will put the category numbers that were entered in `L1` on the *X* axis, and the percentages on the *Y* axis, so our plot definition screen is at right. At this point, it is good practice to check that no functions are defined on the o screen itself (if so, press □ to erase them) and that the other two plots are "turned off." *TI calculators try to plot everything they know of at once. This can result in error messages and other junk!*

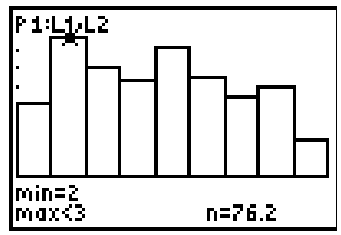




For most plots, pressing θ→ (`ZoomStat`) will suffice. However, for histograms and bar charts this usually does not give a reasonable picture. At right is my `ZoomStat` graph. I have pressed ρ to display the first bar. Notice that the bar's x-values go from 1 to 1.8. We need intervals of width 1 for our bar chart.

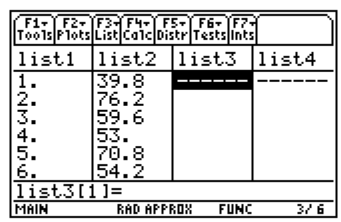


Press π. This allows us to control the values that display on a graph. If you look at the graph above, notice that the endpoint of the first interval (1.8) is not included. We need intervals of length 1. All we need to do here is change `Xmax` (the largest X value that displays to 10 (for consistency) and the bar width `Xscl` to 1. Don't worry about changing `Ymin` and `Ymax`. `Ymin` is negative so that the ρ information does not obscure the graph.

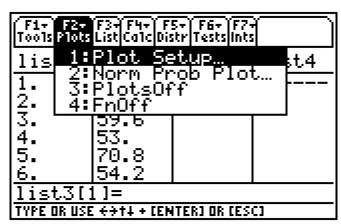


Press σ to display the rescaled graph. *When you have changed a window, always press σ. Pressing θ→ will revert to the default scaling.* Here is the completed graph. If you are copying this onto paper, don't forget to give proper labels to the categories and separate the bars.

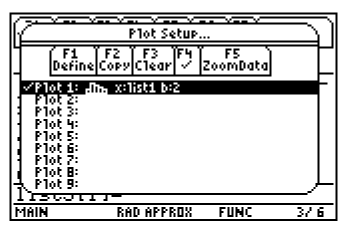
TI-89 Solution



Press ∞O, move the cursor to highlight the `Statistics/List Editor` Application, and press ⊆ to get to the list editor. Here I have entered the data into `list1` (the category numbers) and `list2` (the percents).

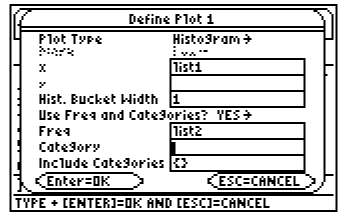


Plot Definition on TI-89 calculators is somewhat different from the 83/84models. Press □ (`Plots`) then ⊆ to select option `1:Plot Setup`.

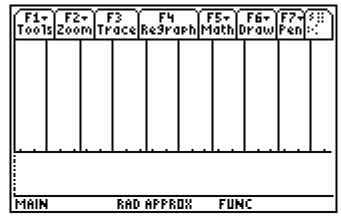


TI-83/84 calculators can have up to 3 plots defined at once, the 89 can have as many as 9. Press ⊆ to select Plot1 to be defined. *TI calculators try to graph everything they possibly can at once.* This can lead to error messages and other types of "junk" on your graph. If any other plots have been defined (and you wish to preserve them) move the highlight to the plot line and press □ to uncheck them so they won't try to be displayed. Similarly, you should check that the # function editor (press ∞□) is cleared.

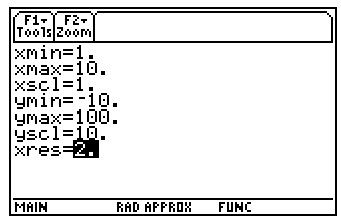
Press □ to select `Plot 1` to be defined. The plot type is a histogram, our category “labels” were stored in `list1`. TI-89 calculators must have the bar (bucket) with explicitly defined. Here, I have set it to one. Since we have another list of the frequencies of each category, I have set `Use Freq and Categories` to `YES` and specified that the frequencies are in `list2`. Remember, press □| (°) to locate the list names. Press ⊆ to finish the plot definition. You will be returned to the `Plot Setup` screen.



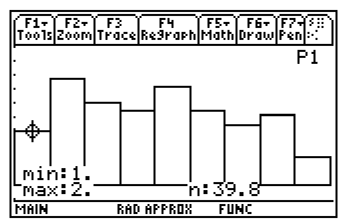
If I press □ (`ZoomData`) at this point, this is the graph I see. TI-89 calculators (especially on bar charts like these and histograms also) need some extra help in setting their windows because of the tabs at the top of the screen. Press ∞□ (`WINDOW`).



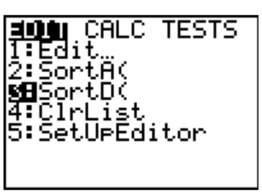
At right, I’ve changed the window settings. The minimum (`xmin`) and maximum (`xmax`) *x*-axis values to display have been set to 1 and 10 because our lowest category number was 1 and the bars will have width 1. `ymin` and `ymax` have been changed as well to allow the full bar graph to display. We need a “roomy” `ymax` so the tabs don’t hide the top of the plot.

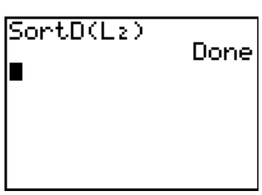


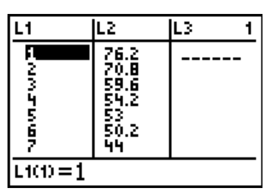
Press ∞□ to display the graph. Press □ (`Trace`) to display the values shown by each bar. The upper end of each bar (2 in the first bar highlighted) is not included in the bar. This will become important in histograms. In copying your graph onto paper, don’t forget to use the proper category labels, and to separate the bars.

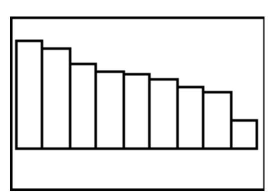


Example 1.2 Women’s Degrees: A Pareto Chart of Categorical Data. Next, we make a *Pareto* chart, a bar chart with the bars ordered by height. To do so, we simply use the `SortD(` command from the `STAT EDIT` screen to sort the data in list `L2` into descending order, then regraph using the same window settings as before by pressing σ. On a TI-89, find the sort command on the □ (`List`) option `2:Ops` menu.







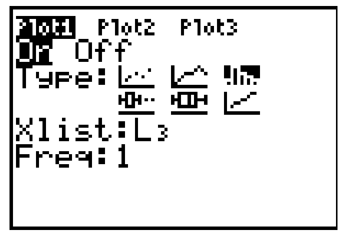


If copying this onto paper, be careful to match the bars with their correct labels (degrees)!
Example 1.3 The Density of the Earth: Making a Histogram. Make a histogram of Cavendish's measurements of the density of the Earth.

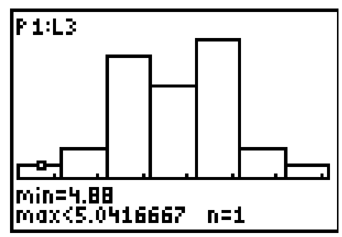
5.50	5.61	4.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

Solution. First, we enter the data into a list. Here we have entered the data in list L3.

First, define the plot as at right. In the bar chart above, we had a separate list of the frequencies of each type of degree, now each value represents one observation. To change the Freq: to 1, press □ to change the cursor from ¬ mode, or your typed ℵ will be a Y.

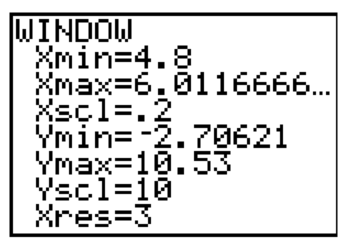


Press θ→ to display the default graph. Then press ρ to display the actual bars and counts. As we saw with bar charts, the calculator needs some help setting a reasonable window. The length of this first bar (and succeeding bars) does not make logical sense. Your instructor (or text) may indicate what intervals should be used on any given problem, but in general we want these to make some type of logical sense so as to clearly convey the information, and use a reasonable number of bars.

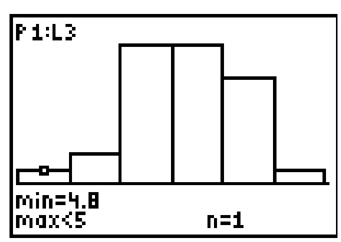


How many bars we need is matter of judgment, but usually we want between 5 and 20 bars. The old "rule of thumb" was to start by thinking of *n* (the sample size) divided by 5. Here, our *n* is 29, so about 6 bars would be reasonable. Looking at the minimum value, we need to make our minimum somewhat smaller, and find a bar width that is intuitive.

Here, I have chosen to set Xmin to 4.8 and use a bar width (Xscl) of .2. You normally will not need to change the *Y* settings unless a scale change makes the top of a bar disappear; in that case, increase Ymax.



Here is my final calculator graph. To copy it onto paper, use ρ and the right arrow key to move across the bars. Each interval on the *x*-axis goes from the min to the max shown and contains the number of observations indicated by n= at the lower right. *Be careful the x-axis is a number line; do not display the values as "category labels!"*



This distribution is unimodal (one-peaked) and possibly a little left-skewed, due to the two short bars at left.

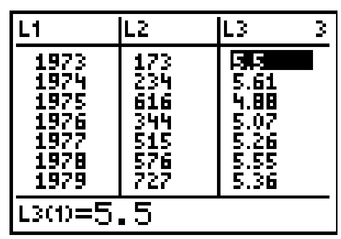
If you are using a TI-89, follow the instructions given above for creating a bar chart. You will need to set the bucket width on the plot definition screen to something reasonable, and modify the `WINDOW` settings to get a final graph.

Example 1.4 Rock Sole Recruitment: A Time Plot. The next exercise demonstrates how to use a time plot to view data observations that are made over a period of time.

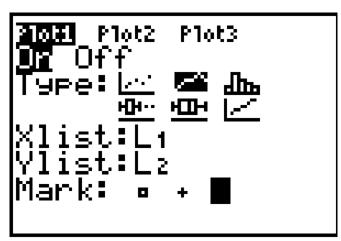
Here are data on the recruitment (in millions) of new fish to the rock sole population in the Bering Sea between 1973 and 2000. Make a time plot of the recruitment.

Year	Recruitment	Year	Recruitment	Year	Recruitment	Year	Recruitment
1973	173	1980	1411	1987	4700	1994	505
1974	234	1981	1431	1988	1702	1995	304
1975	616	1982	1250	1989	1119	1996	425
1976	344	1983	2246	1990	2407	1997	214
1977	515	1984	1793	1991	1049	1998	385
1978	576	1985	1793	1992	505	1999	445
1979	727	1986	2809	1993	998	2000	676

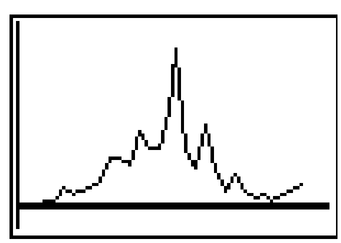
Solution. Enter the data into two lists. For the years, it is easiest to use the `seq(` function from the `LIST` menu as described on page 10. Here, I have used lists `L1` and `L2`.



On the `STAT PLOTS` menu, select `Plot1` and define it as at right. This second plot type is a connected scatter plot. If you are using a TI-89, this is the `xy-line` plot type. The years go on the *x*-axis of a time plot, so `L1` is the `Xlist` variable, and `L2` (the rockfish) is the `Ylist`. The last data point mark (the single pixel) is recommended for this type of plot, since we really are most interested in seeing the pattern of variation and not the individual points.

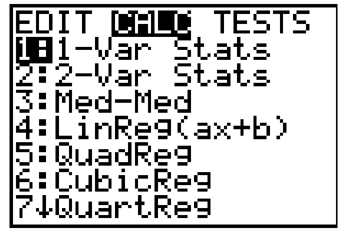


Press θ→ to display the graph. There is no need to adjust the window for this type of plot. We clearly see the rockfish recruitment was small from 1973 but increased fairly steadily until recruitment peaked in 1987. The fish population thereafter seems to have collapsed. A fisheries scientist would be interested in finding out possible reasons to explain



this phenomenon.

1.2 Describing Distributions with Numbers

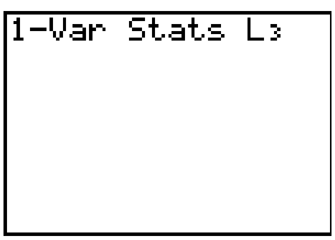


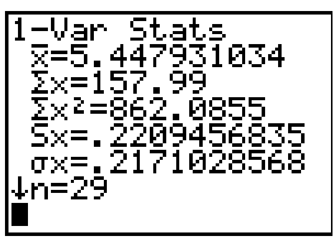
In this section, we will use the `1-Var Stats` command from the `STAT CALC` menu to compute the various statistics of a data set including the mean, standard deviation, and five-number summary. We also will use boxplots and modified boxplots to view these statistics as another way to picture a distribution.

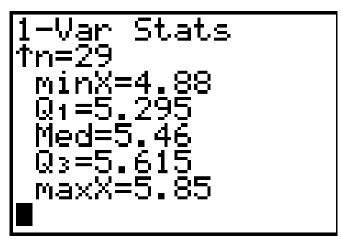
Example 1.5 The Density of the Earth: Finding Summary Statistics. Find $\bar{x}$ and s for Cavendish's data used in Example 1.3 on page 19. Also give the five-number summary and create a boxplot to view the spread.

5.50	5.61	4.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

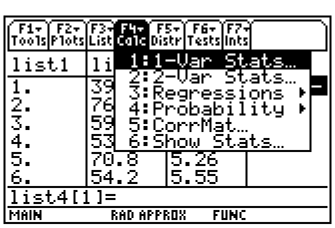
TI-83/84 Solution. Since we still have this data in list `L3`, we compute the statistics by pressing □ then ~ to `CALC`, and finally ⊆ since option `1:1-Var Stats` is highlighted. This transfers the shell of the command to the home screen. We finish the command by telling the calculator which list we are interested in. In this case `L3` will be entered as ψℜ. *The calculator defaults to calculating statistics for* `L1`. *It is good practice to always specify the list you are working with.* Press ⊆ to perform the calculations. The values of $\bar{x}$ and s are then displayed. Scroll down to see the five-number summary.

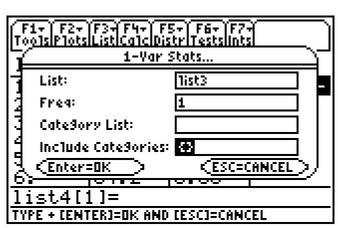


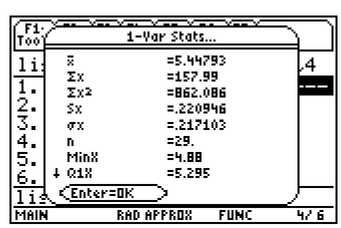




TI-89 Solution. The data have been entered into `list3`. Press □ for the `Calc` Menu. Option 1:1-Var Stats is highlighted, so press ⊆. Locate the list name for which you want Statistics on the ° screen, highlight the list name and press ⊆ to move its name onto the definition screen. Finally, press ⊆ to execute the command. As with the other models, use the down arrow key to complete display of the five-number summary.



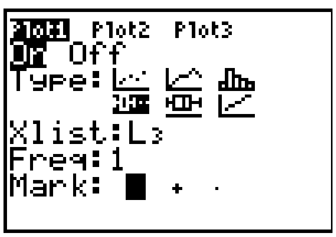


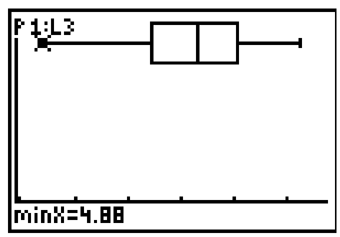


Notice that many of the values have numerous decimal places given. *Your instructor will most likely specify a rounding rule, but the usual rule is to report one more significant digit than was in the original data*. Also, two standard deviation values are given. The first, `Sx`, is the standard deviation that is denoted by s in the text. Thus, we obtain $\bar{x}$ = 5.448 and s = 0.221 with a five-number summary of 4.88 – 5.295 – 5.46 – 5.615 – 5.85.

Example 1.6 The Density of the Earth: A Boxplot. TI calculators will do two different boxplots: the original that simply uses the five-number summary, and the modified boxplot that uses the 1.5*IQR criteria to identify outliers. We always recommend using the latter, since if a long tail appears, we usually want to know if it's real, or due to outliers. Also, certain distributions that contain outliers may not look like they exist due to scaling.

To make a boxplot of data in a single list, define the plot as below. If you are using a TI-89, use plot type `5:Mod Box Plot`. For this type of plot, we do not recommend the single pixel mark for outliers; it is too difficult to see. Select either the box or cross. Press θ→ (□ for `ZoomData` on a TI-89) to display the plot. There is no need to adjust windows. As always, you can use ρ to locate the values on the plot.



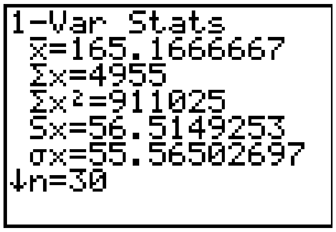


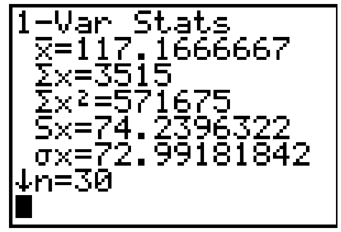
Here we see a longer, left tail to the distribution, indicating more spread of the lowest values. However, since we asked the calculator to identify any outliers, we know there are none as none are indicated.

Example 1.7 Study Time: Side-by-Side Boxplots. The data below give the nightly study time claimed by samples of first-year college men and women. We will first compute the summary statistics for each distribution. We will also make side-by-side boxplots to compare these distributions.

Women					Men				
180	120	180	360	240	90	120	30	90	200
120	180	120	240	170	90	45	30	120	75
150	120	180	180	150	150	120	60	240	300
200	150	180	150	180	240	60	120	60	30
120	60	120	180	180	30	230	120	95	150
90	240	180	115	120	0	200	120	120	180

Solution. We enter the data sets into separate lists and then use the `1-Var Stats` command on each list. We have entered the women's data values in to `L1` and the men's into `L2`.

Women: 

Men: 

The women study longer, on average and have a smaller standard deviation than the men. Using the down arrow, we see that the median for women (175) is also higher than the median for the men (120).

Note: If we have two data sets with an equal number of measurements (like here), then we can compute the statistics of both simultaneously with the `2-Var Stats` command from the `STAT CALC` menu. In this case, enter `2-Var Stats L1,L2`. However, this command does not display the five-number summaries.

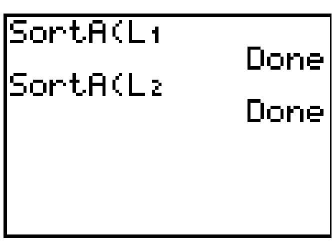
From the five-number summaries, we can compute boundaries (or fences) according to the 1.5**IQR* rule to determine outliers. In each case, we need the values of Q1 and Q3. For the women's study times, Q1 is 120 and Q3 is 180, so these boundaries are

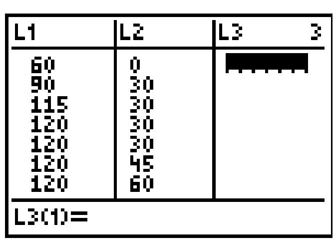
$$120 - 1.5*(180 - 120) = \mathbf{30} \quad \text{and} \quad 180 + 1.5*(180 - 120) = \mathbf{270}$$

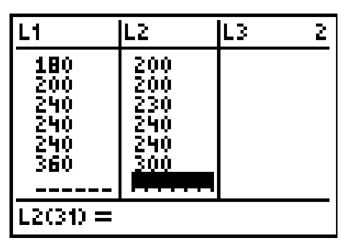
For the men's study times, Q1 is 60 and Q3 is 150, so the fences are:

$$60 - 1.5*(150 - 60) = \mathbf{-75} \quad \text{and} \quad 150 + 1.5*(150 - 60) = \mathbf{285}$$

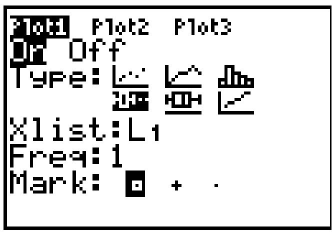
Now we can determine the suspected outliers. For the women, these outliers are any times below 30 minutes or above 270 minutes, while for the men they are any times below –75 minutes or above 285 minutes. To see these values more quickly, we can use the `SortA(` command from the `STAT EDIT` menu to sort each list into increasing order. Enter `SortA(L1` then `SortA(L2`. Because these lists have the same size, we can also enter the command `SortA(L1,L2`. In each case, there are no low outliers, but the time 360 is a high outlier for the women and the time 300 is a high outlier for the men.

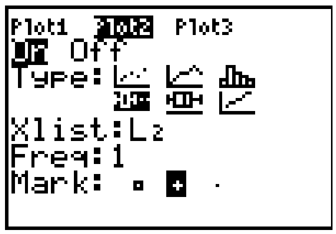


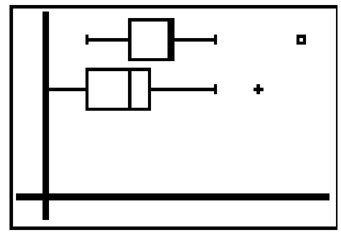




We could have found these outliers much more readily by creating side-by-side boxplots of our two lists of data. This is an exception to the one-plot-at-a-time rule. Below, I define `Plot1` to use the women's data and `Plot2` to use the men's data. The different symbol for any outliers found helps to distinguish which plot is which, but ρ also informs you which list a plot is using. Use the down arrow to move between the plots. Press θ→ to display to graphs. We can clearly see the outlier in both distributions. Notice that Q3 for the women is very close to the median.

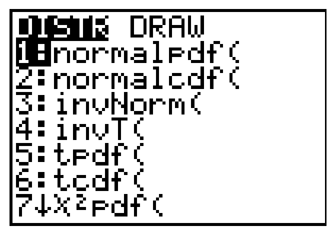






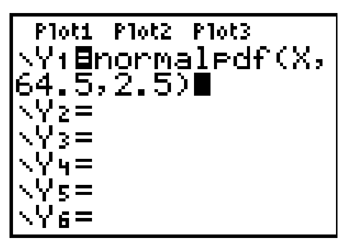
1.3 Density Curves and Normal Distributions

TI calculators have several commands in the `DISTR` menu that can be used for graphing normal distributions, computing normal probabilities, and making inverse normal calculations. In this section, we demonstrate these various functions.

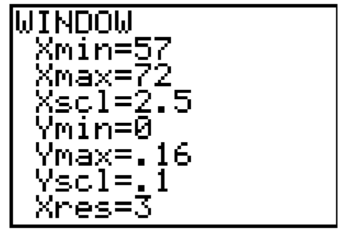


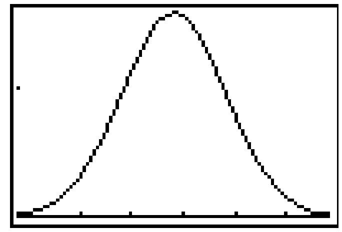
Example 1.8 Women's Heights: Plotting and Shading a Normal Distribution. The distribution of heights of young women are approximately normal with mean μ = 64.5 inches and standard deviation σ = 2.5 inches. (a) Plot a density curve for this *N*(64.5, 2.5) distribution. (b) Shade the region and compute the probability of heights that are within one standard deviation of the mean.

TI-83/84 Solution. (a) We must enter the normal density function `normalpdf(X,`μ`,`σ`)` and adjust the window settings before graphing. Press o to bring up the function definition screen. Now press ψ□ for the `DISTR` menu. Option 1 is `normalpdf`, so press ⊆ to select it and transfer the shell back to the `Y=` screen. Finish defining the function by entering the mean and standard deviation.



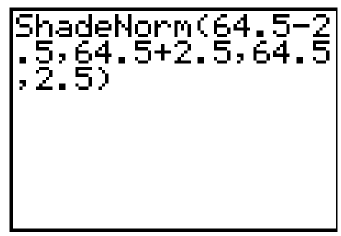
Make certain that all `STAT PLOTS` are turned off. Press ψo andmthen ∂ to turn all plots off. We will need to size the `WINDOW` for this plot. Press π. From the 68-95-99.7 Rule, we know that almost all the area under a Normal curve is within 3 standard deviations of the mean. Three standard deviations here is 3*2.5 = 7.5. We add and subtract that from the mean of 65, so have an `Xmin` of 57 and `Xmax` of 72. We have set `Xscl` to 2.5, so we will see a tick mark for each standard deviation above and below the mean. No probability distribution can be negative (we can't see negative frequencies of something!) so we set `Ymin` to 0. Setting `Ymax` for a plot like this is a little harder. The whole area under the curve is 1, so we know `Ymax` should be less than 1. Play around some to find a good value. Here, I've set `Ymax` to .16.

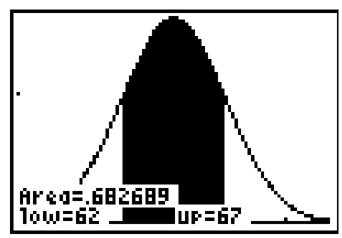




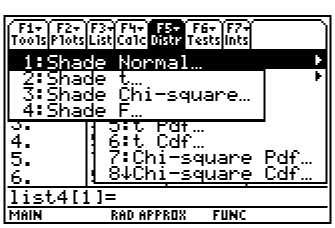
Finally, press σ to display the distribution curve.

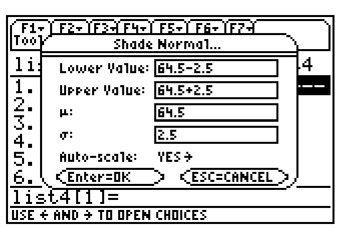
b) Return to the Home screen and press ψ□ for the DISTR menu. Press the right arrow key to DRAW. Press ⊆ to select option 1:ShadeNorm. You can figure out the heights that are one standard deviation above and below the mean explicitly before entering the command, or you can let the calculator do this for you. Here I'm letting the calculator do that for me. The parameters of the command are low-end, high-end, mean, and sigma. Be sure to separate these with commas. We press ⊆ and the desired area will be shaded on the graph created in step a. The calculator also tells us the explicit ends of the area of interest and that 68.27% of women have heights between 62" and 67".

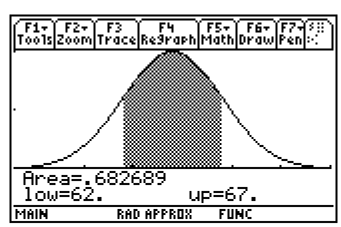




TI-89 Solution. In the Statistics/List Editor, press □ (`Distr`). The first option is Shade. Press the right arrow to see more options under this. Press ⊆ to select `Shade Normal`. Enter the lower and upper values of interest (either explicitly, or as I've done here as $\mu-\sigma$ and $\mu+\sigma$) then the values of μ and σ. Select `YES` to `Auto-scale` the plot by pressing the right arrow and moving the cursor to highlight `YES`; this saves having to figure out the window on your own.



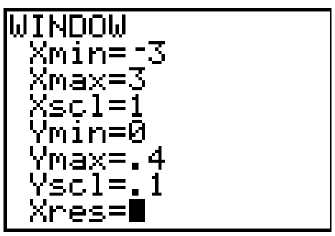


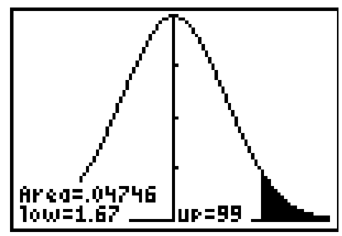


Example 1.9 The Standard Normal Distribution. Let $Z \sim N(0, 1)$ be the standard normal distribution. Shade the areas and find the proportions for the regions (a) $Z > 1.67$, (b) $-2 < Z < 1.67$.

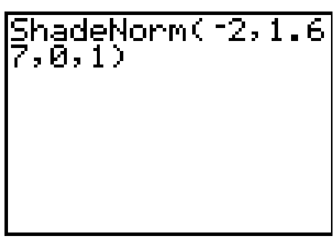
As above, if you are working with a TI-83, you will need to set the window. With a TI-89, this follows immediately from the example above — let the calculator figure out the window for you.

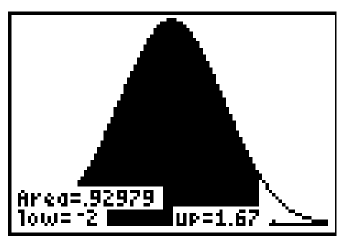
Again, since almost all the area under the curve is within 3 standard deviations of the mean, I have set `Xmin` to –3 and `Xmax` to 3 with `Xscl` 1 (the standard deviation). `Ymin` is 0, and `Ymax` is .4. This window is good for standard Normal distributions. Technically, since the curve extends to ∞ in the positive direction, the high end of interest should be ∞ (1E99 on the calculator, entered as ♦⊥oo), but practically any large number will do. There is almost no area in a normal curve more than 99 standard deviations above the mean.





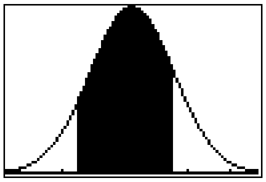
(b) Before drawing a new graph, we will need to clear the shaded part from part a, or we'll just accumulate shading. Press ψ□ (`DRAW`) and press ⊆ to execute option `1:ClrDraw`.





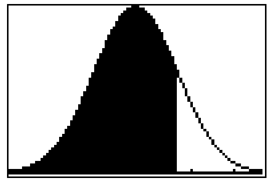
The Normal Distribution and Inverse Normal Commands

For any $N(\mu,\sigma)$ distribution X, we can find probabilities directly with the built-in `normalcdf(` command from the `DISTR` menu. On a TI-89, the command is option 4 on the `Distr` menu. Fill in the boxes as prompted; they will look just like the ones on the `Shade Norm` screen. The command on a TI-83/84 is used as follows.



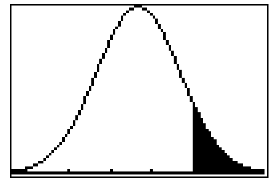
$P(a < X < b)$

`normalcdf`(a,b,μ,σ)



$P(X < k) = P(X \le k)$

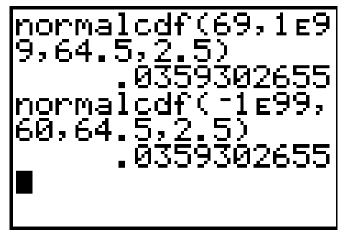
`normalcdf`(–1E 99,k,μ,σ))



$P(X > k) = P(X \ge k)$

`normalcdf`(k,1E 99,μ ,σ))

Example 1.10 More Women's Heights. Find the proportion of American women that are taller than 5'9" (69"). Also find the proportion of women that are shorter than 5' (60 inches).

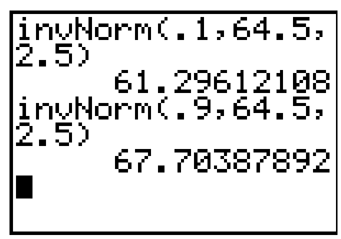


The screen at right shows the commands and results for these questions. Instead of using 1E99 and –1E99, we could also have used a large positive value and a large negative value with no loss of accuracy. We see that about 3.6% of American women are taller than 5'9"; the same proportion are shorter than 5'.

Inverse Normal Calculations

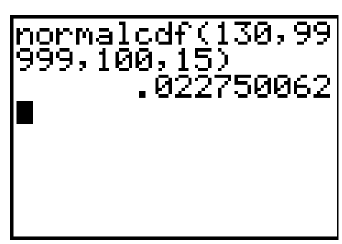
To find the value x for which $P(X \le x)$ equals a desired proportion p (an inverse normal calculation), we use the command `invNorm(`p, μ, σ). This is option 3 on the TI-83/84 `DISTR` menu. On a TI-89, press the right arrow to expand the Inverse menu where this is option 1. The following examples demonstrate these commands.

Example 1.11 Women's Heights: Finding Percentiles. How short are the shortest 10% of American women? How tall are the tallest 10% of American women?



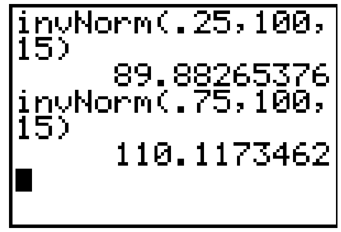
Solution: To find the heights of the shortest 10%, we have $P(X \le x) = 0.1$. My first calculation shows that these women are shorter than about 61.3". To find the heights of the tallest 10% of American women, we could use symmetry of Normal distributions (move an equal distance above the mean), but we recognize that $P(X \le x) = 0.9$. The results of my calculation indicate that the tallest 10% are at least 67.7" tall.

Example 1.12 IQ Scores. The Weschler Adult Intelligence Scale (WAIS) provides IQ scores that are normally distributed with a mean of 100 and a standard deviation of 15. (a) What percent of adults would score 130 or higher? (b) What scores contain the middle 50% of all scores?



Solution. (a) We let $X \sim N(100,15)$ and enter the command at right `normalcdf`(130,99999,100,15). We find that about 2.28% of people should have IQs of at least 130.

(b) If 50% of scores are between *a* and *b*, then 25% of scores are below *a* and 25% of scores are above *b*. So *a* is the inverse normal of 0.25 and *b* is the inverse normal of 0.75. We see that (practically speaking) the middle 50% of people have IQs between 90 and 110.



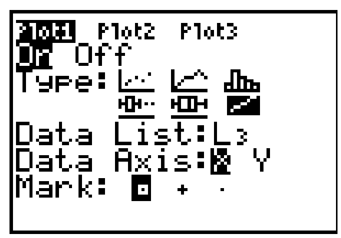
Normal Quantile Plots

Normal Quantile plots on TI-calculators plot the data values against the *z*-score that value would have if the data came from a normal distribution. This used to be an extremely tedious plot to create by hand, but technology makes these easy. If the data do come from an (approximately) normal distribution, the plot of points should be a straight line.

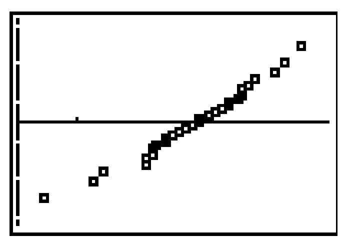
Example 1.13 The Density of the Earth: A Normal Plot. Make a normal quantile plot of Cavendish's data that were used in Examples 1.3 and 1.5.

5.50	5.61	4.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

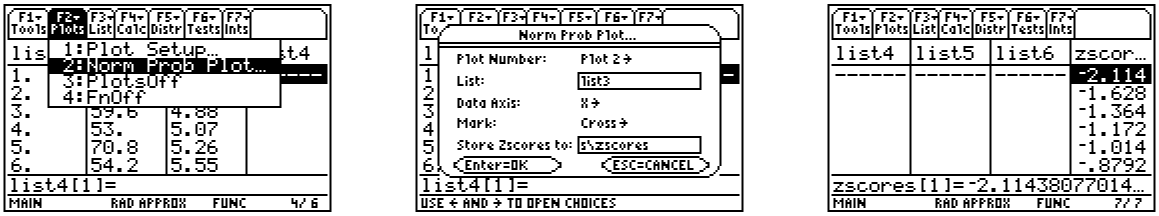
TI-83/84 Solution. We still have these data in list `L3`. Normal quantile plots are the last plot type on a plot definition screen. You have the option of either the *x*- or *y*-axis containing the data, which is a matter of personal preference. Since we hope to see a straight line result, it really makes no difference, but this author is accustomed to data on the *x*-axis, so we have selected that. You (as always) have the option of selecting the mark for each data point.

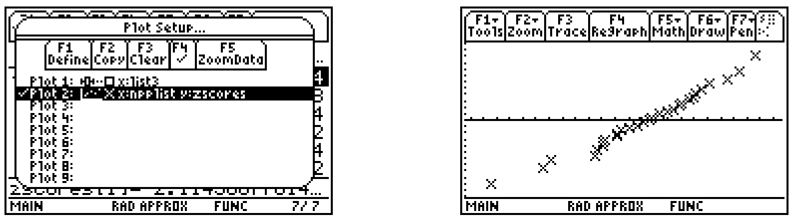


Press θo to display the graph. Recall from the boxplot of these data (page 22) that there was a long left tail. In this plot, these data values seem to split away somewhat from the bulk of the data. While they weren't outliers, they are a little too far out for this data to be perfectly normal. We might be happy to call it approximately normal.



TI-89 Solution. On TI-89 calculators there is an intermediate step in creating a normal quantile plot that makes the *z*-score computations more explicit. Since we still have Cavendish's data in `list3`, from the Statistics/List Editor application press □ for the `Plots` menu. Select option `2:Norm Prob Plot`. You can select the plot number (it defaults to one higher than what is already defined), the list to use (use ° to insert the list name), the axis that will contain the data, and a list to contain the *z*-scores (we recommend taking the defaults). Press ⊆ to perform the calculations. *Z*-scores for each data value are stored and displayed on the editor screen. Return to the `Plot Setup` menu. Uncheck using □ or clear any other plots. Move the cursor to highlight the normal plot, then press □ to display the final plot.

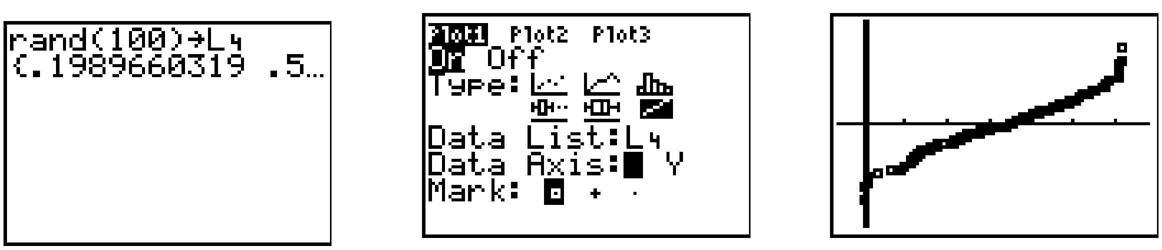




Example 1.14 A Normal Quantile Plot for a Uniform Distribution. Certain types of distributions have classical normal quantile plots. Generate 100 observations from a Uniform[0,1] distribution and display its normal plot.

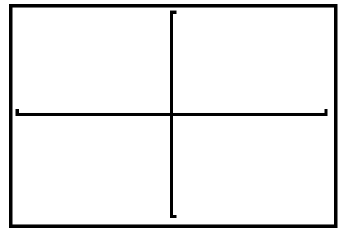
Solution. Press □. Arrow to `PRB`. The first option on this menu generates uniform random numbers on [0,1]. Since we want 100 of them, we will use the command `rand(100)↵L4`.

The calculator will echo back the first one or two of these. To look at them further, use the editor. Define the normal plot as before to use this list. Press θo to display the plot. Here we see the classical bends of a uniform distribution at each end of the plot. This is because uniform random variables have abrupt ends — there is no gradual tapering of the distribution as there is with a normal distribution.

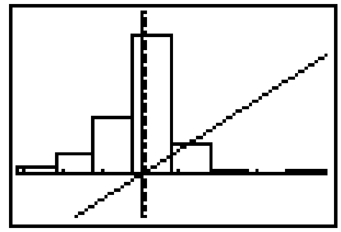


1.4 Common Errors

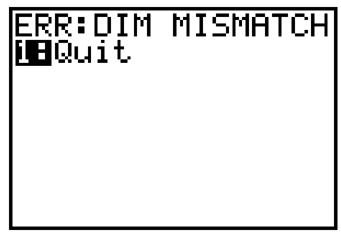
There's no picture!
Seeing something like this (or a blank screen) is an indication of a windowing problem. This is usually caused by pressing σ using an old setting. Try pressing θ→ to display the graph with the current data. This error can also be due to having failed to turn the plot "On."



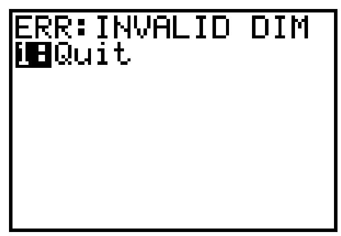
What's that weird line (or curve)?
There was a function entered on the o screen. The calculator graphs everything it possibly can at once. To eliminate the line, press o. For each function on the screen, move the cursor to the function and press □ to erase it. Then redraw the desired graph by pressing σ.



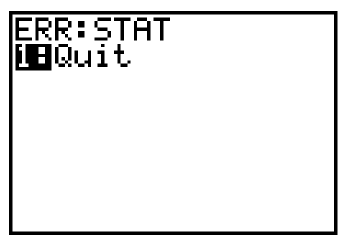
Err: Dim mismatch
This common error results from having two lists of unequal length. Here, it pertains either to a histogram with frequencies specified or a time plot. Press ⊆ to clear the message, then return to the statistics editor and fix the problem.



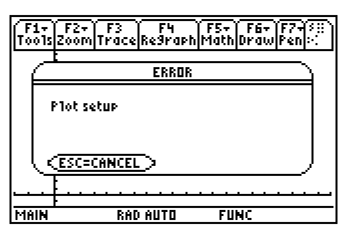
Err: Invalid dim
This problem is generally caused by reference to an empty list. Check the statistics editor for the lists you intended to use, then go back to the plot definition screen and correct them.



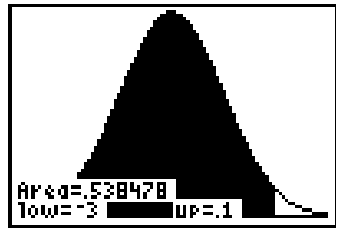
Err: Stat
This error is caused by having two stat plots turned on at the same time. What happened is the calculator tried to graph both, but the scalings are incompatible. Go to the STAT PLOT menu and turn off any undesired plot.



Plot setup
This is the TI-89 equivalent of the STAT error above. It is caused by having two stat plots turned on at the same time. The calculator tried to graph both plots, but the scalings are incompatible. Go to the Stat plots menu and turn off any undesired plot by moving the cursor to that plot, and pressing □.

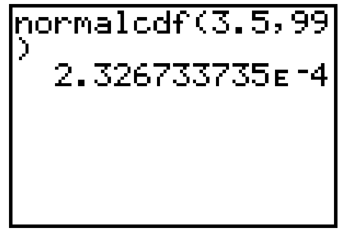


Why is my curve all black?



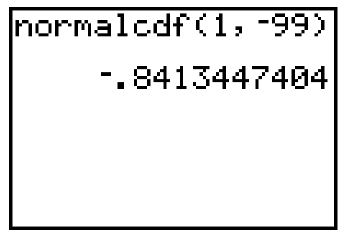
For the standard normal curve, the graph indicates that well over half of the area is of interest between –3 and .1. The message at the bottom says the area is 53.8%. This is a result of having failed to clear the drawing between commands. Press ψ□ then ⊆ to clear the drawing, then reexecute the command.

Probability more than 1



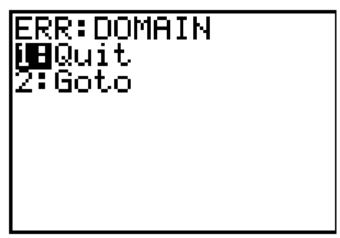
This is not possible. If the results look like the probability is more than one, check the right side of the result for an exponent. Here it is –4. That means the leading 2 is really in the fourth decimal place, so the probability is 0.0002. The chance of a variable being more than 3.5 standard deviations above the mean is about 0.02%.

Negative probability



This is not possible. The low and high ends of the area of interest have been entered in the wrong order. As the calculator does a numerical integration to find the answer, it doesn't care, but you should.

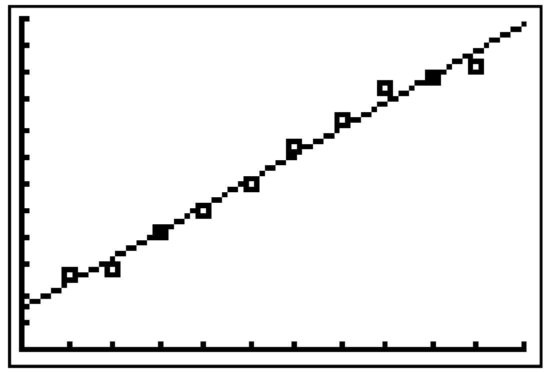
Err: Domain



This message comes as a result of having entered the invNorm command with parameter 90. (You wanted to find the value that puts you into the top 10% of women's heights, so 90% of the area is to the left of the desired value.) The percentage must be entered as a decimal number. Reenter the command with parameter .90.

CHAPTER

2



Looking at Data—Exploring Relationships

Introduction

In this chapter, we use TI calculators to graph the relationship between two quantitative variables using a scatterplot. We then show how to compute the correlation and find the least-squares regression line through the data. Last, we show how to work with the residuals of the regression line.

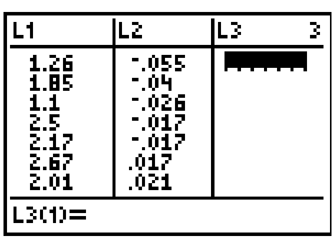
2.1 Scatterplots

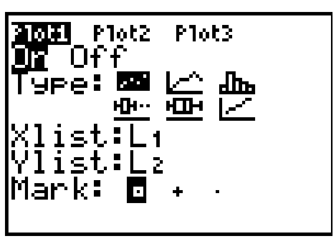
We begin by showing how to make a scatterplot of two quantitative variables along the *x* and *y* axes so that we may observe if there is any noticeable relationship. In particular, we look for the strength of the linear relationship.

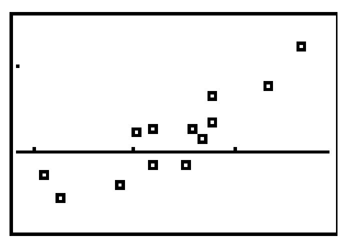
Example 2.1 Brain Activity and Stress. It has been suggested that emotional stress leads to increased brain activity. Make a scatterplot of brain activity level against social distress score.

Subject	Social distress	Brain activity	Subject	Social distress	Brain activity
1	1.26	–0.055	8	2.18	0.025
2	1.85	–0.040	9	2.58	0.027
3	1.10	–0.026	10	2.75	0.033
4	2.50	–0.017	11	2.75	0.064
5	2.17	–0.017	12	3.33	0.077
6	2.67	0.017	13	3.65	0.124
7	2.01	0.021			

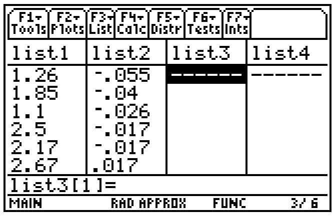
TI-83/84 Solution. We first enter the data into the `STAT EDIT` screen. Here we use `L1` for the social distress scores, plotted on the *x* axis, and use `L2` for the brain activity levels that will be plotted on the *y* axis. Define `PLOT1` on the `STAT PLOTS` menu to use the first (scatterplot) plot type as below. You have your choice of marks for each data point, but we do not recommend the last (single pixel) as it is hard to see. Press θo to display the plot.

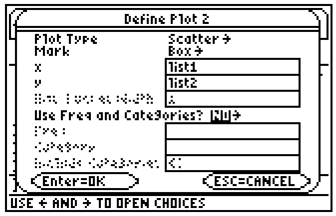


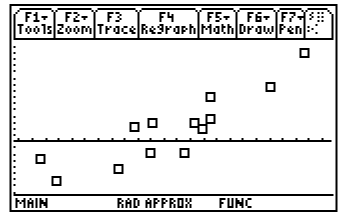




TI-89 Solution. With a TI-89, the sequence required to generate the scatterplot is very similar . We have entered the distress scores into `list1` and the brain activity levels into `list2`. Press □ for `Plots` and select option `1:Plot Setup`. Clear or uncheck (□) any unneeded plots. Select a plot to define and press □ to proceed. The plot type is `Scatter`. You can use the right arrow key to expand the plot type options, if needed. Select a plot symbol for each data point in a like manner. Use ψ| (°) to access the lists of variable names and select `list1` for the `x` variable and `list2` for the `y` variable. We are not using separate lists of frequencies or categories, so this should be set to `NO`. Press ⊆ to complete the definition and return to the Plot Setup screen. Press □ to display the plot.





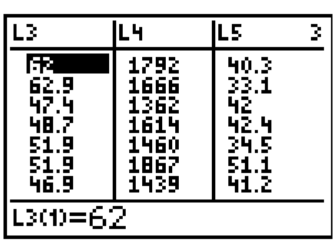


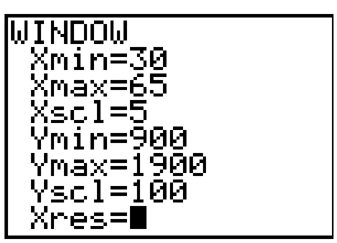
We see that as the social distress score increases, the brain activity level generally tends to increase also.

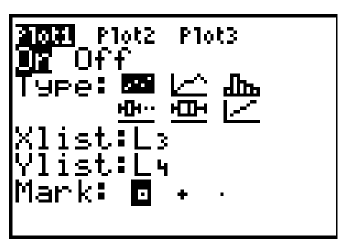
Example 2.2 Body Mass and Gender. We would like to examine whether or not there seems to be a difference in metabolic rate versus body mass for males and females. We will use two scatterplots with different symbols for males and females, and display both at once.

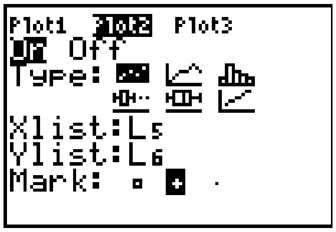
Sex	Mass	Rate	Sex	Mass	Rate
M	62.0	1792	F	40.3	1189
M	62.9	1666	F	33.1	913
M	47.4	1362	F	42.0	1418
M	48.7	1614	F	42.4	1124
M	51.9	1460	F	34.5	1052
M	51.9	1867	F	51.1	1347
M	46.9	1439	F	41.2	1204
			F	54.6	1425
			F	50.6	1502
			F	36.1	995
			F	42.0	1256
			F	48.5	1396

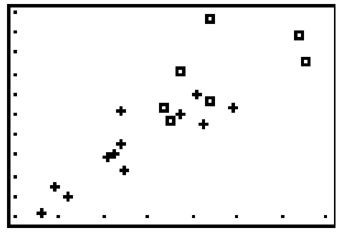
Solution. We first enter the mass and rate of just the males into lists `L3` and `L4` respectively. Then we enter the mass and rate of the females into lists `L5` and `L6`. Here, we will need to adjust the `WINDOW` so that the *X* range includes all the masses and the *Y* range includes all the rates. We adjust the `STAT PLOT` settings for `Plot1` to obtain the scatterplot of `L3` versus `L4` (the males), and adjust the `STAT PLOT` settings for `Plot2` to obtain the scatterplot of `L5` versus `L6` (the females). Since we have explicitly sized the window, press σ to display the plots.









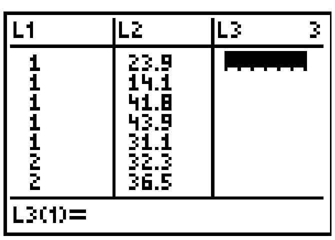


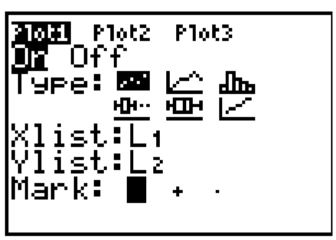
We see a linear pattern between the two genders. This is clearly stronger for women who have both lower body mass and metabolism. The men's data at the upper right is much more scattered (a weaker relationship).

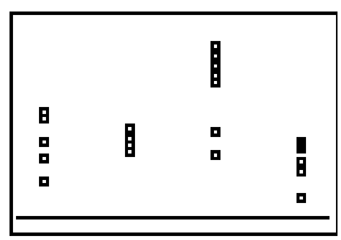
Example 2.3 Sector Fund Returns. How does the return on investment vary among sector mutual funds? Data in the table below are annual total returns from several sector funds. We will make a plot of the total return against market sector. We'll also compute the mean return for each sector, add the means to the plot, and connect the means with line segments, so the averages will stand out more

Market sector	Fund returns (percent)						
Consumer	23.9	14.1	41.8	43.9	31.1		
Financial services	32.3	36.5	30.6	36.9	27.5		
Technology	26.1	62.7	68.1	71.9	57.0	35.0	59.4
Natural resources	22.9	7.6	32.1	28.7	29.5	19.1	

Solution. We will plot the market sectors on the *x* axis as the values 1, 2, 3, and 4. Because there are multiple returns for each sector, we enter each of the values 1 through 4 as many times into list L1 as there are returns for that sector. We enter the corresponding returns into list L2. We define the scatterplot as we have done previously, and press θo to display the plot (there is no need to specially size this one).

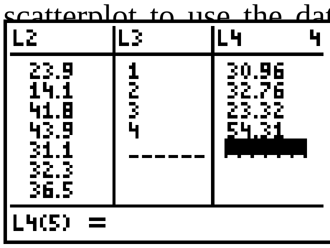


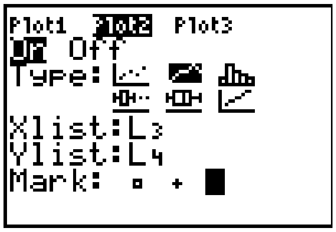


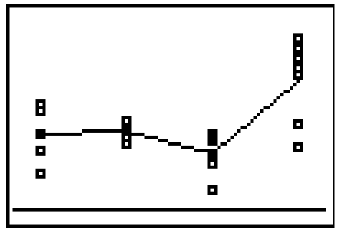


So far, we see that the returns in sector 3 (Technology) are much more variable than the others — there is potential for both greater and smaller returns. Natural resources seems to have the lowest returns, and Financial Services the most consistent (least variability).

Data for each sector was reentered into a new list and the mean calculated for each. We have now entered the category numbers and means. We will define a new connected scatterplot to use the data in lists L3 and L4, this scatterplot will use the single pixel ...t to have the appearance of an extra data point. Since the graph ...or these data, pressing either σ or θ→ will display the plot.







2.2 Correlation

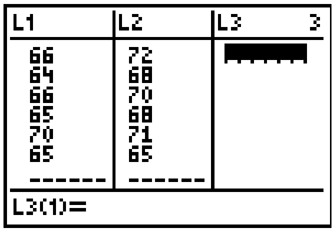
In this section, we compute the correlation coefficient *r* between paired quantitative variables. For all TI calculator models, to obtain correlations, the linear regression must be computed

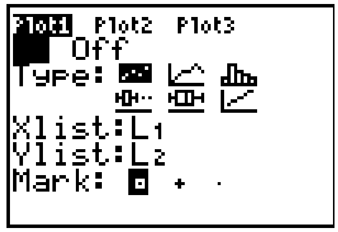
If you are using a TI-83 or -84, we must make sure that the calculator's diagnostics are turned on. Enter the `CATALOG` (2⊇) and press □ to advance to functions beginning with the letter D. Scroll down until you find the `DiagnosticOn` command. Press ÷ to bring the command to the `Home` screen, then press ÷ again to execute the command.

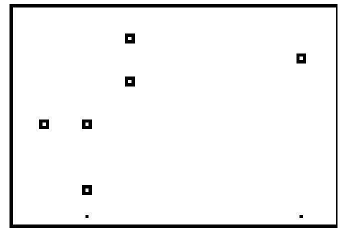
Example 2.4 Dates' Heights. The table below gives the heights in inches for a sample of women and the last men whom they dated. (a) Make a scatterplot. (b) Compute the correlation coefficient *r* between the heights of these men and women. (c) How would *r* change if all the men were 6 inches shorter than the heights given in the table?

Women (x)	66	64	66	65	70	65
Men (y)	72	68	70	68	71	65

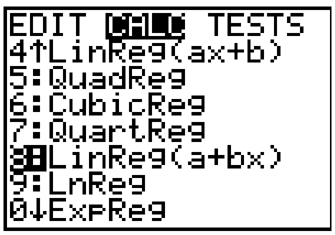
TI-83/84 Solution. (a) We enter the heights of the women into list `L1` and the heights of the men into list `L2`, adjust the `STAT PLOT` settings, and display the graph using θ→. If you have just completed working through the previous example, be sure to turn `Plot2 OFF`.

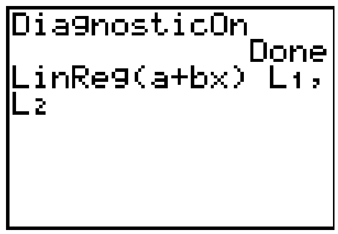


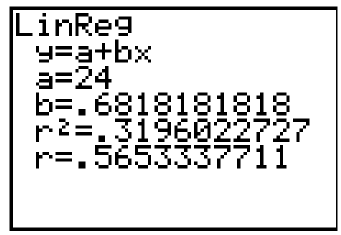




(b) To compute the correlation, we use the `LinReg(a+bx)` command (option 8) from the `STAT CALC` menu. Enter the command `LinReg(a+bx) L1,L2`. The `LinReg(ax+b)` command (item 4) will also compute the correlation. We are actually computing the linear regression here (but won't examine it yet) in order to get the correlation.

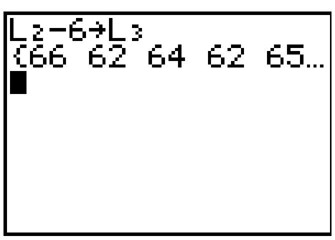


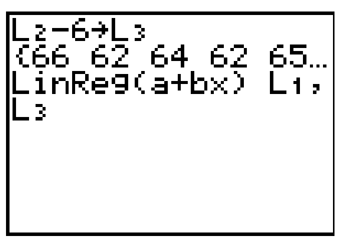


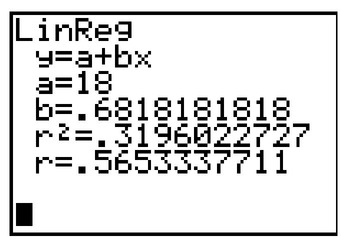


The correlation between the men's and women's heights is r = 0.565.

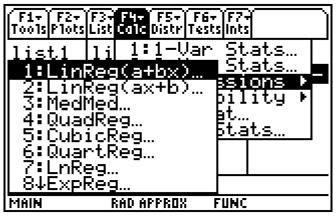
(c) In order to get heights for all the males that are 6 inches shorter than they originally were, we use the calculator to subtract 6 from each entry in `L2` and store the result into `L3`. The ♣ key is above × at the lower left of the keyboard. After pressing ÷ to perform the calculation, the first few values of the new list are shown. To view the complete list, use the editor. Then we use the command `LinReg(a+bx) L1,L3` to see that r has not changed.

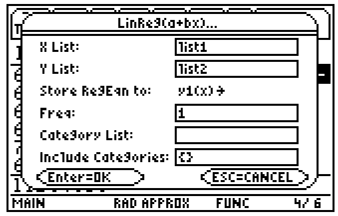


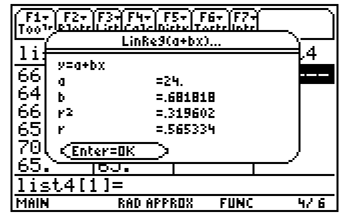




TI-89 Solution: The data have been entered into `list1` and `list2`. Press □ for the `Calc` menu and select option `3:Regressions`. The submenu of all possible regressions will open. Select option `1:LinReg (a+bx)`. In the dialog box, we have given our list names, and also asked that the regression equation be stored in the `y1(x)` function. Press ÷ for the results.







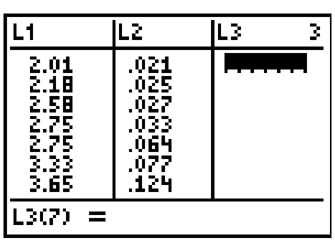
2.3 Least-Squares Regression

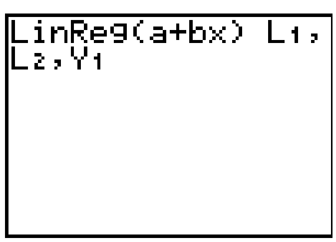
In this section, we will compute the least-squares line of two quantitative variables and graph it through the scatterplot of the variables. We will also use the line to predict the *y*-value that should occur for a given *x*-value.

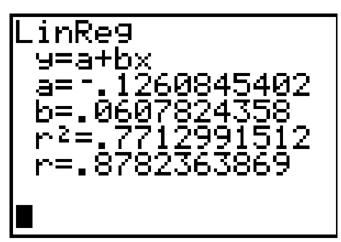
Example 2.5 More Brain Activity and Stress. The data from Example 2.1 are repeated below. (a) What is the equation of the least-squares regression line for predicting brain activity from a social distress score? Make a scatterplot with this line drawn on it. (b) Use the equation of the regression line to get the predicted brain activity level for a distress score of 2. (c) What percent of the variation in brain activity among these subjects is explained by the straight-line relationship with social distress score?

Subject	Social distress	Brain activity	Subject	Social distress	Brain activity
1	1.26	–0.055	8	2.18	0.025
2	1.85	–0.040	9	2.58	0.027
3	1.10	–0.026	10	2.75	0.033
4	2.50	–0.017	11	2.75	0.064
5	2.17	–0.017	12	3.33	0.077
6	2.67	0.017	13	3.65	0.124
7	2.01	0.021			

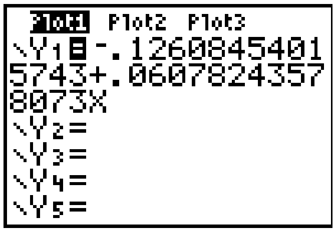
Solution. (a) We obtain the linear regression line using the same `LinReg(a+bx)` command that computes the correlation. After entering data into lists, say `L1` and `L2`, we would enter the command `LinReg(a+bx) L1,L2`. However, in order to store the equation and be able to use it for graphing or prediction, we will add a new parameter to the LinReg command. Add a comma after `L2` and press □, press ~ to move the highlight to Y-VARS, press ÷ to select option 1:Function and ÷ to select `Y1`. Finally, press ÷ to perform the calculations.

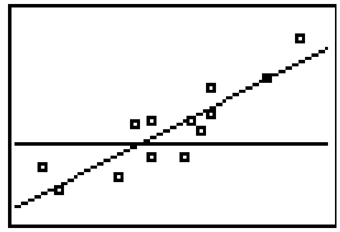




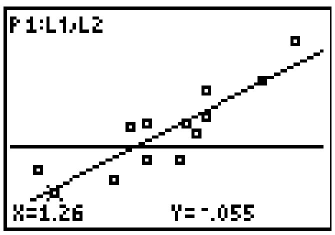


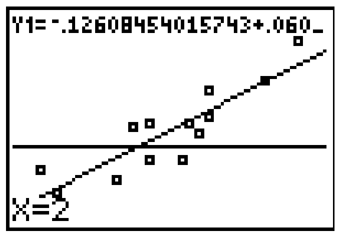
You can see the equation of the line stored by pressing o. Our equation is Brain-Activity = –0.126+0.061*Distress. To graph the data along with the regression line, press θ→ (□ for `ZoomData` from the `Plot Setup` Screen on a TI-89).

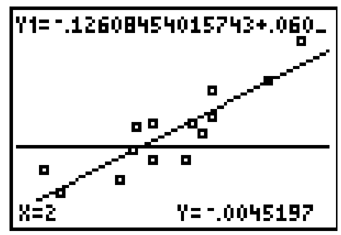




(b) With the equation of the regression line computed and displayed on the graph as above, we can use the ρ (□ on a TI-89) to evaluate the function for a specific *x* value. Press ρ. At the upper right, the active plot is displayed, which indicates we are tracing the scatterplot. Press □ to switch to the line. The top line will change to display the equation of the line. Type in 2 (the distress value we want to find a brain activity for) and press ÷. We see that a distress score of 2 gives a predicted brain activity of –0.0045.

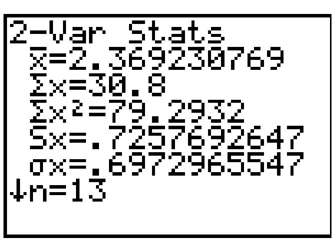


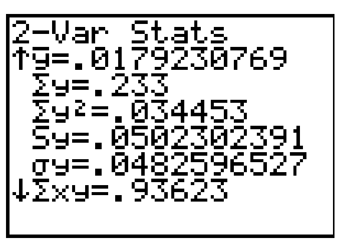


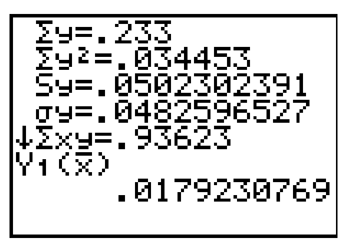


Alternately, we can access the `Y1` function from the `Home` screen. To do so, press □, arrow right to `Y-VARS`, enter ♦ for `Function`, enter ♦ for `Y1`, then enter the command `Y1(2)`. Press ÷ to perform the calculation.

We can also verify that the point $(\bar{x}, \bar{y})$ is on the regression line; but first we must compute the statistics. We can do so simultaneously with the `2-Var Stats` command from the `STAT CALC` menu, because the two data sets are the same size. Enter the command `2-Var Stats L1,L2`. Then enter `Y1(`$\bar{x}$`)` by recalling $\bar{x}$ from the □ `Statistics` menu.







(c) With the calculator's diagnostics turned on, the `LinReg(a+bx)` command also displays the values of r and r^2. In this case, $r^2 = 0.7713$. Thus, 77.13% of the variation in brain activity among these subjects is explained by the straight-line relationship with social distress score.

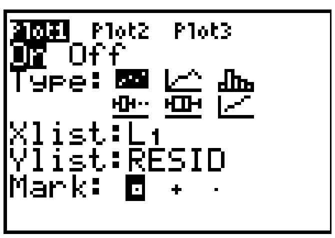
2.4 Cautions About Correlation and Regression

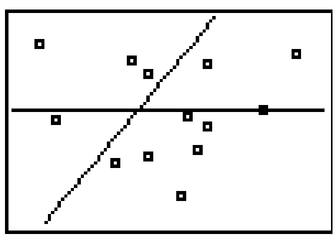
We now complete an exercise to demonstrate how to work with the residuals of a least-squares regression line. These are computed for each regression and stored in a list called `RESID`. Ideally, these will be randomly distributed around the x-axis ($y = 0$ line). Any indications of pattern (curvature, widening (or narrowing) as x-values increase) show a violation of assumptions for the regression. They can also be used to help identify outliers and potentially influential points in a regression.

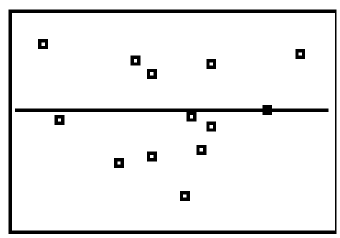
Example 2.6 More Brain Activity and Stress. We continue with our example on distress and brain activity. We want to obtain the residuals scatterplot against our x (predictor) variable.

Solution. We first define the scatterplot as shown below. To access the list named RESID, press 2□, and locate the list name. Press ÷ to transfer the list name to the plot definition screen. Press θ→ to display the plot. Oops! The regression line shown on our plot below is *not* part of the residuals plot! Go to the o screen and press □ to erase the equation. Redisplay the plot using either θ→ or σ.

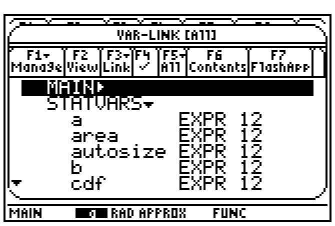
This plot is pretty much ideal. Points are relatively evenly spread around the $y = 0$ axis and seem randomly distributed.

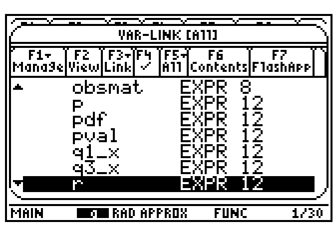


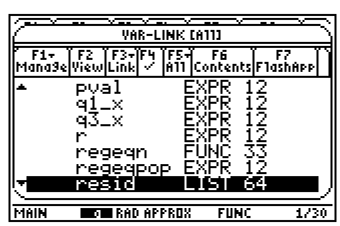




TI-89 Note: To locate the `resid` list on the ° screen, press A to collapse the `MAIN` folder. Press the down arrow to get inside the `STATVARS` folder, then press ♥ (R in alpha mode) to get to the portion of these variables that begin with the letter r. Move the arrow down until you find the list and press ÷ to select it and transfer this name to the plot definition.



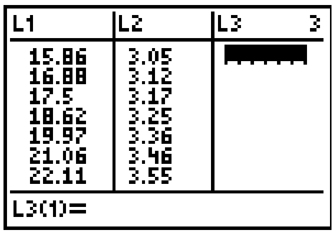


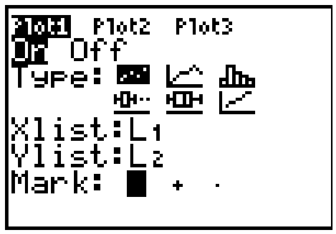


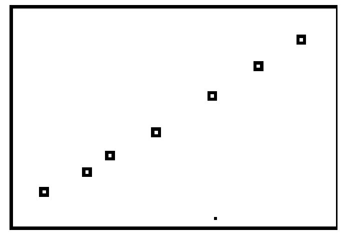
Example 2.7 Runners' Stride Rates. The following table gives the speeds (in feet per second) and the mean stride rates for some of the best female American runners. We'll perform a complete analysis of these data, attempting to predict stride rate (*y*) using the speed (*x*) of the runner.

Speed	15.86	16.88	17.50	18.62	19.97	21.06	22.11
Stride rate	3.05	3.12	3.17	3.25	3.36	3.46	3.55

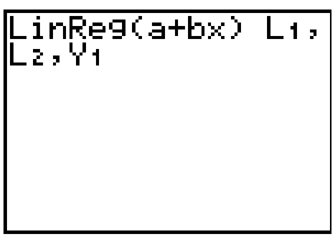
Solution. The data have been entered into `L1` and `L2`. The initial data plot is defined and displayed using θ→.

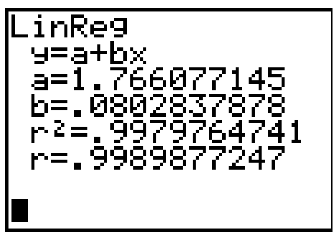


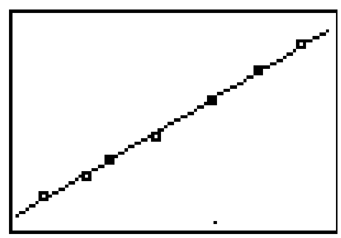




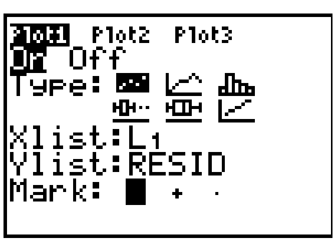
The plot appears to be very linear. We'll calculate the regression, storing the equation into `Y1`, and display the line on the graph.

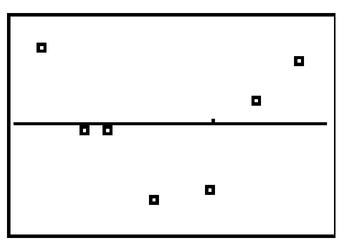






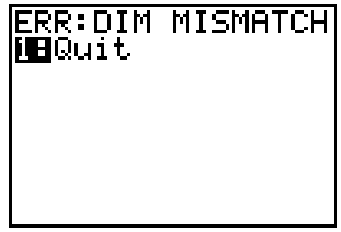
Our regression equation is Stride-rate = 1.766 + 0.080*Speed. This appears to be an extremely strong relationship, since the linear regression explains 99.9% (r^2) of the variation in stride rate. We're not done, however. Examining the residuals plot should always be the next step — is our model the "correct" one for these data? The residuals plot is defined and displayed using θ→.



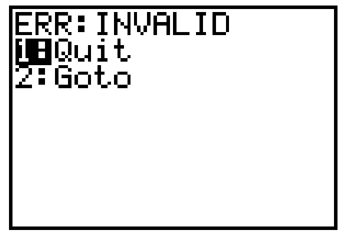


These residuals indicate a clear curve. A straight line model is *not* the correct one!

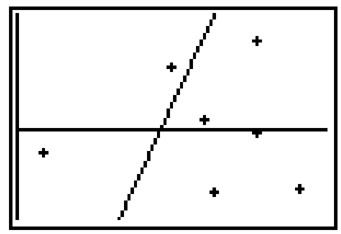
2.5 Common Errors



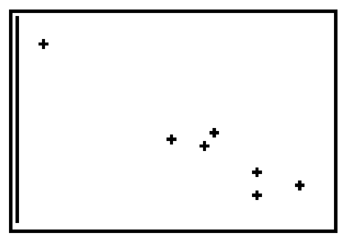
Err:Dim mismatch
We've seen this one before. Press ⊆ to quit. This error means the two lists referenced (either in a plot or a regression command) are not the same length. Go to the STAT editor and fix the problem.



Err: Invalid
This error is caused by referencing the function for the line when it has not been stored. Recalculate the regression being sure to store the equation into a *y* function.



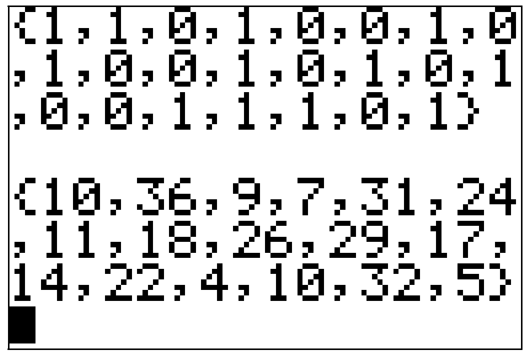
What's that weird line?
This error can come either in a data plot (an old line still resides in the o screen) or the stored regression line is showing in the residuals plot, as shown here. As stated earlier, the regression line is *not* part of the residuals plot and shows only because the calculator tries to graph everything it knows about. Press o followed by □ to erase the unwanted equations, then redraw the graph by pressing σ.



This doesn't look like a residuals plot!
Residuals plots *must* be centered around $y = 0$. This error is usually caused by confusing which list contains the *y*'s and which the *x*'s in computing the regression or in defining the residuals plot. Go back and check which list is which, and recompute the regression, or redefine the plot.

CHAPTER

3



Producing Data

Introduction

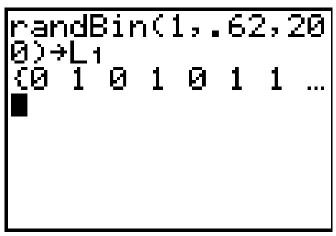
In this chapter, we use TI calculators to simulate the collection of random samples. We also provide a supplementary program that can be used to draw a random sample without repeats from a list of integers numbered from m to n. Good data collection practice involves randomly selecting individuals from the population, or randomly assigning treatments in a controlled experiment. The randomization can be done with a random digits table, a calculator, or a computer. When your text says “start on line xx of table b” the sample drawn in that manner will *not* be random — this is merely a mechanism to be able to write an answer for the back of the book.

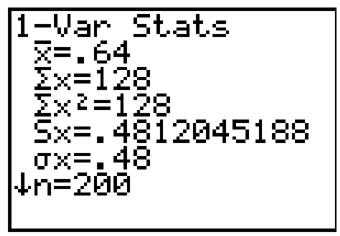
3.1 First Steps

In this section, we demonstrate how to generate count data, or *Bernoulli trials*, for a specified proportion *p*. The data simulates observational “Yes/No” outcomes obtained from a random survey. To generate the data, we use the `randBin` command from the `MATH PRB` menu (option 7).

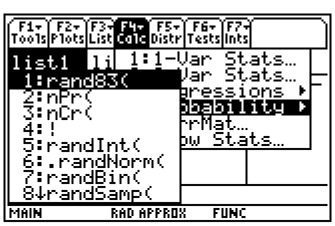
Example 3.1 Simulating a Survey. Suppose that 62% of students hold a part-time or full-time job at a particular university. Simulate the results of a random survey of 200 students and determine the sample proportion of those who have a job.

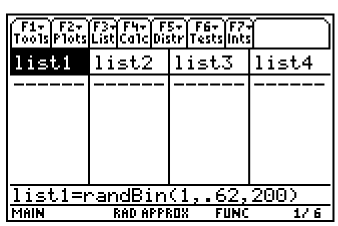
TI-83/84 Solution. To generate a random list of 1 and 0 responses (“Yes/No”), enter the command `randBin(1,p,n)↵L1`, where *p* is the specified proportion of yeses (these will be 1s in our data), and *n* is the desired sample size. Here, use `randBin(1,.62,200)↵L1`. Then enter the command `1-Var Stats L1.` Notice after the command is executed, the first few values are shown on the screen. To be able to see the whole list, use the Statistics Editor. In our sample, we actually had 64% “successes.”

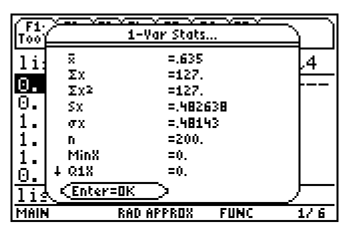




TI-89 Solution. In the `Stats/List Editor` application, move the cursor to highlight the list name that will hold our random data. Press □ for the `Calc` menu, arrow down to option `4:Probability`, and press the right arrow to expand the options. Press ← to select randBin. Complete the command by entering the parameters 1,.62,200 then close the parentheses and press ⊆ to execute. You will see the first few entries in the generated list, and can page down to view the entire list, if desired. We show the results of performing 1-Var Stats on this generated data to show that results are random — here, we had 63.5% 1s.



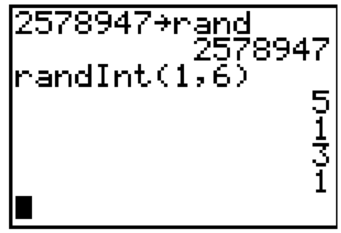




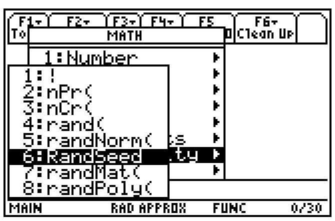
Using the calculator to generate random “samples” is not truly random. These are really *pseudo*random numbers. The calculator uses a value called a “seed” to control the sequence. This seed is changed each time a random number is generated, so results should appear random each time.

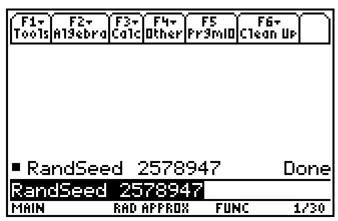
Example 3.2 Changing the Seed. We really want random numbers each time. However, just as your text will instruct you to "start in line xx" when using the table of random digits, your instructor may ask that you use a particular seed so that each student will get the same random numbers. Good seed numbers are large, odd, and preferably prime numbers. The method of setting the seed varies with calculator type. We'll follow the seed setting with a `randInt` command (also found on the `MATH PRB` menu) to simulate throwing a die.

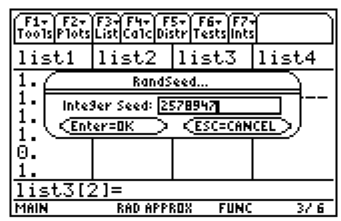
TI-83/84 Solution. Type the desired seed number on the home screen. Press ↵. Then, from the `MATH PRB` menu, select option `1:rand`. Press ⊆ and the calculator will echo the seed number back. Immediately follow this with the `randInt(1,6)` to generate numbers randomly from 1 to 6 (inclusive). To keep generating values, keep pressing ⊆. If you use my seed as at right, you will get the same "random" numbers.



TI-89 Solution. The seed can be set from the home screen using option `6:randseed` from the `Math Probability` menu. Type in the desired seed and press ⊆ to see the `Done` message. This can also be done in the Stats/List Editor application using option `A:randseed` from the `Calc Probability` menu.

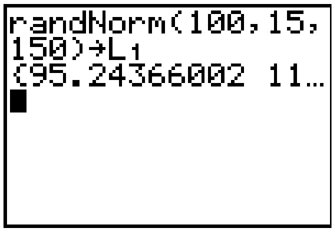


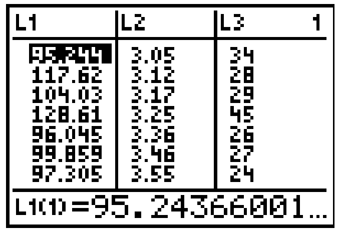


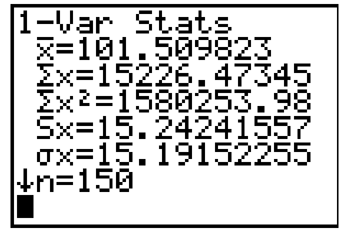


Example 3.3 Simulating IQ Scores. Generate 150 observations from a *N*(100,15) distribution. This distribution will mimic scores for individuals on the Wechsler Adult Intelligence Scale. Compute the sample statistics to compare $\bar{x}$ with 100 and to compare *s* with 15.

Solution. From the `MATH PRB` menu, select `6:randNorm(`. Complete the command by entering the parameters `100,15,150)↵L1` and then compute the sample statistics. If you are using a TI-89, the procedure is essentially the same as used above — select option `6:.randnorm` from the `Calc Probability` menu. Notice that on the home screen, many more decimal places are shown than in the list editor. One point to make here, is that normal random variables are truly continuous (many decimal places are possible) while the IQ scores are really discrete (or at least rounded).







My sample mean (101.5) and standard deviation (15.24) are close to the parameter values, but not exactly the same.

3.2 Design of Experiments

Assigning treatments for a randomized experiment (or selecting individuals for a sample) can easily be done using the `randInt` command we just used. One small drawback is that "selected" individuals can be selected more than once.

Below, we provide a TI-83/84 supplementary program that can be used to choose subjects at random from an enumerated group without any repeated numbers. The TI-89 has a built-in function `randSamp` on the Calc Probability menu that samples with replacement from a list of numbers.

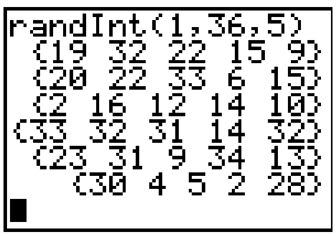
The RANDOM Program

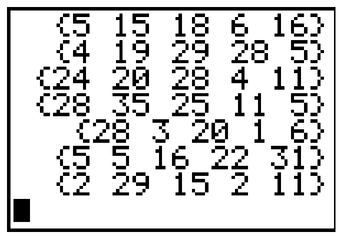
PROGRAM:RANDOM :Disp "LOWER BOUND" :Input M :Disp "UPPER BOUND" :Input N :Disp " HOW MANY?" :Input R :ClrList L4 :seq(J,J,M,N)↵L1 :For(**I**,1,R) :ClrList L5 :randInt(1,N-M+2-**I**) ↵A :L1(A) ↵L4(**I**) :1↵K :While K<A	:L1(K) ↵L5(K) :1+K↵K :End :A↵K :While K≤N-M+1-**I** :L1(K+1) ↵L5(K) :1+K↵K :End :L5↵L1 :End :L4↵L1 :ClrList L5,L4 :ClrHome :Output(1,2,L1) :Stop

The `RANDOM` program can be used to choose a random subset of *k* subjects from a group that has been numbered from *m* to *n*. It also can be used to permute an entire set of *n* subjects so that the group can be assigned randomly to blocks. The program displays the random choices and also stores the values into list `L1`.

Example 3.4 Random Treatment Assignment. We have 36 subjects numbered 1 through 36 available for a small clinical trial. We want to assign them randomly to four treatment groups, each of size 9.

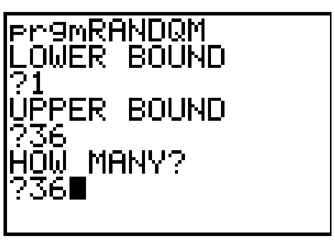
Basic Solution. Use the `randInt(` function from the MATH PRB menu to assign the subjects. The first nine numbers generated will receive the first treatment; the second nine the second treatment, and the third nine will get the third treatment. At that point, anyone not already assigned will get the fourth treatment. We could generate a long list of numbers into a list as has been done previously (this is preferred if using a TI-89), or simply keep hitting ⊆ for numbers singly, but for ease of viewing, we have used the command randInt(1,36,5) which will give 5 numbers each time we press ⊆.

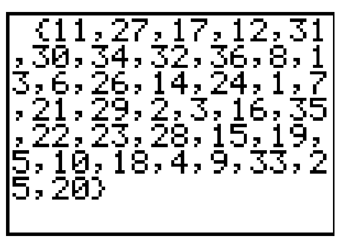


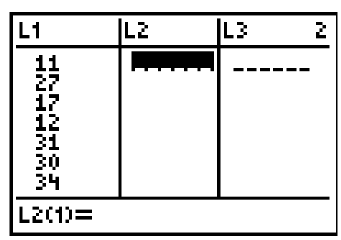


Looking at the values in the first screen, subjects numbered 19, 32, 22, 15, 9, 20, 33, 6, and 2 will receive the first treatment. Subjects 16, 12, 14, 10, 31, 23, 34, 13, and 30 will receive the second treatment. The third treatment will be assigned to subjects 4, 5, and 28. Having generated more sets, we'll complete the treatment 3 group with subjects 18, 29, 24, 11, 35, and 25. Treatment 4 will be assigned to the previously unassigned subjects 1, 3, 7, 8, 17, 21, 26, 27, and 36. While it's fairly easy to do this way, the program can make life a little simpler since there will be no duplicates.

Program Solution. We execute the `RANDOM` program by numbering the subjects from 1 to 36 and choose all 36.



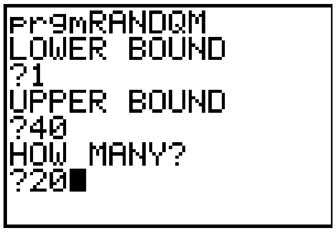


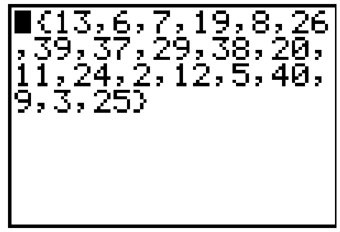


The 36 subjects have been permuted so that they can be assigned randomly to our four groups. Simply use consecutive groups of nine for these groups: {11, 27, 17, 12, 31, 30, 34, 32, 36}, {8, 13, 6, 26, 14, 24, 1, 7, 21}, {29, 2, 3, 16, 35, 22, 23, 28, 15}, {19, 5, 10, 18, 4, 9, 33, 25, 20}.

Example 3.5 More Treatment Assignments. For an experiment with two treatments, we will randomly choose 20 subjects from a group of 40 to receive treatment A. The remaining subjects will receive treatment B.

Solution. We label the subjects from 1 to 40 and then randomly choose 20 for treatment A.



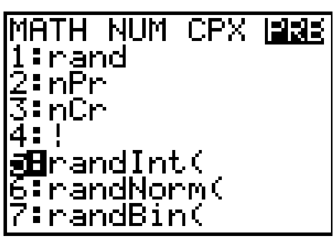


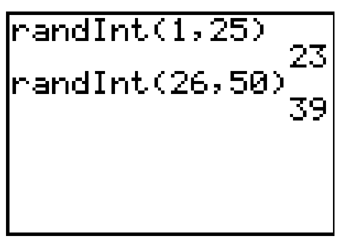
3.3 Sampling Design

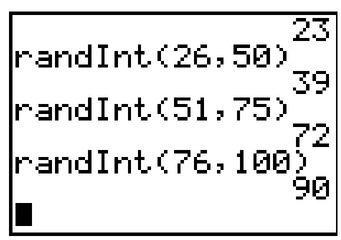
The **RANDOM** program can be used to choose a simple random sample from a designated population. This program chooses the sample all at once without repeated choices. But instead, we may want to use a systematic random sample by drawing one subject at a time from sequential groups. The following exercise demonstrates this process.

Example 3.6 Systematic Sampling. Choose a systematic random sample of four addresses from a list of 100.

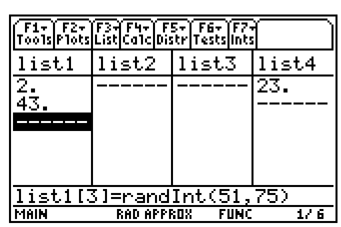
TI-83/84 Solution. Because the list of 100 divides evenly into four groups of 25, we will choose one address from each of the groups 1–25, 26–50, 51–75, and 76–100. And because we are choosing only one number at a time, we can use the `randInt(` command from the `MATH PRB` menu. To choose one integer from *a* to *b*, enter the command `randInt(a,b)`. We use the command four times as shown below to obtain the desired sample.







TI-89 Solution. Place the cursor in the first entry slot in a list. Press (`Calc`), ψ to expand the `Probability` menu, and ζ to select the `randInt` function. Enter the parameters (low number to select, high number to select), close the parentheses and press ⊆. You can repeat these steps changing the low and high numbers to generate the desired four random numbers as above.

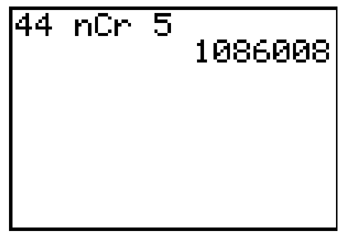


Number of Ways to Choose

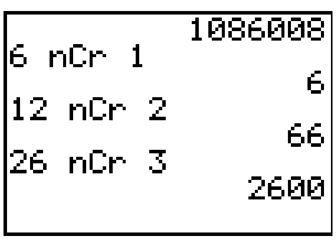
When choosing a simple random sample of size *r* from a population of size *n*, we generally choose without repeats and without regard to order. Such a choice is called a *combination*. The number of possible combinations (often called "*n* choose *r*") can be computed with the `nCr` option from the `MATH PRB` menu. On a TI-89, this is available on the homescreen (2ζ for the `Math` menu) or in the `Stats/List Editor` application from the □ (`Calc`) `Probability` menu.

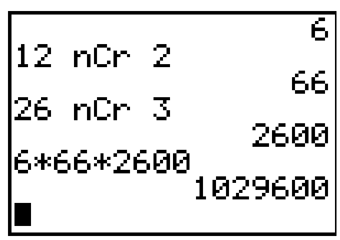
Example 3.7 How Many Samples? (a) How many ways are there to choose five blocks from a group of blocks labeled 1–44? (b) How many ways are there to choose a stratified sample of five blocks so that there is one chosen from blocks 1–6, two chosen from blocks 7–18, and three chosen from blocks 19–44. (c) In each case, choose such a sample.

Solution. (a) There are "44 choose 5" ways to pick five blocks at random from 44, which is computed by `44 nCr 5` on the calculator. Thus, there are 1,086,008 possible samples in this case.

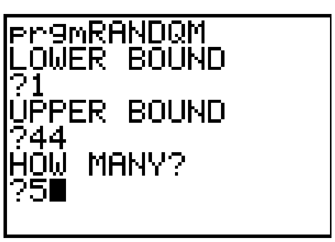


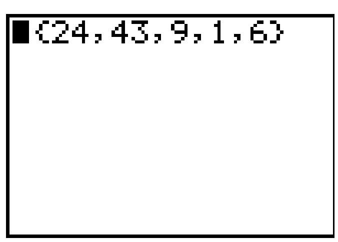
(b) We choose one from the first group of 6 blocks, choose two from the second group of 12 blocks, and choose three from the third group of 26 blocks. The total number of ways to choose in this manner is given by (6 nCr 1)*(12 nCr 2)*(26 nCr 3) = 6 * 66 * 2600 = 1,029,600.

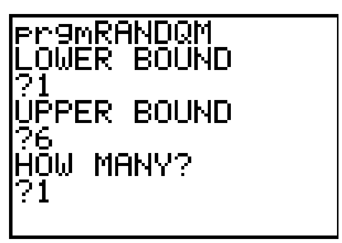


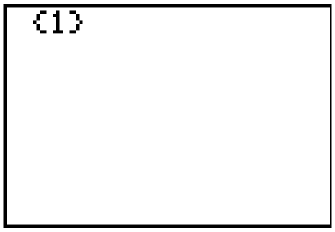


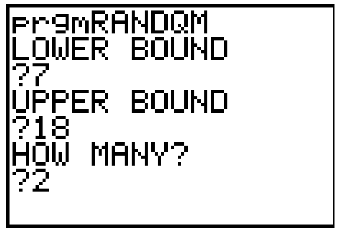
(c) We can choose the samples in each case with the **RANDOM** program. We do not show the last selection.

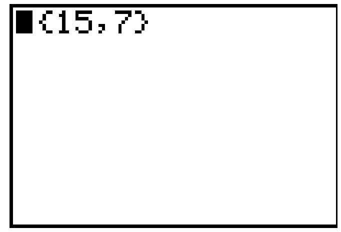










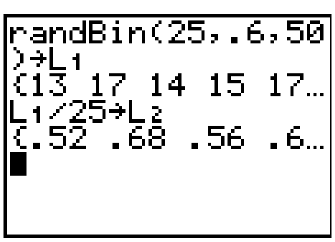


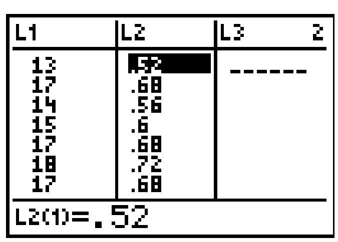
3.4 Toward Statistical Inference

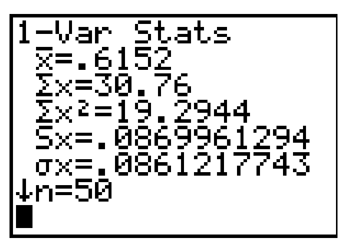
In Section 3.1, we used the command `randBin(1,p,n)↵L1` to generate a random sample of "Yes/No" responses. In this section, we will demonstrate how to simulate the collection of multiple samples. In particular, we are concerned with the total number of "Yes" responses, the sample proportion for each sample, and the resulting average of all sample proportions. This simulation also can be made with the `randBin(` command from the `MATH PRB` menu.

Example 3.8 Simulating Coin Flips. (a) We have a coin for which the probability of heads is 0.60. We toss the coin 25 times and count the number of heads in this sample. Then we repeat the process for a total of 50 samples of size 25. Simulate the counts of heads for these 50 samples of size 25, compute the sample proportion for each sample, and make a histogram of the sample proportions.

Solution. One simulated sample of counts can be obtained with the command `randBin(25,0.6)`. But because we want 50 samples of size 25, we will use the command `randBin(25,0.6,50)↵L1` in order to generate the counts and store them into list `L1`. Then the command `L1/25↵L2` will compute the sample proportion for each sample and store the results in list `L2`. By computing the sample statistics on list `L2`, we obtain the average of all the sample proportions. This example works the same way on a TI-89; however, you must be in the `Stats/List Editor` Application and have a list name highlighted to issue the `randBin` command. Similarly, highlight list name `list2`, for example, to perform the division.

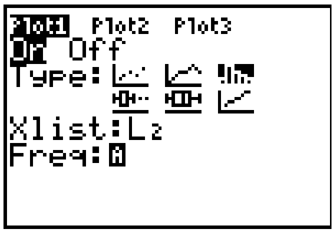


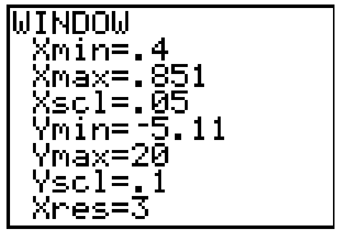


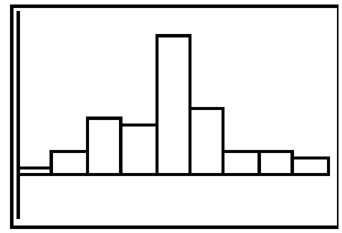


By observing the list of sample proportions in `L2`, we see that we may almost never have a sample proportion that equals the real proportion of 0.60. However, the *average* of all 50 sample proportions was 0.6152, which is very close to 0.60. Lastly, adjust the

WINDOW and STAT PLOT settings to see a histogram of the sample proportions. You could use θo here to display the histogram, but remember the intervals taken by the calculator are usually not logical. Start there, and then adjust Xmin and Xscl(the bar width) to something that makes sense.

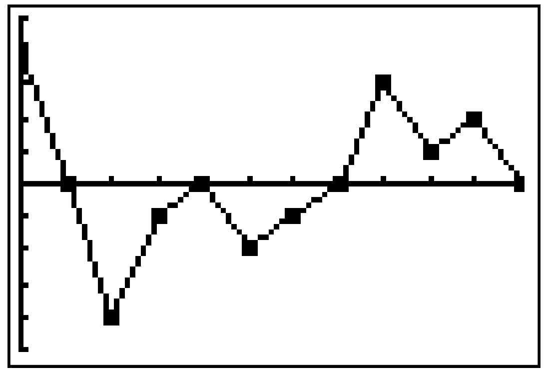






CHAPTER

4



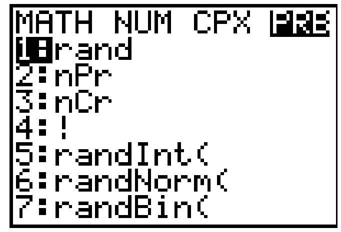
Probability: The Study of Randomness

Introduction

In this chapter, we show how to use TI calculators to generate some random sequences. We then see how to make a probability histogram for a discrete random variable and how to compute its mean and standard deviation. We conclude with a program for the Law of Total Probability and Bayes' Rule.

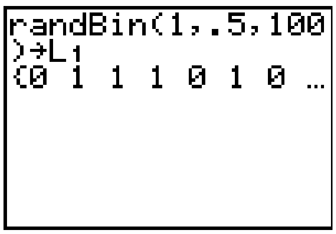
4.1 Randomness

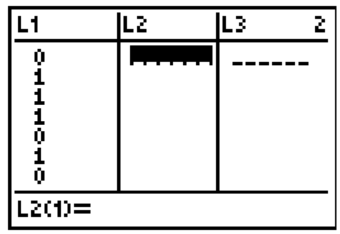
In this section, we work some exercises that use the calculator to generate various random sequences. We shall need the `randBin(` and `randInt(` commands from the `MATH PRB` menu.



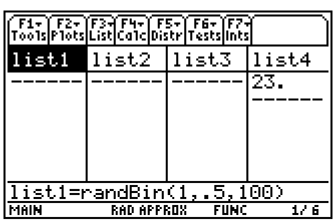
Example 4.1 Simulating Free Throws. Simulate 100 free throws shot independently by a player who has 0.5 probability of making a single shot. Examine the sequence of hits and misses.

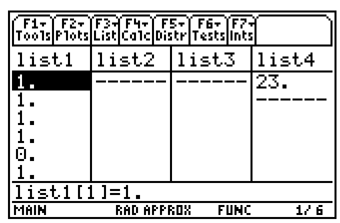
Solution. This is similar to Example 3.1 of the preceding chapter. The command `randBin(1,.5,100)↲L1` will generate and store a list of 100 "1s and 0s" to represent the hits and misses. The results can be viewed on the **STAT EDIT** screen.





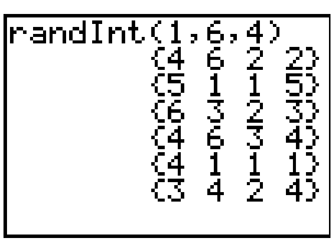
On a TI-89 the solution is similar, but enter the command with the cursor highlighting a list name in the `Stats/List Editor` application. Find the randBin command in the □ (`Calc`), `Probability` menu.

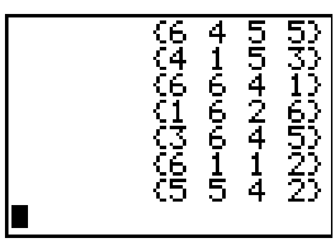


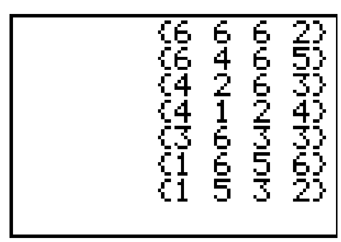


Example 4.2 Simulating Dice. Simulate rolling four fair dice over and over again. What percentage of the time was there at least one "6" in the set of four rolls?

TI-83/84 Solution. The command `randInt(j,k)` generates a random integer from *j* to *k*. The command `randInt(j,k,n)` generates *n* such random integers. Here we enter the command `randInt(1,6,4)` to simulate four rolls of dice numbered 1 to 6. After entering the command once, keep pressing ⊆ to reexecute the command.

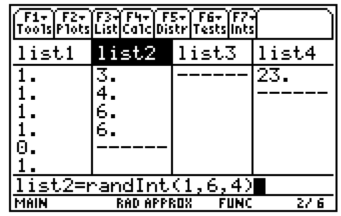






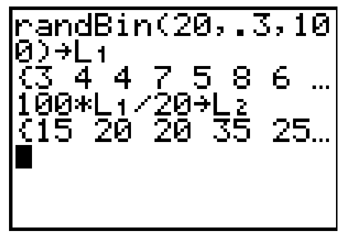
In the 20 sets shown above, there are 13 sets with at least one "6," which gives an estimate of 65%.

TI-89 Solution. This is not as elegant as with the other models. The TI-89 will only generate random numbers of this type when the command is entered while a list name is highlighted. Keep reentering the command to repeat the simulation of rolling four dice.

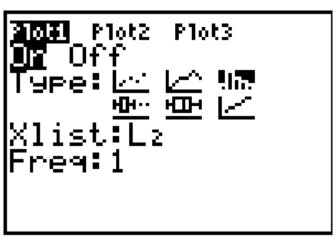


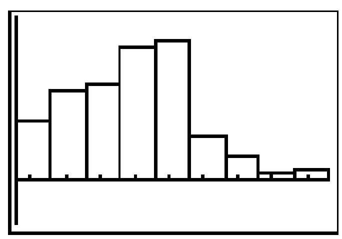
Example 4.3 Simulating Binomial Counts. Simulate 100 binomial observations each with $n = 20$ and $p = 0.3$. Convert the counts into percents and make a histogram of these percents.

Solution.
The command `randBin(20,0.3,100)↲L1` will generate the observations and put them in list `L1`, and the command `100*L1/20↲L2` will put the percents into list `L2`. On a TI-89, enter the commands with a list name highlighted in the `Stats/List Editor` application.



Next, adjust the `STAT PLOT` settings to make a histogram of the percents in list `L2`. If necessary, use `WINDOW` to adjust the plot for logical intervals.



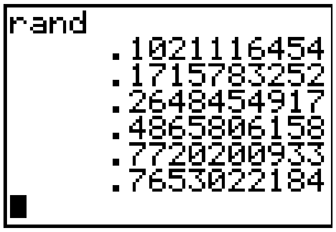


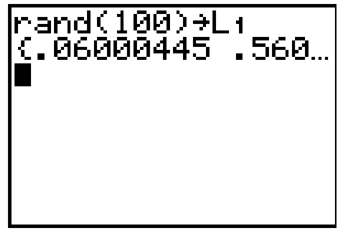
4.2 Probability Models

In this section, we demonstrate some of the basic concepts of probability models.

Example 4.4 Generating Random Numbers (a) Generate random numbers between 0 and 1. (b) Make a histogram of 100 such randomly generated numbers.

Solution. (a) Simply enter the command `rand` from the `MATH PRB` menu. On a TI-89, this can be done either on the home screen, or in the Stats/List Editor application. To store a list on the TI-89, use the command with the list name highlighted as has been shown before. After the command has been entered once, keep pressing ⊆ to continue generating more random values between 0 and 1. (b) Enter the command `rand(100)↲L1` to generate and store 100 random numbers between 0 and 1.



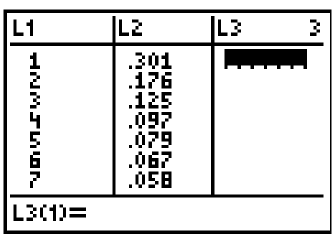


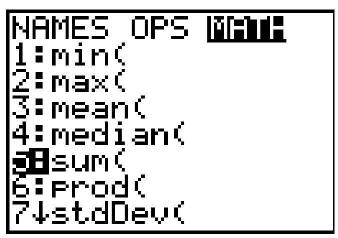
Example 4.5 Benford's Law. The first digit *v* of numbers in legitimate records often follows the distribution given in the table below, known as Benford's Law. (a) Verify that the table defines a legitimate probability distribution. (b) Compute the probability that the first digit is 6 or greater.

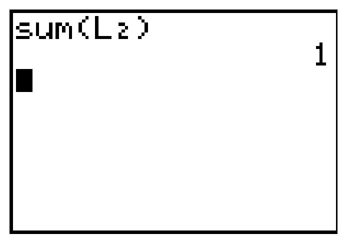
First digit *v*	1	2	3	4	5	6	7	8	9
Probability	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Solution. (a) To sum these probabilities, we can enter the values into a list and use the command `sum(` from the `LIST MATH` menu.

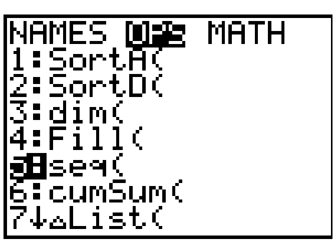
First, enter the values of the digits into list `L1` (optional at this point, but we'll use it later) and enter the probabilities into list `L2`. To verify that the table gives a legitimate probability distribution, enter the command `sum(L2)`, which sums the values in list `L2`. If you are using a TI-89, move the cursor to an empty list entry, and find the command in the □ `List` menu.

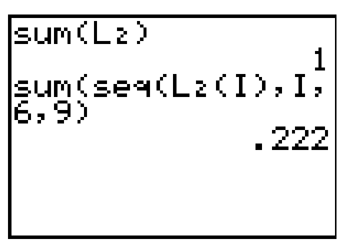






(b) For this part of the question, we only want to sum some of the entries. After summing all the probabilities, press 2⊆ to recall the previous command, then edit it to `sum(seq(L2(I),I,6,9))` to sum the sixth through ninth probabilities. We find the `seq(` command shell on the `LIST OPS` menu. We see that the probability of the first digit being 6 or greater is 0.222. You will need to press the □ key in order to type the I's in the command.





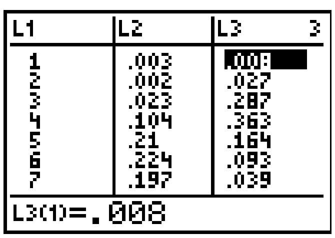
4.3 Random Variables

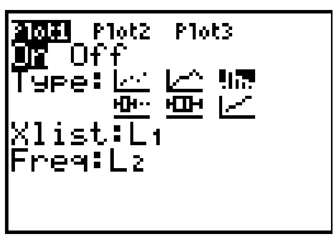
In this section, we work on exercises that compute various probabilities involving random variables. We begin with an exercise on constructing a probability histogram.

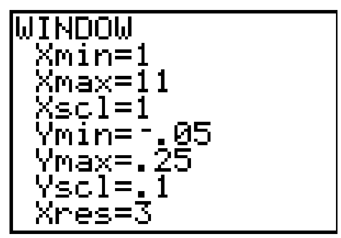
Example 4.6 How Many Rooms? The table below gives the distributions of rooms for owner-occupied units and for renter-occupied units in San Jose, California. Make probability histograms of these two distributions. If *X* represents the number of rooms in a randomly chosen owner-occupied unit, compute $P(X > 5)$.

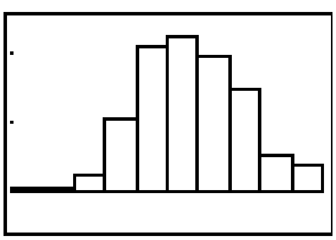
Rooms	1	2	3	4	5	6	7	8	9	10
Owned	0.003	0.002	0.023	0.104	0.210	0.224	0.197	0.149	0.053	0.035
Rented	0.008	0.027	0.287	0.363	0.164	0.093	0.039	0.013	0.003	0.003

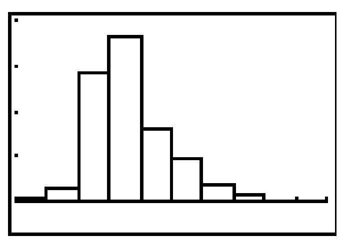
Solution. We enter the values of the rooms into list `L1`, the owner probabilities into list `L2`, and the renter probabilities into list `L3`. Then we make separate histograms for an `Xlist` of `L1` with frequencies of either `L2` or `L3`. Since the number of rooms is discrete, we'll set the window so each bar has width 1. The left histogram is for the owned units, the right for rented units. Note: We had to change `Ymax` for the rented units, as their largest frequency is higher than in the owned distribution.









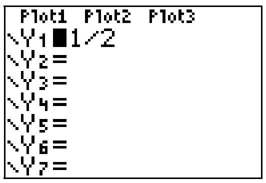


After displaying these histograms, you will need to set Freq: on the plot definition screen back to a 1 by pressing □ℵ.

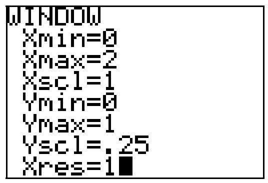
The value $P(X > 5)$ is equivalent to $P(6 \le X \le 10)$. This value can be computed with the command `sum(seq(L2(I),I,6,10))`as described above, which gives 0.658.

Example 4.7 Uniform Distributions. Let $Y \sim U[0, 2]$. (a) Graph the density curve. (b) Find $P(0.5 < Y < 1.3)$.

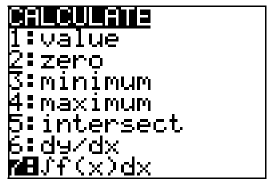
Solution. For $Y \sim U\,[0, 2]$, the height of the density curve is 1/(2 – 0) = 0.5. We simply enter this function into the o screen and use item 7 from the CALC menu to compute the area between the values 0.5 and 1.3. The steps are similar on a TI-89 (using the top row buttons preceded by ∞). Use the □ (Math) button to find the numerical integration command.



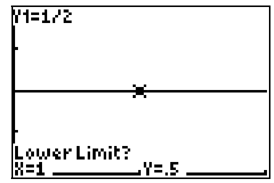
Enter 1/2 into Y= screen.



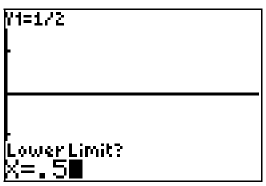
Set WINDOW with X from 0 to 2.



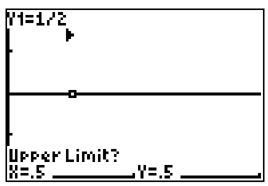
Press CALC (2nd TRACE), then 7.



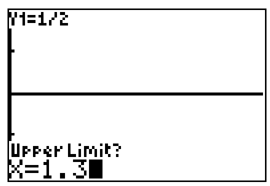
When screen appears, type .5



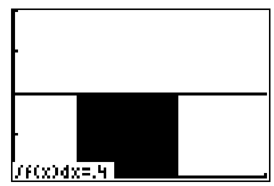
Press ENTER after typing .5.



When new screen appears, type 1.3.



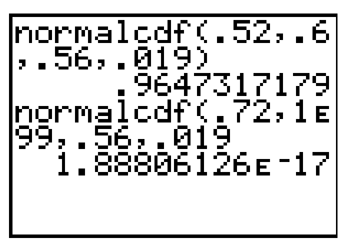
Press ENTER after typing 1.3.



$P(0.5 \le Y \le 1.3)$ is shown as 0.4.

We conclude this section by working an exercise using the normal density curve that reviews the normal distribution calculations from Section 1.3 of this manual.

Example 4.8 Voting in Oregon. After an election in Oregon, voter records showed that 56% of registered voters actually voted. A survey of 663 registered voters is conducted and the sample proportion $\hat{p}$ of those who claim to have voted is obtained. For all random samples of size 663, these values of the sample proportions $\hat{p}$ will follow an approximate normal distribution with mean μ = 0.56 and standard deviation σ = 0.019. Use this distribution to compute $P(0.52 < \hat{p} < 0.60)$ and $P(\hat{p} \ge 0.72)$.



Solution. We use the `normalcdf(` command from the `DISTR` menu. For $\hat{p} \sim N(.56,.019)$, we can find $P(0.52 < \hat{p} < 0.60)$ = using the command `normalcdf(.52,0.6,.56,.019)`. To find $P(\hat{p} \ge 0.72)$, use `normalcdf(.72,1E99,.56,.019)`, although practically any large positive value (such as

999999) would do. We find there is a 96.5% chance that between 52% and 60% would say they had voted. There is practically no chance that more than 72% would say they had voted.

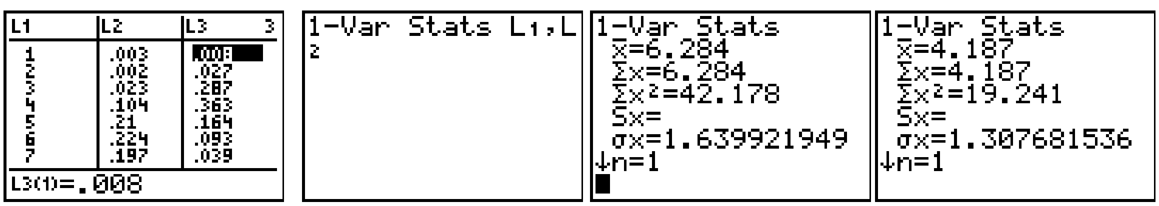
4.4 Means and Variances of Random Variables

We now show how to compute the mean and standard deviation of a discrete random variable for which the range of measurements and corresponding probabilities are given.

Example 4.9 More About Rooms. The table below gives the distributions of the number of rooms for owner-occupied units and renter-occupied units in San Jose, California, which we saw earlier. We want to calculate the mean and the standard deviation of the number of rooms for each type.

Rooms	1	2	3	4	5	6	7	8	9	10
Owned	0.003	0.002	0.023	0.104	0.210	0.224	0.197	0.149	0.053	0.035
Rented	0.008	0.027	0.287	0.363	0.164	0.093	0.039	0.013	0.003	0.003

Solution. If they are not still there, enter the measurements (rooms) into list `L1` and the probabilities into lists `L2` and `L3`. For the owner-occupied units, enter the command `1-Var Stats L1,L2`. We see that the average number of rooms for owner-occupied units is μ = 6.284 with a standard deviation of σ = 1.64. For the renter-occupied units, enter the command `1-Var Stats L1,L3` to obtain μ = 4.187 and σ = 1.3077. Remember, since this is a probability distribution, we want parameter values. Notice the calculator is smart enough to skip the sample standard deviations!

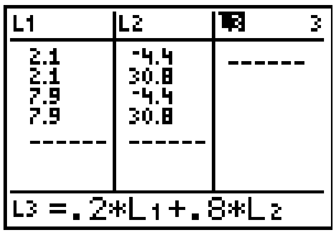


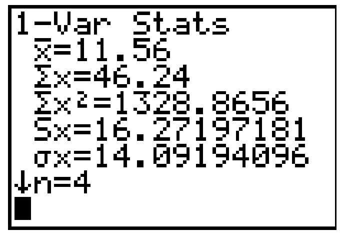
Mean and Standard Deviation of an Independent Sum

Here we show how to verify the mean and standard deviation of a random variable of the form $Z = aX + bY$, where X and Y are independent.

Example 4.10. Suppose that $Z = 0.2X + 0.8Y$, where μ_X = 5, $\sigma_X = 2.9$, $\mu_Y = 13.2$, and $\sigma_Y = 17.6$. Assuming that X and Y are independent, find the mean and standard deviation of Z.

Solution. We can think of X as taking two values $5-2.9$ = 2.1 and $5+2.9$ = 7.9. Likewise, we can consider Y to assume only the values $13.2-17.6$ = –4.4 and $13.2+17.6$ = 30.8. We enter these values of X and Y into lists `L1` and `L2`; however, we list each X value twice in consecutive fashion (2.1, 2.1, 7.9, 7.9), and we list the Y values twice in alternating fashion (–4.4, 30.8, –4.4, 30.8). Next, we enter the command `0.2*L1 + 0.8*L2` while name `L3` is highlighted to send the possible values of Z to list `L3`. Finally, enter `1-Var Stats L3` to compute the mean and standard deviation of Z.





We thereby can verify that $\mu_Z = 0.2\mu_X + 0.8\mu_Y$ and, because X and Y are independent, that $\sigma_Z = \sqrt{0.2^2\sigma^2{}_X + 0.8^2\sigma^2{}_Y}$. However, if $Z = aX + bY$ and X and Y are *not* independent, then $\sigma_Z = \sqrt{a^2\sigma^2{}_X + b^2\sigma^2{}_Y + 2\rho a\sigma_X b\sigma_y}$.

4.5 General Probability

We conclude this chapter with a short program for the Law of Total Probability and Bayes's rule. The **BAYES** program given below computes and displays the total probability $P(C)$ according to the formula

$$P(C) = P(A_1)P(C \mid A_1) + P(A_2)P(C \mid A_2) + ... + P(A_n)P(C \mid A_n)$$

Before executing the **BAYES** program, enter given probabilities $P(A_1)$, . . . , $P(A_n)$ into list `L1` and the given conditionals $P(C \mid A_1),..., P(C \mid A_n)$ into list `L2`.

The program also stores the probabilities of the intersections, $P(C \cap A_1),..., P(C \cap A_n)$ in list `L3`, and stores the reverse conditionals $P(A_1 \mid C),..., P(A_n \mid C)$ in list `L4`. Finally, the conditional probabilities $P(A_1 \mid C'),..., P(A_n \mid C')$ are stored in list `L5`, and the conditional probabilities $P(C \mid A_1'),..., P(C \mid A_n')$ are stored in list `L6`. To help the user keep track of which list contains which probabilities, the program displays a description upon completion.

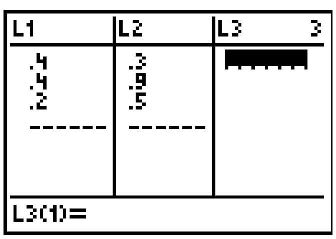
The program, with appropriate list name changes, will also run on a TI-89.

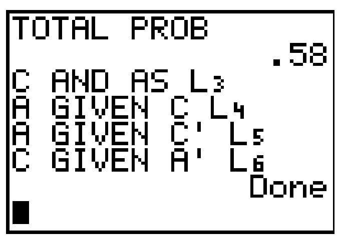
The BAYES Program

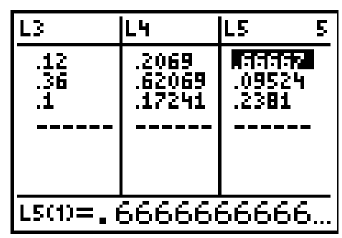
```
PROGRAM:BAYES
:L1*L3↵L3
:sum(L3) ↵T
:L3/T↵L4
:L1*(1-L2)/(1-T) ↵L5
:T*(1-L4)/(1-L1)↵L6
:Disp "TOTAL PROB"
:Disp round(T,4)
:Disp "C AND As : L3"
:Disp "As GIVEN C : L4"
:Disp "As GIVEN C' : L5"
:Disp "C GIVEN A's : L6"
```

Example 4.11 Will He Win? The voters in a large city are 40% white, 40% black, and 20% Hispanic. A mayoral candidate expects to receive 30% of the white vote, 90% of the black vote, and 50% of the Hispanic vote. Apply the **BAYES** program to compute the percent of the overall vote that the candidate expects, and to analyze the other computed conditional probabilities.

Solution. Here we let A_1 = white voters, A_2 = black voters, and A_3 = Hispanic voters. We enter the probabilities of these events into list `L1`. We let C be the event that a person votes for the candidate. Then $P(C \mid A_1)$ = 0.30, $P(C \mid A_2)$ = 0.90, and $P(C \mid A_3)$ = 0.50, and we enter these conditional probabilities into list `L2`. Then we execute the **BAYES** program.







We first see that the candidate can expect to receive 58% of the overall vote. This result is obtained by $P(C) = .4*.3 + .4*.9 + .2*.5 = 0.58$. The other computed probabilities are also stored in the designated lists.

List `L3` contains the probabilities of the intersections. The probability that a voter is white and will vote for the candidate is $P(A_1 \cap C)$ = 0.12. The probability that a voter is black and will vote for the candidate is $P(A_2 \cap C)$ = 0.36. The probability that a voter is Hispanic and will vote for the candidate is $P(A_3 \cap C)$ = 0.1. These values are obtained with the formula $P(A_i \cap C) = P(A_i)P(C \mid A_i)$, and are found by multiplying the terms in lists `L1` and `L2`.

List `L4` contains the reverse conditional probabilities of being white, black, Hispanic given that one *will* vote for the candidate. This list is the direct application of Bayes's rule. These conditional probabilities are $P(A_1 \mid C)$ = 0.2069, $P(A_2 \mid C)$ = 0.62069, and

$P(A_3 \mid C)$ = 0.17241, respectively. This list is obtained by dividing the respective intersection probabilities in list L3 by $P(C)$ = 0.58.

List L5 contains the reverse conditional probabilities of being white, black, Hispanic given that one will *not* vote for the candidate. These values are $P(A_1 \mid C')$ = 0.66667, $P(A_2 \mid C')$ = 0.09524, and $P(A_3 \mid C')$ = 0.2381, respectively.

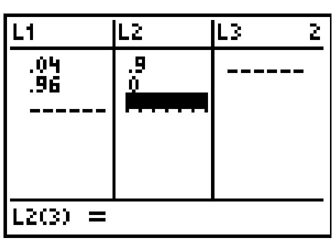
List L6 (not shown) contains the conditional probabilities of voting for the candidate given that a voter is *not* white, *not* black, and *not* Hispanic. These values are $P(C \mid A_1')$ = 0.76667, $P(C \mid A_2')$ = 0.36667, and $P(C \mid A_3')$ = 0.6, respectively.

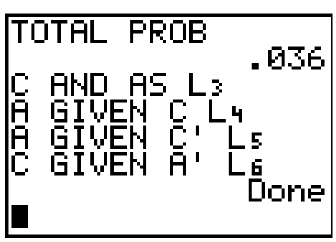
We note that given any conditional probability $P(C \mid D)$, then the complement conditional probability is given by $P(C' \mid D) = 1 - P(C \mid D)$. Thus, lists for complement conditional probabilities do not need to be generated. For example, the respective conditional probabilities of *not* voting for the candidate given that one is white, black, and Hispanic are respectively 0.70, 0.10, and 0.50. These values are the complement probabilities of the originally given conditionals.

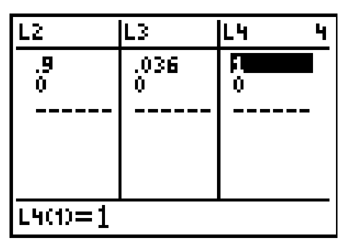
Example 4.12 Cystic Fibrosis. The probability that a randomly chosen person of European ancestry carries an abnormal CF gene is 1/25. If one is a carrier of this gene, then a test for it will be positive 90% of the time. If a person is not a carrier, then the test will never be positive. Jason tests positive. What is the probability that he is a carrier?

Solution. We let C be the event that a person of European ancestry tests positive, let A_1 be carriers of the gene, and let A_2 be noncarriers. Because $P(A_1)$ = 1/25, then $P(A_2)$ = 24/25. Also, $P(C \mid A_1)$ = 0.90, $P(C \mid A_2)$ = 0. We are trying to compute $P(A_1 \mid C)$, which is the probability of being a carrier given that one has tested positive.

We enter the probabilities $P(A_1)$ and $P(A_2)$ into list L1, enter the conditional probabilities $P(C \mid A_1)$ and $P(C \mid A_2)$ into list L2, and execute the **BAYES** program.





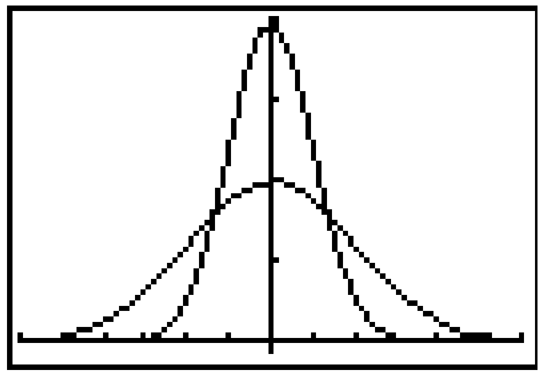


List L4 contains the conditional probabilities $P(A_i \mid C)$. From the first entry, we see that

$P(A_1 \mid C)$ = 1. Because Jason has tested positive, there is a 100% chance that he is a carrier. (He must be a carrier because it is impossible for noncarriers to test positive.) This value is also given by $P(A_1 \mid C) = \frac{P(A_1 \cap C)}{P(C)} = \frac{1/25 * 0.90}{1/25 * .90 + 24/25 * 0} = 1$.

CHAPTER

5



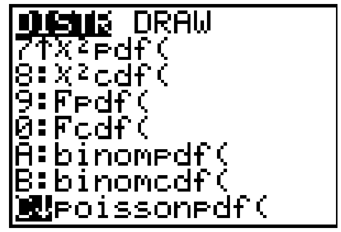
Sampling Distributions

Introduction

In this chapter, we show how to compute probabilities involving binomial distributions, Poisson distributions, and the sample mean $\overline{x}$.

5.1 Sampling Distributions for Counts and Proportions

We begin by demonstrating how to compute various probabilities for a given binomial distribution. To do so, we will need the `binompdf(` and `binomcdf(` commands from the `DISTR` menu. Actual location of these depends on the calculator model.

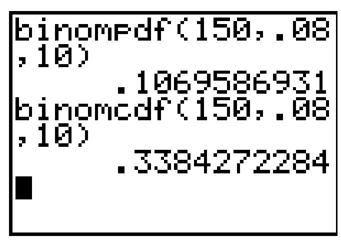


Binomial Probabilities

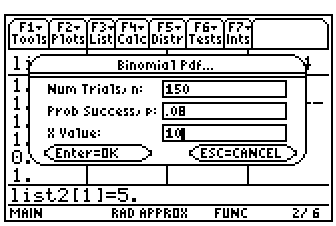
For a binomial distribution $X \sim B(n, p)$, we compute the probability of exactly k successes, $P(X = k)$, by entering the command `binompdf(n,p,k)`. The probability $P(X \le k) = P(0 \le X \le k)$ of at most k successes is computed with the command `binomcdf(n,p,k)`. The probability of there being at least k successes is given by $P(X \ge k) = 1 - P(X \le k-1)$, and is computed with the command `1-binomcdf(n,p, k-1)`. The following three examples demonstrate these calculations. TI-89 calculators explicitly ask for the low and high ends of interest to be included in the calculation, so there is no need for any subtraction.

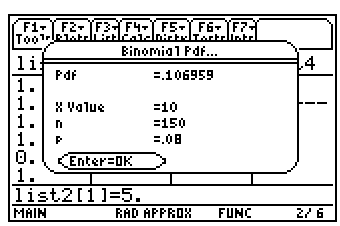
Example 5.1 Auditing Sales. An audit examines a simple random sample of 150 out of 10,000 available sales records. Suppose that in fact, 800 of the 10,000 sales are incorrectly classified. What is the probability we find *exactly* 10 misclassified records? What is the probability we find *at most* 10 misclassified records? Since 800/10000 = 0.08, we let X = the number of misclassified records, and $X \sim$ B(150, 0.08). Calculate $P(X = 10)$ and $P(X \le 10)$.

TI-83/84 Solution. Simply enter the commands `binompdf(150,.08,10)` and `binomcdf(150,.08,10)` to obtain $P(X = 10) \approx 0.1070$ and $P(X \le 10) \approx 0.3384$. As always, ask your instructor how many decimal places to report.

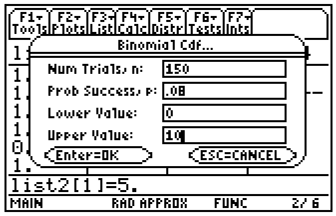


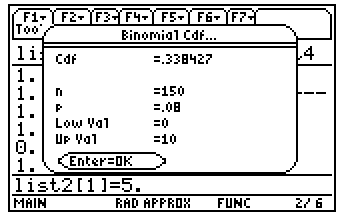
TI-89 Solution. In the Stats/List Editor application, press □ for the Distr menu. Arrow down to find Binomial Pdf. Press ⊆ to select it. You will have a dialog box like the one below that explicitly asks for n, p, and k. Press ⊆ to perform the calculation.





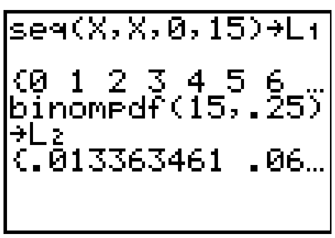
`Binomial Cdf` is similar, however. In this case, we specify the low and high ends of interest.

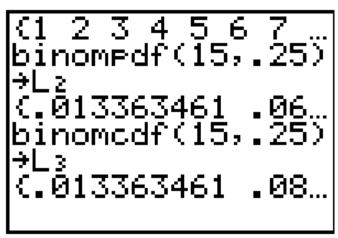


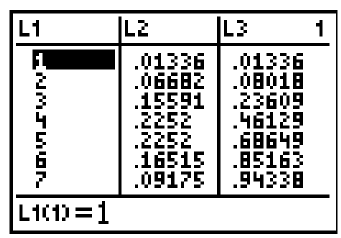


Example 5.2 Guessing on a Quiz. Suppose a multiple choice quiz has 15 questions, each with 4 possible answers. If a student is purely guessing, what is the probability of at most one correct answer? Let $X \sim B(15, 0.25)$. Make a probability table and probability histogram of the distribution. Also make a table of the cumulative distribution and use it to find $P(X \leq 1)$**.**

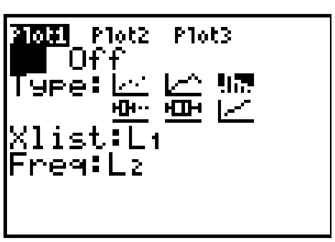
Solution. Because there are $n = 15$ attempts, the possible number of successes range from 0 to 15. So we first enter the integer values 0, 1, . . . , 15 into list `L1`. We can do so directly or we can use the command `seq(X,X,0,15)` from the `LIST` menu with list name `L1` highlighted by the cursor. Next, we use the commands `binompdf(15,.25)↵L2` and `binomcdf(15,.25)↵L3` to enter the probability distribution values $P(X = k)$ into list `L2` and the cumulative distribution values $P(X \leq k)$ into list `L3`. If you are using a TI-89, enter all these commands with the receiving list name highlighted by the cursor. From the values in list `L3`, we can see that $P(X \leq 1) \approx 0.0802$.

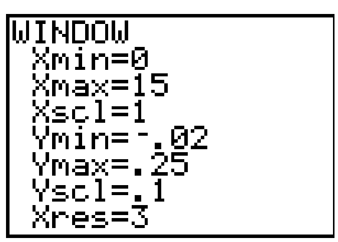


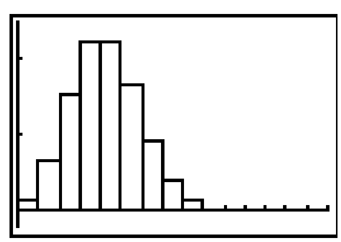




To see a probability histogram of the distribution, adjust the `STAT PLOT` settings for a histogram of `L1` with frequencies `L2`. We note that the probabilities beyond $X = 9$ will not be observable, because those probabilities are so small. Once again, for this histogram, since we want bars of width 1, the θ→ picture will need adjusting for an `Xscl` of 1. I have shown my final `WINDOW` settings below.

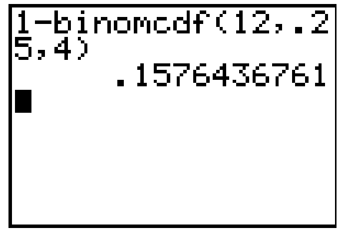






Example 5.3 Free Throws. Suppose a basketball player makes 75% of his free throws. In one particular game, he missed 5 of 12 attempts. Is it unusual to perform this poorly? We let X = number of missed free throws, so X ~ B(12,.25), and compute $P(X \geq 5)$.

Solution. We use the probability of the complement to obtain $P(X \geq 5) = 1 - P(X \leq 4)$, which is computed using `1-binomcdf(12,.25,4)` ≈ 0.1576. If you are using a TI-89, simply enter a low end of 5 and a high end of 12.



Probabilities for $\hat{p}$

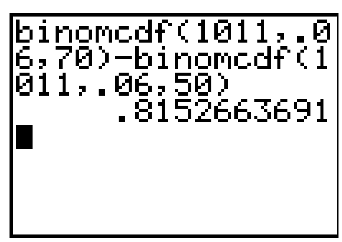
The next two examples show how to make probability calculations for a sample proportion $\hat{p}$ by converting to a binomial probability.

Example 5.4 Clothes Shopping. Suppose that 60% of all adults agree that they like shopping for clothes, but often find it frustrating and time-consuming. In a nationwide sample of 2500 adults, let $\hat{p}$ be the sample proportion of adults who agree with this response. Compute $P(\hat{p} \geq 0.58)$, the chance that more than 58% in your sample will agree that clothes shopping can be frustrating and time-consuming.

Solution. Because 58% of 2500 is 1450, we must compute $P(X \geq 1450)$, where $X \sim B(2500, 0.60)$. Instead, we may compute $1 - P(X \leq 1499)$ using `1-binomcdf(2500,.6,1499)` ≈ 0.5087.

Example 5.5 Betting on Football. A Gallup poll of size n = 1011 found 6% of the respondents said they bet on college football. Assuming a true proportion of p = 0.06, what is the probability that a sample proportion $\hat{p}$ lies between 0.05 and 0.07?

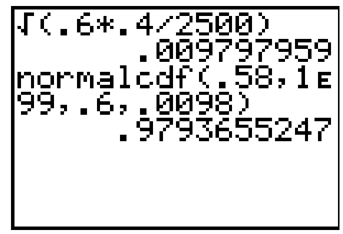
Solution. For n = 1011 and p = 0.06, then $P(0.05 < \hat{p} < 0.07)$ = $P(.05 * n < X < .07 * n)$ = $P(50.55 < X < 70.77)$. Because there can't be fractions of "successes" this becomes $P(51 \leq X \leq 70) = P(X \leq 70) - P(X \leq 50)$, where $X \sim$ B(1011, 0.06). We find this value by entering `binomcdf(1011,.06,70)–binomcdf(1011,.06,50)`, and obtain a probability of about 0.8153.



Normal Approximations

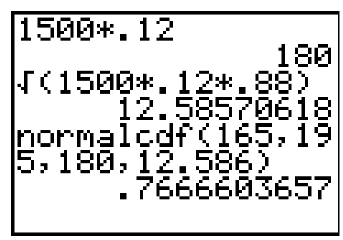
We conclude this section by showing how to approximate a sample proportion probability and a binomial probability with a normal distribution.

Example 5.6 More on Clothes Shopping. With n = 2500 and p = 0.60 as in Example 5.4 above, use the approximate distribution of $\hat{p}$ to estimate $P(\hat{p} > 0.58)$.



Solution. The distribution of $\hat{p}$ is approximately normal with $\mu = p = 0.60$ and
$\sigma = \sqrt{p(1-p)/n} = \sqrt{.6*.4/2500} = 0.0098$.
Thus, $P(\hat{p} \geq 0.58) \approx P(Y \geq 0.58)$, where Y ~ N(0.60, 0.0098).
The command `normalcdf(.58,1E99,.6,.0098)` gives a probability of about 0.9794.

Example 5.7 Checking for Survey Errors. One way of checking the effect of undercoverage, nonresponse, and other sources of error in a sample survey is to compare the sample with known facts about the population. About 12% of American adults are black. The number X of blacks in a random sample of 1500 should therefore be binomial with n = 1500 and p = 0.12. (a) What are the mean and standard deviation of X?
(b) Use the Normal approximation to find the chance there are between 165 and 195 blacks in our survey.



Solution. (a) The mean is $\mu = np = 1500*0.12 = 180$, and the standard deviation is
$\sigma = \sqrt{np(1-p)} = \sqrt{1500*.12*.88} = 12.586$.

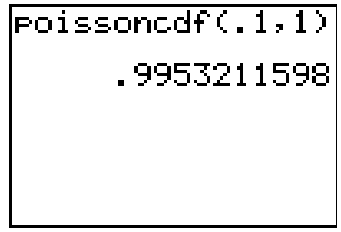
(b) We now let $Y \sim N(180, 12.586)$. Then $P(165 < X < 195)$ is found using `normalcdf(165,195,180,12.586)`. We see that the desired probability ≈ 0.7667.

5.2 Poisson Random Variables

Poisson random variables generally come from one of two situations. We have a binomial random variable with large n and small p so that the expected number of "successes" is small, or we have a process, such as telephone calls to a switchboard, where we only observe the successes. Poisson random variables have only one parameter, λ , the average "rate" of the process. In the binominal setting, $\lambda = np$.

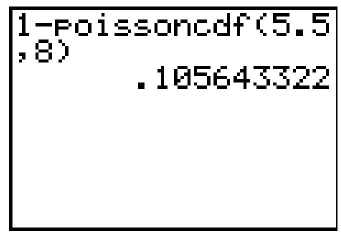
Example 5.8 Mumps Outbreaks. Mumps is an acute viral infection that is generally mild, and in about 20% of infected individuals, even asymptomatic (meaning they do not know they have the disease). However, severe complications can arise. Mandatory vaccinations have largely eradicated the disease in the United States. For the state of Iowa, the average monthly number of reported cases is about 0.1 per month. Assuming that cases are independent, what is the probability that in a given month, there will be no more than one case of mumps in Iowa?

Solution. The rate of the process is $\lambda = 0.1$ per month. We let X be the number of cases in a month and want to know $P(X \leq 1)$. Just as in the binomial case, this is a cumulative probability (the probability of either 0 or 1 case of mumps), so we use Poissoncdf(from the DISTR menu. Parameters are λ and k. The probability of at most one case of mumps per month in Iowa is about 0.9953.



Example 5.9 ATM Use. Suppose the number of people who use a certain ATM machine can be modeled as a Poisson process. The average number of users is 5.5 per hour. What is the probability that in a given hour, more than 8 will use the machine?

Solution. The rate of the process is $\lambda = 5.5$ per hour. If X is the number of users of the machine in an hour, we want to find $P(X > 8)$. Just as with binomial probabilities, we find this probability using the complements rule as $1 - P(X \leq 8)$. The probability of more than 8 ATM users in an hour for this machine is about 10.56%.



5.3 The Sampling Distribution of a Sample Mean

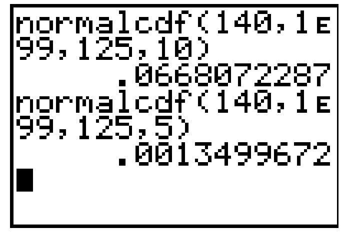
We now show how to compute various probabilities involving the sample mean $\bar{x}$. To do so, we make use of the fact that for random samples of size n from a $N(\mu, \sigma)$ distribution, the sample mean $\bar{x}$ follows a $N(\mu,\ \sigma/\sqrt{n})$ distribution. According to the Central Limit Theorem, when samples are "large," $\bar{x}$ also follows a $N(\mu,\ \sigma/\sqrt{n})$ distribution.

Example 5.10 Measuring Blood Glucose. Sheila's glucose level one hour after ingesting a sugary drink varies according to the Normal distribution with μ = 125 mg/dl and σ = 10 mg/dl.

(a) If a single glucose measurement is made, what is the probability that Sheila measures above 140?
(b) What is the probability that the sample mean from four separate measurements is above 140?

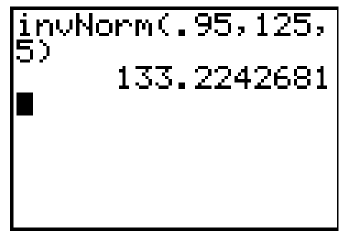
Solution. (a) We compute $P(X > 140)$ for $X \sim N(125,10)$ with the command `normalcdf(140,1E99,125,10)`. Then, $P(X > 140) \approx 0.0668$.

(b) For an SRS of size $n = 4$, $\bar{x}$ has a mean of $\mu = 125$ and a standard deviation of $\sigma/\sqrt{n} = 10/\sqrt{4} = 5$. So now we compute $P(\bar{x} > 140)$ for $\bar{x} \sim N(125, 5)$ and obtain a value of about 0.00135.



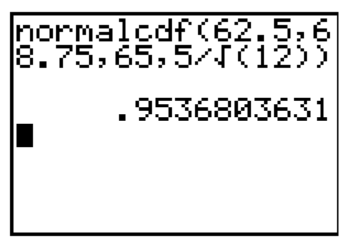
Example 5.11 More Blood Glucose. Sheila's glucose level one hour after ingesting a sugary drink varies according to the normal distribution with $\mu = 125$ mg/dl and $\sigma = 10$ mg/dl. What is the level L such that there is only 0.05 probability that the mean glucose level of four test results falls above L for Sheila's glucose level distribution?

Solution. As in the previous example, $\bar{x} \sim N(125, 5)$. So we must find the inverse normal value L for which $P(\bar{x} > L) = 0.05$ or, equivalently, $P(\bar{x} \le L) = 0.95$. We compute this value with the `invNorm(` command from the `DISTR` menu by entering `invNorm(.95,125,5)`. We see that only about 5% of the time should $\bar{x}$ be larger than $L = 133.224$.



Example 5.12 Egg Weights. The weight of eggs produced by a certain breed of hen is normally distributed with a mean of 65 g and a standard deviation of 5 g. For random cartons of 12 eggs, what is the probability that the weight of a carton falls between 750 g and 825 g?

Solution. If the total weight of 12 eggs falls between 750 g and 825 g, then the sample mean $\bar{x}$ falls between 750/12 = 62.5 g and 825/12 = 68.75 g. So, we compute $P(62.5 < \bar{x} < 68.75)$ for $\bar{x} \sim N(65, 5/\sqrt{12})$ using the command `normalcdf(62.5,68.75,65,5/√(12))`. Note: If you are using a TI-89, you will first have to find $\sigma(\bar{x}) = 5/\sqrt{12} = 1.443$.

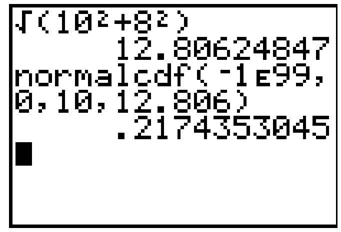


Sum of Independent Normal Measurements

Let $X \sim N(\mu_X, \sigma_X)$ and let $Y \sim N(\mu_Y, \sigma_Y)$. Assuming that X and Y are independent measurements, then $X \pm Y$ follows a $N\left(\mu_X \pm \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$ distribution.

Example 5.13 A Golf Competition. Tom and George are playing in the club golf tournament. Their scores vary as they play the course repeatedly. Tom's score X has the $N(110, 10)$ distribution and George's score Y has a $N(100, 8)$ distribution. If the scores are independent, then what is the probability that Tom's score is less than George's?

Solution. The value $P(X < Y)$ is equivalent to $P(X - Y < 0)$. So we need to use the distribution of the difference $X - Y$ which is $N(110 - 100, \sqrt{10^2 + 8^2}) = N(10, 12.806)$. Using the command `normalcdf(⁻1E99,0,10,12.806)`, we find that $P(X - Y < 0) \approx 0.2174$. Although George has a lower average, Tom should beat him about 21% of the time.



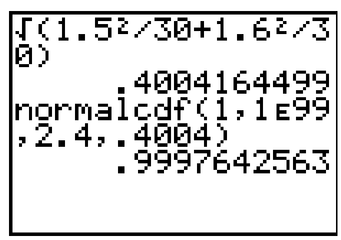
Sum and Difference of Sample Means

Let $\bar{x}$ be the sample mean from an SRS of size n from a $N(\mu_X, \sigma_X)$ distribution, and let $\bar{y}$ be the sample mean from an independent SRS of size m from a $N(\mu_Y, \sigma_Y)$ distribution. Then, the sum/difference $\bar{x} \pm \bar{y}$ follows a $N\left(\mu_X \pm \mu_Y, \sqrt{\sigma_X^2 / n + \sigma_Y^2 / m}\right)$ distribution.

Example 5.14 Credible Sources. In a randomized comparative experiment, students read ads that cited either the *Wall Street Journal* or the *National Enquirer*. They were asked to rate the trustworthiness of the source on a 7 point scale. Let $\bar{y}$ be the sample mean from the group who read the ad citing the *Journal* and $\bar{x}$ be the sample mean from a group of size 30 who read the ad citing the *Enquirer*. Suppose the population of all student scores for the *Journal* have a $N(4.8, 1.5)$ population, and the population of all student scores for the *Enquirer* are $N(2.4, 1.6)$. What is the distribution of $\bar{y} - \bar{x}$? Find $P(\bar{y} - \bar{x}) \geq 1$.

Solution. First,

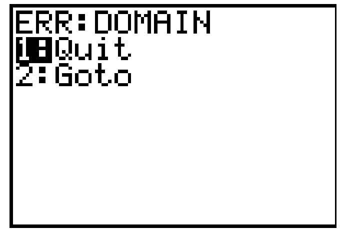
$\bar{y} - \bar{x} \sim N(4.8 - 2.4, \sqrt{\frac{1.5^2}{30} + \frac{1.6^2}{30}}) = N(2.4, 0.4004)$. We find $P(\bar{y} - \bar{x}) \geq 1$ using `normalcdf(1,1E99,2.4,0.4004)` and obtain a value of 0.9998. It's virtually certain the *Journal* will have a higher mean score at least one point higher than the *Enquirer*.



5.4 Common Errors

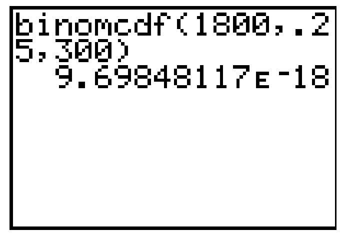
Err:Domain

This error is normally caused in these types of problems by specifying a probability as a number greater than 1 (in percent possibly instead of a decimal) or a value for *n* or *x* that is not an integer. Reenter the command giving *p* in decimal form. Pressing N will return you to the input screen to correct the error. This will also occur in older TI-83 calculators if *n* is too large in a binomial calculation; in that case, you need to use the normal approximation.



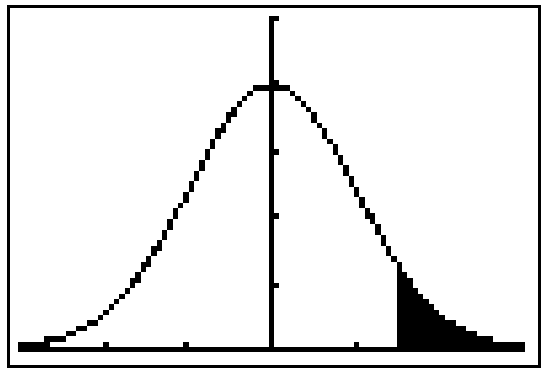
Probabilities larger than 1

This is not possible. As we've said before, if it appears more than 1 on the first glance, check the right hand side. This value is 9.7 x 10^{-18} or 17 0s followed by the leading 9.



CHAPTER

6



Introduction to Inference

Introduction

In this chapter, we show how to use TI calculators to compute confidence intervals and conduct hypothesis tests for the mean μ of a normally distributed population. We begin our study of inference with the unrealistic assumption that the population standard deviation, σ, is known. This assumption will be relaxed in Chapter 7.

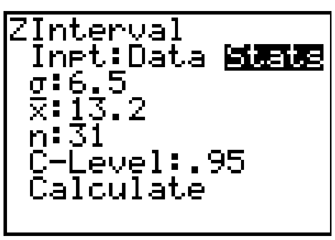
6.1 Confidence Intervals with σ Known

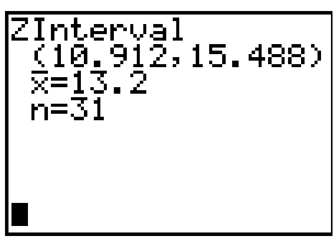
In this section, we show how to compute a confidence interval for the mean of a normal population with known standard deviation σ. To do so, we will use the built-in `ZInterval` feature (item 7) from the `STAT TESTS` menu (item 1 on the □ `Ints` menu on a TI-89). The following two exercises demonstrate how to use this feature with summary statistics and with a data set.

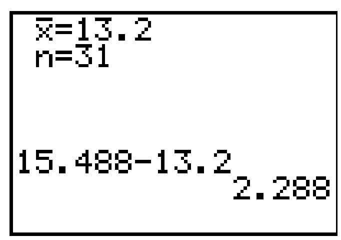
Example 6.1 Bone Turnover. In a study of bone turnover in young women, serum TRAP (a measure of bone resorption) was measured in 31 subjects and the mean was 13.2 Units/liter. Assume that the standard deviation is known to be 6.5 Units/liter. Give the margin of error and find a 95% confidence interval for the mean of all young women represented by this sample.

TI-83/84 Solution. Select the `ZInterval` option, and set `Inpt` to `Stats` by moving the cursor to highlight this option and pressing ÷, since we do not have the actual data entered in a list. Enter the given values of 6.5 for σ, 13.2 for $\bar{x}$, and 31 for *n*. Enter the desired confidence level, then press ÷ with `Calculate` highlighted. We obtain a 95% confidence interval of (10.912, 15.488). We normally report one more significant digit than was in the information provided, but check with your instructor for his or her rounding rules. Based on this information, I am 95% confident the mean TRAP level for young women represented by this sample is between 10.92 and 15.49 Units/liter.

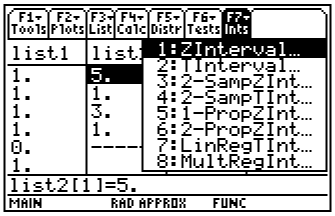
Because the confidence interval is of the form $\bar{x} \pm m$, we can find the margin of error *m* by subtracting $\bar{x}$ from the right endpoint of the interval: 15.488 – 13.2 = 2.288.

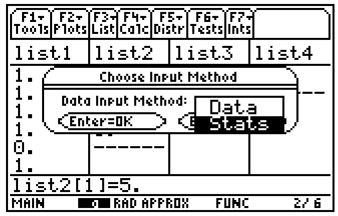


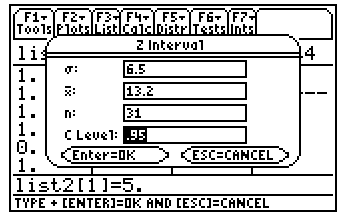




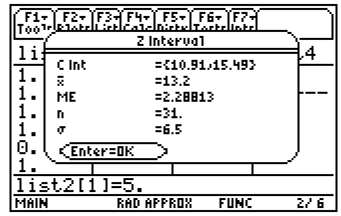
TI-89 Solution. Select `1:ZInterval` from the `Ints` menu. You first see a screen asking for the input method. Press the right arrow to expand the options and move the cursor to the appropriate choice. Here we have Stats, so that should be highlighted. Press ÷ to proceed to the next dialog box and enter the summary statistics given above.







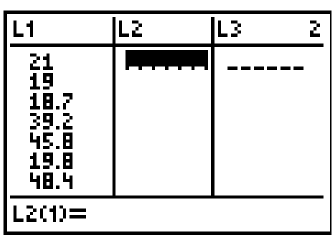
When finished, press ÷ to compute and display the interval. Notice this calculator gives the margin of error explicitly as ME = 2.28813.

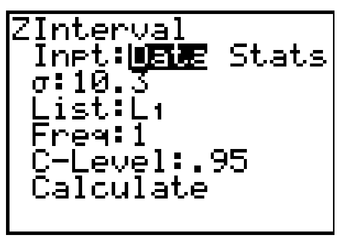


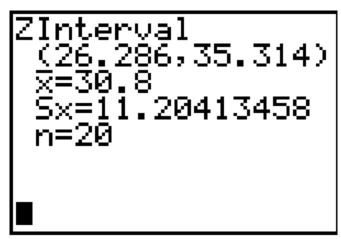
Example 6.2 Fuel Efficiency. Here are the values of the average speed (in mph) for a sample of trials on a vehicle undergoing a fuel efficiency test. Assume that the standard deviation is 10.3 mph. Estimate the mean speed at which the vehicle was driven with 95% confidence.

21.0	19.0	18.7	39.2	45.8	19.8	48.4	21.0	29.1	35.7
31.6	49.0	16.0	34.6	36.3	19.0	43.3	37.5	16.5	34.5

Solution. First, enter the data into a list, say list `L1`. Next, bring up the `ZInterval` screen, set `Inpt` to `Data`, and enter the given value of 10.3 for σ. Set `List` to `L1` with `Freq` 1 (since each value represents one observation), enter the desired confidence level, and press ÷ on `Calculate`. Based on this information, I am 95% confident the mean speed was between 26.29 and 35.31 miles per hour.







Choosing the Sample Size

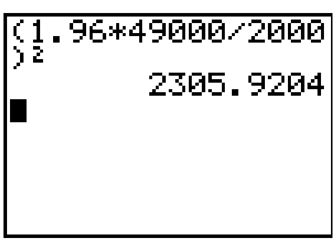
Suppose we want to find the minimum sample size n that will produce a desired margin of error m with a specific level of confidence. To do so, we can use the formula $n \geq \left(\frac{z^* \sigma}{m} \right)^2$, where z^* is the appropriate critical value. We also could use a program that computes the (rounded-up) sample size. To execute the **ZSAMPSZE** program that follows, we simply input the values of the standard deviation σ, the desired margin of error, and the desired confidence level in decimal.

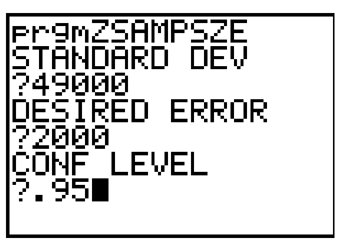
The ZSAMPSZE Program

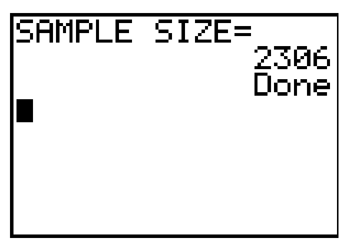
```
PROGRAM:ZSAMPSZE
:Disp "STANDARD DEV."
:Input S
:Disp "DESIRED ERROR"
:Input E
:Disp "CONF. LEVEL"
:Input R
:invNorm((R+1)/2,0,1)↵Q
:(Q*S/E)ϒ ↵M
:If int(M)=M
:Then
:M↵N
:Else
:int(M+1) ↵N
:End
:ClrHome
:Disp "SAMPLE SIZE="
:Disp int(N)
```

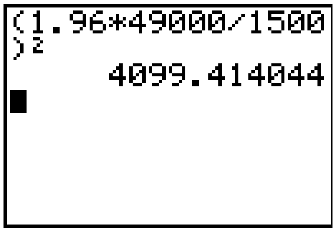
Example 6.3 Student Debt. Suppose we want a margin of error of $2000 with 95% confidence when estimating the mean debt for students completing their undergraduate studies. We have information that the standard deviation is about $49,000. (a) What sample size is required? (b) What sample size would be required to obtain a margin of error of $1500?

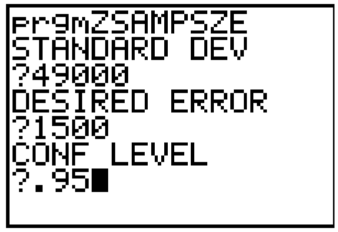
Solution. The critical value for 95% confidence is $z^*= 1.96$. Using this value in the formula $n \geq \left(\frac{1.96 * 49000}{2000} \right)^2$, with σ = 49,000 and m = 2000, we obtain a necessary sample size of n = 2306. The same result is obtained by using the **ZSAMPSIZE** program. Working part (b) similarly with m = 1500, we obtain a required sample size of n = 4100.

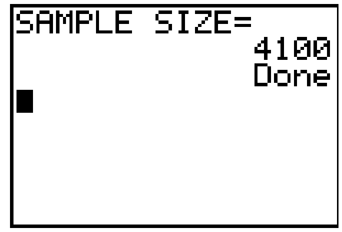










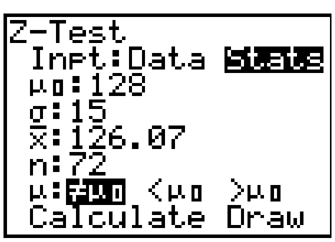


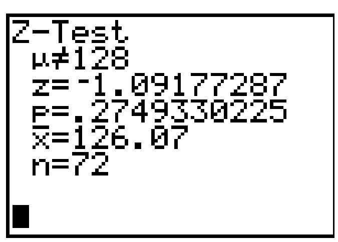
6.2 Tests of Significance

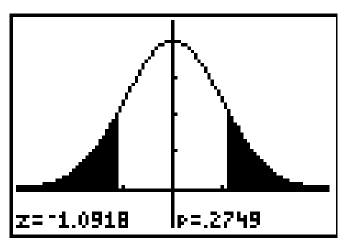
We now show how to perform one-sided and two-sided hypothesis tests about the mean μ of a normally distributed population for which the standard deviation σ is known. To do so, we will use the `Z-Test` feature (item 1) from the `STAT TESTS` menu. This menu is found using 2□=□ in the Stats/List Editor on a TI-89. We can use this feature to work with either summary statistics or data sets.

Example 6.4 Executives' Blood Pressure. The mean systolic blood pressure for males 35 to 44 years of age is 128 and the standard deviation is 15. But for a sample of 72 company executives in this age group, the mean systolic blood pressure is $\bar{x}$ = 126.07. Is this evidence that the company's executives in this age group have a different mean systolic blood pressure from the general population?

Solution. To test if the mean is *different* from 128, we use the null hypothesis H_0: μ = 128 with a two-sided alternative H_A: $\mu \neq 128$. Bring up the `Z-Test` screen and adjust the `Inpt` to `STATS`, which allows us to enter the statistics. Enter the values μ_0 = 128, σ = 15, $\bar{x}$ = 126.07, and `n` = 72. Set the alternative to $\neq \mu_0$ then press ⊆ on either `Calculate` or `Draw`.



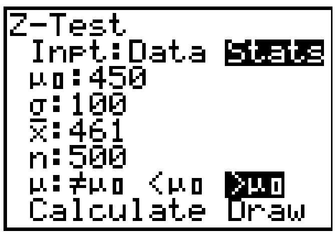


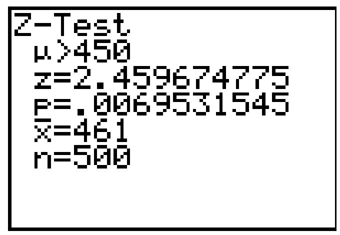


We obtain a *z* test statistic of –1.0918 and a *p*-value of 0.2749. For this two-sided test, the *p*-value comes from the sum of both the right- and left-tail probabilities: $P(z < -1.0918) + P(z > 1.0918)$. If the true mean for all the company's executives in this age group were equal to 128, then there would be a 27.49% chance of obtaining an $\bar{x}$ as far away as 126.07 with a sample of size 72. This rather high *p*-value does not give us good evidence to reject the null hypothesis.

Example 6.5 California SATs. An SRS of 500 California high school seniors gave an average SAT mathematics score of $\bar{x}$ = 461. Is this good evidence against the claim that the mean for all California seniors is no more than 450? Assuming that σ = 100 for all such scores, perform the test H_0: μ = 450, H_A: μ > 450. Give the *z* test statistic and the *p*-value.

Solution. Bring up the `Z-Test` screen from the `STAT TESTS` menu and check that `Inpt` is `STATS`. Enter the value of μ_0 = 450 and the summary statistics, set the alternative to $>\mu_0$, then scroll down to `Calculate` and press ⊆.



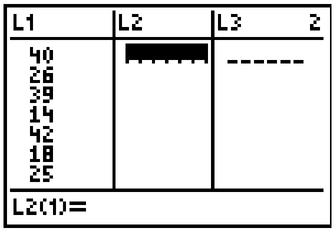


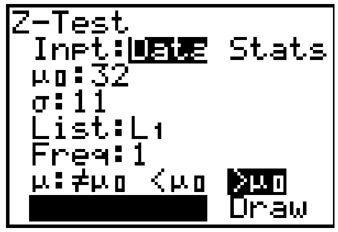
We obtain a *z* test statistic of 2.46 and a *p*-value of 0.00695. Because the *p*-value is so small, we have significant evidence to reject H_0. If the true mean were 450, then there would be only a 0.00695 probability of obtaining a sample mean as high as $\bar{x}$ = 461 with an SRS of 500 students.

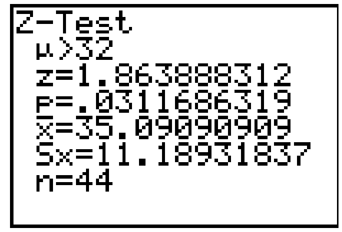
Example 6.6 DRP Scores. The following table gives the DRP scores for a sample of 44 third-grade students in a certain district. It is known that σ = 11 for all such scores in the district. A researcher believes that the mean score of all third-graders in this district is higher than the national mean of 32. State the appropriate H_0 and H_A, then conduct the test and give the *p*-value.

40	26	39	14	42	18	25	43	46	27	19
47	19	26	35	34	15	44	40	38	31	46
52	25	35	35	33	29	34	41	49	28	52
47	35	48	22	33	41	51	27	14	54	45

Solution. Here we test H_0: μ = 32 with a one-sided alternative H_A: μ > 32. Enter the data into a list, say list `L1`, then call up the `Z-Test` screen and change `Inpt` to `Data`. Enter the values μ_0 = **32** and σ = **11**, set the list to `L1` with frequencies 1, and set the alternative to $>\mu_0$. Press ⊆ on `Calculate` or `Draw`.







We obtain a p-value of 0.0312. If the average of the district were equal to 32, then there would be only a 3.117% chance of a sample group of 44 averaging as high as $\bar{x} = 35.09$. There is evidence to reject H_0 and conclude that the district's average is higher than 32.

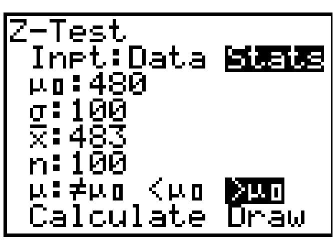
6.3 Use and Abuse of Tests

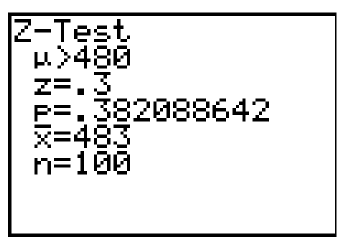
We continue with two more exercises that illustrate how one must be careful in drawing conclusions of significance.

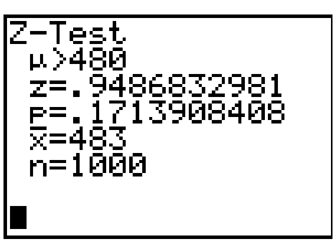
Example 6.7 SAT Coaching. Suppose that SATM scores vary normally with $\sigma = 100$. Calculate the p-value for the test of H_0: $\mu = 480$, H_A: $\mu > 480$ in each of the following situations:

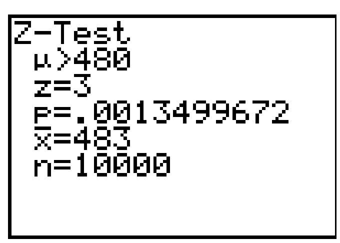
(a) A sample of 100 coached students yielded an average of $\bar{x} = 483$.
(b) A sample of 1000 coached students yielded an average of $\bar{x} = 483$.
(c) A sample of 10,000 coached students yielded an average of $\bar{x} = 483$.

Solution. We adjust the settings in the `Z-Test` screen from the `STAT TESTS` menu and calculate, changing n each time. Below are the results using the three different sample sizes. Notice that with increasing sample size, the p-value changes dramatically. We see that the rise in the average score to $\bar{x} = 483$ is significant (p very small) only when the results stem from the very large sample of 10,000 coached students. With the sample of only 100 students, there is 38.2% chance of obtaining a sample mean as high as $\bar{x} = 483$, even if the true mean were still 480.



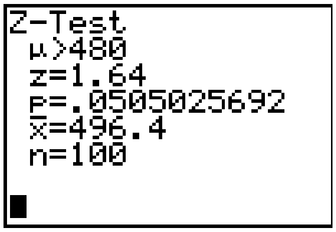


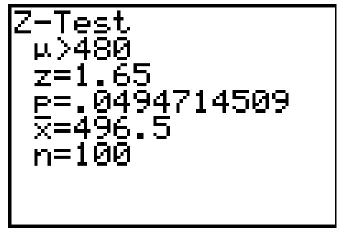




Example 6.8 More SAT Coaching. For the same hypothesis test as in Example 6.7 above, consider the sample mean of 100 coached students. (a) Is $\bar{x}$ = 496.4 significant at the 5% level? (b) Is $\bar{x}$ = 496.5 significant at the 5% level?

Solution: we perform the `Z-Test` for both values of $\bar{x}$:





In the first case, p = 0.0505 > 0.05; so the value of $\bar{x}$ = 496.4 is not significant at the 5% level. However, in the second case, p = 0.04947 < 0.05; so the value of $\bar{x}$ = 496.5 is significant at the 5% level. However, for SATM scores, there is no real "significant" difference between means of 496.4 and 496.5.

6.4 Power and Inference as a Decision

Power is the probability of correctly rejecting a null hypothesis. It is a function of the significance level of the test (with larger α we will correctly reject H_0 more often, but will also wrongly reject it as well), the sample size n, and the distance between H_0 and the true value. We conclude this chapter with some examples on computing the power against an alternative.

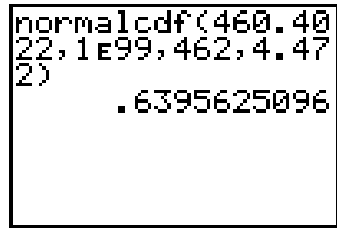
Example 6.9 More California SATs. In Example 6.5 we considered the hypotheses H_0: $\mu = 450$, H_A: $\mu > 450$. A sample of size n = 500 is taken from a normal population having $\sigma = 100$. If we performed our test at the 1% level of significance, find the power of this test against the alternative $\mu = 462$.

Solution. We first find the rejection region of the test at the 1% level of significance. Because the alternative is the one-sided right tail, we wish the right-tail probability under the standard normal curve to be 0.01. This probability occurs at z^* = 2.326. So we reject H_0 if the z test statistic is more than 2.326. That is, we reject if

$$\frac{\bar{x}-450}{100/\sqrt{500}} > 2.326$$

or equivalently if $\bar{x} > 450+2.326*100/\sqrt{500} = 460.4022$. Now we must find the probability that $\bar{x}$ is greater than 460.4022, given that the alternative $\mu = 462$ is true. Given that $\mu = 462$, then $\bar{x} \sim N(462, 100/\sqrt{500} = 4.472)$, and we must compute

$P(\overline{x} > 460.4022)$. To do so, we use the command `normalcdf(460.4022,1E99,462,4.472)` and find that the power against the alternative μ = 462 is about 0.64.

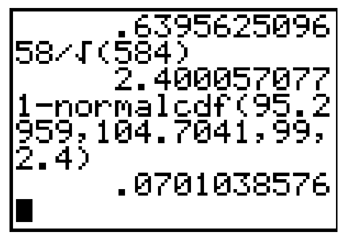


Example 6.10 Two-sided Power. (a) An SRS of size 584 is taken from a population having σ = 58 to test the hypothesis H_0: μ = 100 versus a two-sided alternative at the 5% level of significance. Find the power against the alternative μ = 99.

Solution. Again, we first must find the rejection regions. For a two-sided alternative at the 5% level of significance, we allow 2.5% at each tail. Thus, we reject if the *z* test statistic is beyond $\pm$ 1.96. That is, we reject if

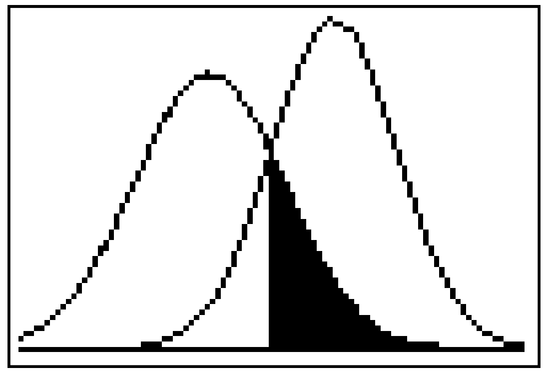
$$\frac{\overline{x}-100}{58/\sqrt{584}} < -1.96 \quad \text{or} \quad \frac{\overline{x}-100}{58/\sqrt{584}} > 1.96$$

Equivalently, we reject if $\overline{x}$ < 95.2959 or if $\overline{x}$ > 104.7041. Now assuming that μ = 99, then $\overline{x} \sim (99, 58/\sqrt{584} = 2.400)$. We now must compute $P(\overline{x} < 95.2959) + P(\overline{x} > 104.7041)$, which is equivalent to $1 - P(95.2959 < \overline{x} < 104.7041)$ = `1-normalcdf(95.2959,104.7041,99,2.400)`. With this command, we see that the power against the alternative μ = 99 is about 0.07. This low power is to be expected, since the difference between 100 and 99 is small.



CHAPTER

7



Inference for Distributions

Introduction

In this chapter, we demonstrate the various *t* procedures that are used for confidence intervals and significance tests about the mean of a normal population for which the standard deviation is unknown. We also consider comparing means from two independent samples and paired samples.

7.1 Inference for the Mean of a Population

We begin with a short program that allows us to find a critical value t^* after specifying the degrees of freedom and confidence level. This is a built-in function on TI-84 and TI-89 calculators. TI-83s need this program (or you could use Table C). This program asks for degrees of freedom rather than sample size, since t distributions are used in inference for more than just a single sample.

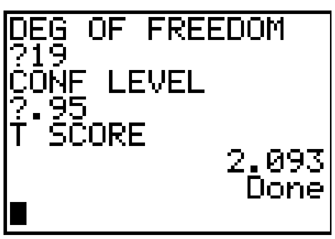
The TSCORE Program

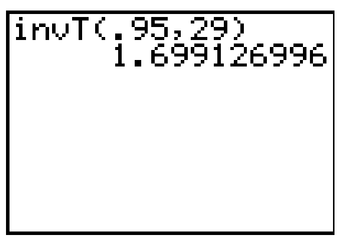
```
PROGRAM:TSCORE
:Disp "DEG. OF FREEDOM"
:Input M
:Disp "CONF. LEVEL"
:Input R
:tcdf(0,X,M)↲Y1
:solve(Y1-R/2,X,2) ↲Q
:Disp "T SCORE"
:Disp round(Q,3)
```

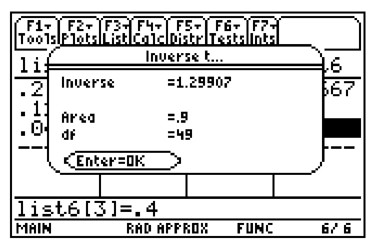
Example 7.1 Finding *t** Find the critical values t^* for confidence intervals for the mean in the following cases:

(a) A 95% confidence interval based on $n = 20$ observations
(b) A 90% confidence interval from an SRS of 30 observations
(c) An 80% confidence interval from a sample of size 50

Solution. The confidence intervals are based on t distributions with $n - 1$ degrees of freedom, since we have single samples. So we need 19 degrees of freedom for part (a), 29 degrees of freedom for part (b), and 49 degrees of freedom for part (c). Below are the outputs of the **TSCORE** program and from the built-in functions. These (like invNorm) use area to the left of the point desired, so for a 90% confidence interval there is .95 to the left of t^* and for 80% confidence there is 90% to the left of t^*.





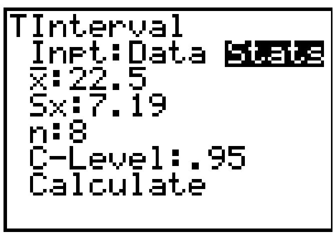


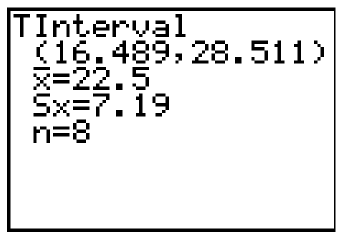
One-Sample *t* Confidence Interval

We now examine confidence intervals for one mean for which we will use the `TInterval` feature (item 8) from the `STAT TESTS` menu (□ `Ints` on a TI-89). As with the `ZInterval` feature that we used in Chapter 6, we can enter the summary statistics or use data in a list.

Example 7.2 Vitamin C in Corn. The amount of vitamin C in a factory's production of corn soy blend (CSB) is measured from eight samples giving $\bar{x}$ = 22.50 (mg/100 g) and s = 7.19. Find a 95% confidence interval for the mean vitamin C content of the CSB produced during this run.

Solution. Call up the `TInterval` feature from the `STAT TESTS` menu. Set `Inpt` to `Stats`, then enter the values of $\bar{x}$, `Sx`, `n`, `C-Level`, and press ⊆ on `Calculate`. We obtain the interval 16.489 to 28.511.

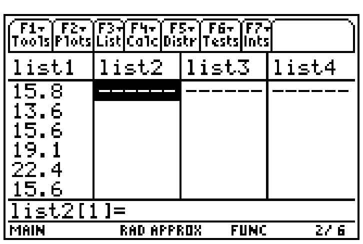


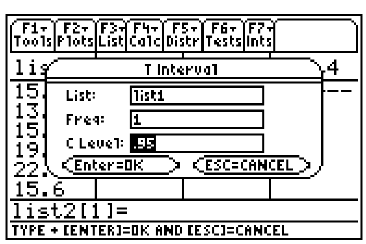


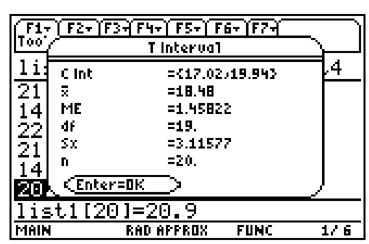
Example 7.3 Fuel Efficiency. Here are the values of the fuel efficiency in mpg for a sample of trials on a vehicle undergoing testing. Find the mean, the standard deviation, the standard error, the margin of error for a 95% confidence interval, and give a 95% confidence interval for the mean mpg of this vehicle.

15.8	13.6	15.6	19.1	22.4	15.6	22.5	17.2	19.4	22.6
19.4	18.0	14.6	18.7	21.0	14.8	22.6	21.5	14.3	20.9

Solution. First, enter the data into a list, say list `L1` (or `list1` on a TI-89). Next, bring up the `TInterval` screen, set `Inpt` to `Data`, set the `List` to `L1` with frequencies 1, enter the desired confidence level, and press ⊆ on `Calculate`. The sample mean, sample deviation, and confidence interval are all displayed.







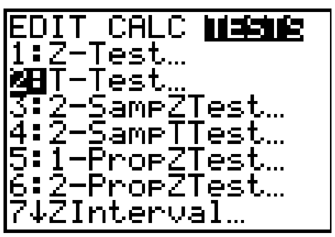
Because the confidence interval is of the form $\bar{x} \pm m$, we can find the margin of error m by subtracting $\bar{x}$ from the right endpoint of the interval: m = 19.938 – 18.48 = 1.458. The TI-89 gives the margin of error explicitly. The standard error is given by $s/\sqrt{n}$ = 3.115766358/$\sqrt{20}$ = 0.6967.

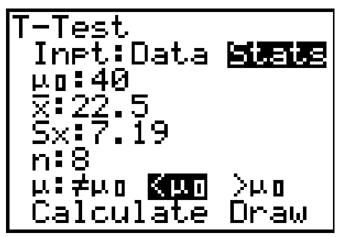
One-Sample *t* test

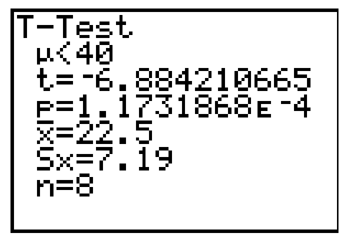
We now perform some significance tests about the mean using the `T-Test` feature (item 2) from the `STAT TESTS` menu.

Example 7.4 More Vitamin C in Corn. Using the vitamin C data of n = 8, $\bar{x}$ = 22.50, and s = 7.19 from Example 7.2 above, test the hypothesis H_0: μ = 40 versus the alternative H_A: μ < 40.

Solution. Bring up the `T-Test` screen from the `STAT TESTS` menu and adjust `Inpt` to `STATS`. Enter the value of μ_0 = 40 and the summary statistics, set the alternative to $< \mu_0$, then scroll down to `Calculate` and press ⊆.







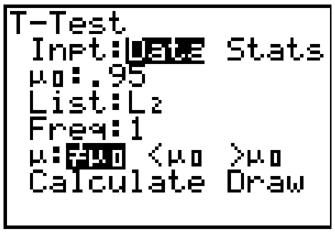
We obtain a *t* test statistic of –6.88 and a *p*-value of 0.0001. Because the *p*-value is so small, we have significant evidence to reject H_0. If the true mean were 40, a sample mean of 22.5 or lower with a sample of size 8 should only happen about 1 in 10,000 times.

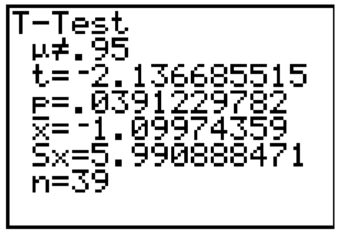
Example 7.5 Stock Returns. An investor sued his broker and brokerage firm because lack of diversification in his portfolio led to poor performance. The following table gives the monthly percentage rates of return for the months in which the account was managed by the broker. For these months, the average of the Standard & Poor's 500-Stock Index was 0.95%. Are these returns compatible with the S&P 500 average? Use the data to test the hypothesis H_0: μ = 0.95 versus the alternative H_A: $\mu \neq 0.95$.

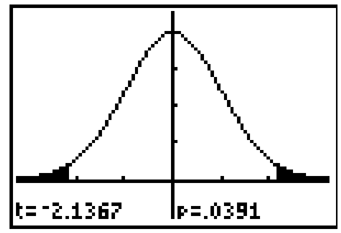
–8.36	1.63	–2.27	–2.93	–2.70	–2.93	–9.14	–2.64
6.82	–2.35	–3.58	6.13	7.00	–15.25	–8.66	–1.03
–9.16	–1.25	–1.22	–10.27	–5.11	–0.80	–1.44	1.28
–0.65	4.34	12.22	–7.21	–0.09	7.34	5.04	–7.24

–2.14	–1.01	–1.41	12.03	–2.56	4.33	2.35	

Solution. Enter the data into a list, say list `L2`, then bring up the `T-Test` screen and set `Inpt` to `Data`. Enter the value of μ_0 = .95, set the list to `L2` with frequency 1, and set the alternative to $\neq \mu_0$. Then press ⊆ on `Calculate` or `Draw`.







We obtain a *p*-value of 0.0391. If the average return were equal to 0.95%, then there would be only a 3.91% chance of a sample of 39 months averaging as far or farther away from that value as –1.1%. There is sufficient evidence to reject H_0 and conclude that the mean monthly return differs from (was worse than) 0.95%.

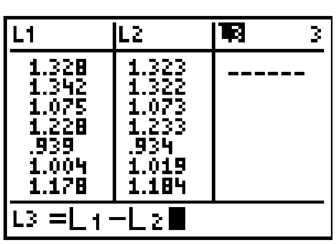
Matched Pair *t* Procedure

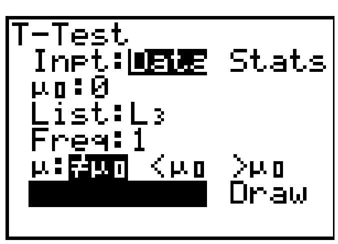
Example 7.6 Are the Technicians Consistent? Two operators of X-ray machinery measured the same eight subjects for total body bone mineral content. We want these two operators to have consistent results when dealing with this test. Here are the results in grams:

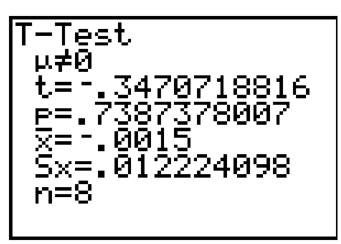
	Subject							
Operator	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	1.328	1.342	1.075	1.228	0.939	1.004	1.178	1.286
2	1.323	1.322	1.073	1.233	0.934	1.019	1.184	1.304

Use a significance test to examine the null hypothesis that the two operators have the same mean. Use a 95% confidence interval to provide a range of differences that are compatible with these data.

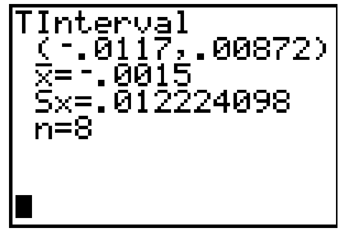
Solution. We consider the average μ_D of the *difference* of the measurements between the operators. First, enter the measurements of Operator 1 into list `L1` and the measurements of Operator 2 into list `L2`. Next, highlight list name `L3` and enter the command `L1-L2` and press ⊆ to store the differences. Then, use a `T-Test` on list `L3` to test H_0: $\mu_D = 0$ versus the alternative H_A: $\mu_D \neq 0$.







We obtain a *p*-value of 0.7387 from a test statistic of $t = -0.347$. Due to the high *p*-value, we can say that there is not a significant average difference. If μ_D were equal to 0, then there would be a 73.87% chance of having an average difference as far away as $\bar{d} = -0.0015$ with a random sample of eight subjects. Next, use the `TInterval` on list `L3` to find a 95% confidence interval for the average difference. Because the interval (–0.0117, 0.00872) contains 0, we have further evidence that the operators could have the same mean.



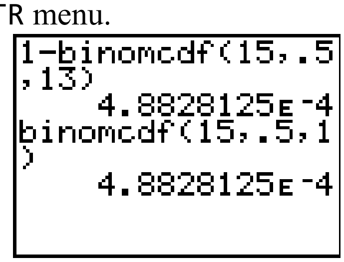
The Sign Test

The sign test is an easy way to perform a hypothesis test about the median of the distribution if the data are not normal. It is based on the binomial distribution. If the null hypothesis is true there should be a 50% chance of observing a response either more than or less than the claimed median.

Example 7.7 The Full Moon and Behavior. A study of dementia patients in nursing homes recorded various types of disruptive behavior every day for 12 weeks. Days were classified as moon days if they were in a three-day period centered at the day of the full moon. For each patient the average number of disruptive behaviors for moon days and for all other days was tracked. Out of 15 patients, 14 had more aggressive behavior on moon days than on other days. Use the sign test on the hypothesis of "no moon effect."

Solution. Because so many patients had a change in behavior, we shall test the hypothesis H_0: $p = 0.50$ with the alternative H_A: $p > 0.50$.

We must compute the probability of there being as many as 14 changes with a *B*(15,0.50) distribution. Equivalently, we can compute the probability of there being as few as one nonchange. Thus, we could compute either $P(B \geq 14) = 1 - P(B \leq 13)$ or $P(B \leq 1)$. To do so, we use the built-in `binomcdf(` command from the `DISTR` menu.



After entering either `1-binomcdf(15,.5,13)` or `binomcdf(15,.5,1)`, we obtain the very low *p*-value of 0.000488. If there were no moon effect, then there would be almost no chance of having as many as 14 out of 15 showing a change. Therefore, we can reject H_0 in favor of the alternative that the moon generally causes more aggressive behavior.

7.2 Comparing Two Means

We next consider confidence intervals and significance tests for the difference of means $\mu_1 - \mu_2$ given two normal populations that have unknown standard deviations. The results are based on independent random samples of sizes n_1 and n_2. For the most accurate results, we can use the `2-SampTInt` and `2-SampTTest` features from the `STAT TESTS` menu.

These features require that we specify whether or not we wish to use the pooled sample variance s_p^2. We should specify "Yes" only when we assume that the two populations have the *same* (unknown) variance. In this case, the critical values *t** are obtained from the $t(n_1 + n_2 - 2)$ distribution and the standard error is $s\sqrt{1/n_1 + 1/n_2}$. When we specify "No" for the pooled variance, then the standard error is $\sqrt{s_1^2/n_1 + s_2^2/n_2}$ and the degrees of freedom *r* are given by

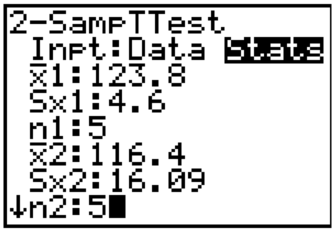
$$r = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

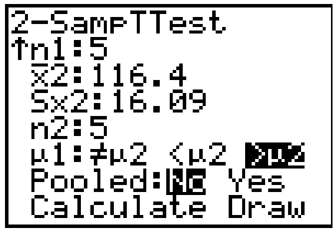
But if the true population standard deviations σ_1 and σ_2 are known, then we should use the `2-SampZInt` and `2-SampZTest` features for our calculations.

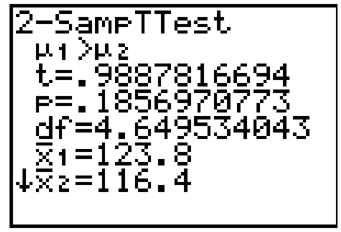
Example 7.8 Does Polyester Decay? How quickly do synthetic fabrics decay in landfills? A researcher buried polyester strips for different lengths of time, then dug up the strips and measured the force required to break them. Breaking strength is easy to measure and a good indicator of decay. One set of strips was dug up after 2 weeks and the other was dug up after 16 weeks. We suspect that decay increases over time. The results are given in the table below. Test the hypothesis H_0: $\mu_1 = \mu_2$ versus H_A: $\mu_1 > \mu_2$, then give a 95% confidence interval for $\mu_1 - \mu_2$.

Group	n	$\bar{x}$	s
2 weeks	5	123.80	4.60
16 weeks	5	116.40	16.09

Solution. Use the `2-SampTTest` feature from the `STAT TESTS` menu, and set `Inpt` to `Stats`. Enter the given statistics, set the alternative, and enter `No` for `Pooled`. Then press ⊆ on `Calculate`.

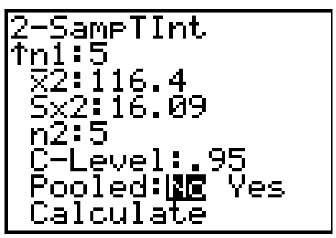


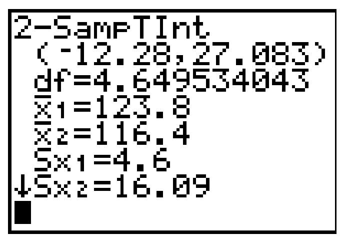




We obtain a *p*-value of 0.1857 from a test statistic of .9888 with 4.6495 degrees of freedom. If the true mean decay rates were equal, then there would be a good chance of our means being as different as they are (the difference is 7.4) with samples of these sizes. We therefore fail to reject H_0 and conclude that there is not a significant difference in decay of polyester with the extra 14 weeks.

To calculate a confidence interval for $\mu_1 - \mu_2$, we use `2-SampTInt.` Our data is still there, so one can simply move the arrow down to enter the desired confidence level and then calculate. We obtain the interval (–12.28, 27.08). These were small samples, which is one reason for the wide interval. We also note that 0 is included in the interval, which supports our conclusion of no significant increase in polyester decay with the additional 14 weeks of burial.





Example 7.9 College Study Habits. The Survey of Study Habits and Attitudes was given to first-year students at a private college. The tables below show a random sample of the scores.

Women's scores

154	109	137	115	152	140	154	178	101
103	126	126	137	165	165	129	200	148

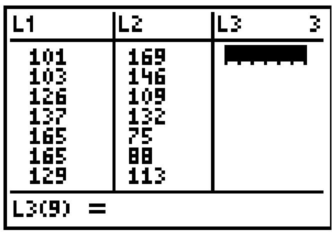
Men's scores

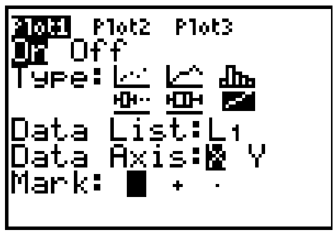
108	140	114	91	180	115	126	92	169	146
109	132	75	88	113	151	70	115	187	104

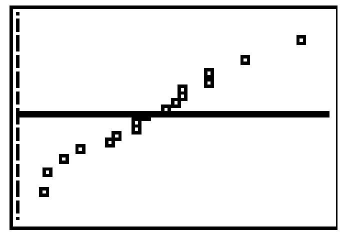
(a) Examine each sample graphically to determine if the use of a *t* procedure is acceptable.

(b) Test the supposition that the mean score for all men is lower than the mean score for all women among first-year students at this college.
(c) Give a 90% confidence interval for the mean difference between the SSHA scores of male and female first-year students at this college.

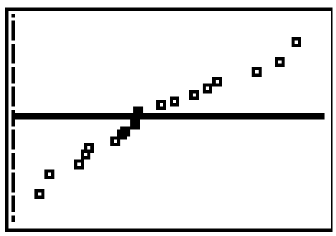
Solution. (a) We shall make normal quantile plots of these data. Enter the women's scores into list L1 and the men's scores into list L2. In the STAT PLOT screen, adjust the Type settings for both Plot1 to the last type for a normal quantile plot, then graph each list separately. Press θ→ to display the plots. The resulting plots appear close enough to linear to warrant use of *t* procedures.





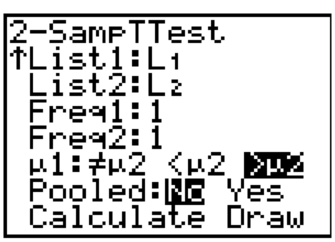


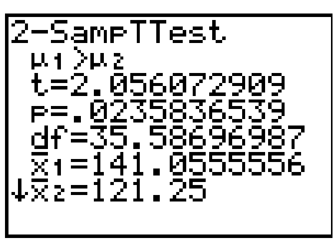
Women

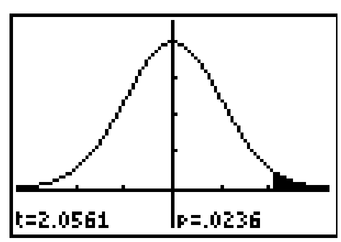


Men

(b) Next, let μ_1 be the mean SSHA score among all first-year women and let μ_2 be the mean score among all first-year men. We shall test the hypothesis H_0: $\mu_1 = \mu_2$ versus the alternative H_A: $\mu_1 > \mu_2$. On the 2-SampTTest screen, set Inpt to Data, enter the desired lists L1 and L2, set the alternative to $> \mu_2$, enter No for Pooled, and press ⊆ on Calculate or Draw.

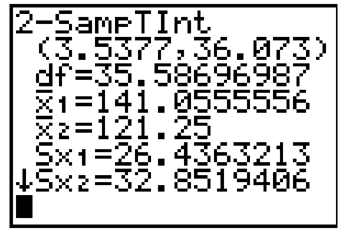






We obtain a *p*-value of 0.02358. If the true means were equal, then there would be only a 2.358% chance of $\bar{x}_1$ being so much larger than $\bar{x}_2$ with samples of these sizes. The relatively low *p*-value gives us evidence to reject H_0 and conclude that $\mu_1 > \mu_2$. That is, the mean score for all men is lower than the mean score for all women among first-year students at this college.

(c) Adjust the settings in the `2-SampTInt` screen and calculate. We obtain (3.5377, 36.073). That is, the mean score of female first-year students should be from about 3.5377 points higher to about 36.073 points higher than the mean score of male first-year students at this college.



Pooled Two-Sample *t* Procedures

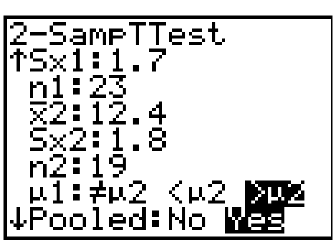
Example 7.10 Infants' Hemoglobin. A study of iron deficiency in infants compared samples of infants following different feeding regimens. Here are the summary results on hemoglobin levels at 12 months of age for two samples of infants:

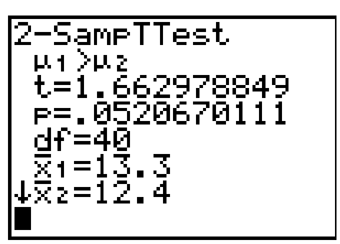
Group	***n***	$\bar{x}$	***s***
Breast-fed	23	13.3	1.7
Formula	19	12.4	1.8

(a) Is there significant evidence that the mean hemoglobin level is higher among breast-fed babies? State H_0 and H_A, and carry out a *t* test.
(b) Give a 95% confidence interval for the difference in mean hemoglobin levels between the two populations of infants.

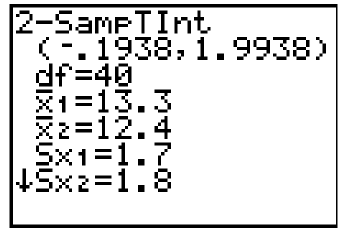
Solution. Let μ_1 be the mean hemoglobin level for all breast-fed babies and let μ_2 be the mean level for all formula-fed babies. Because the sample deviations are so close, it appears that true standard deviations among the two groups could be equal; thus, we may use the pooled two-sample *t* procedures. For part (a), we will test H_0: $\mu_1 = \mu_2$ versus H_A: $\mu_1 > \mu_2$.

Call up the `2-SampTTest` feature, and set `Inpt` to `Stats`. Enter the given statistics, set the alternative, enter `Yes` for `Pooled`, and calculate.





With a *p*-value of 0.052, we conclude that there is not statistical evidence (at the 5% level) to reject H_0. If the true means were equal, then there is greater than a 5% chance of $\bar{x}_1$ being 0.9 higher than $\bar{x}_2$ with samples of these sizes.



(b) Next, calculate a 95% confidence interval with the `2-SampTInt` feature set to `Yes` on `Pooled`. We see that $-0.1938 < \mu_1 - \mu_2 < 1.9938$. That is, the mean hemoglobin level of breast-fed babies could be from 0.1938 lower to 1.9938 higher than the mean level of formula-fed babies. Since 0 is in the interval, we are further convinced that there is not a significant difference between breast-fed and formula-fed babies' hemoglobin levels, on average.

7.3 Optional Topics in Comparing Distributions

We now demonstrate a test for determining whether or not two normal populations have the same variance. If so, then we would be justified in using the pooled two-sample *t* procedures for confidence intervals and significance tests about the difference in means. For the test, we will need the `2-SampFTest` feature from the `STAT TESTS` menu.

One word of caution about this test is in order, however. This is *extremely* sensitive to any departures from normality, and the lack of robustness does not improve substantially with larger sample sizes. As such, it should be used with wariness; it is difficult to tell whether a significant *p*-value is due to a difference in the standard deviations, or to a lack of normality.

The *F* Ratio Test

Example 7.11 More on SSHA. Consider again the data from Example 7.9 regarding the SSHA scores of first-year students at a private college. Test whether the women's scores are less variable than the men's.

Women's scores

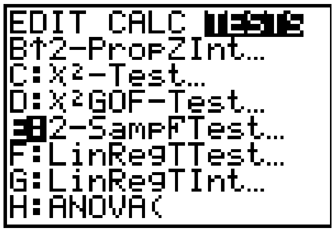
154	109	137	115	152	140	154	178	101
103	126	126	137	165	165	129	200	148

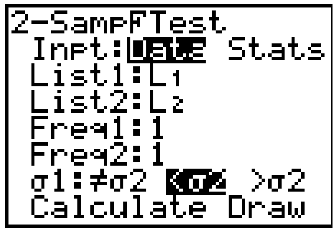
Men's scores

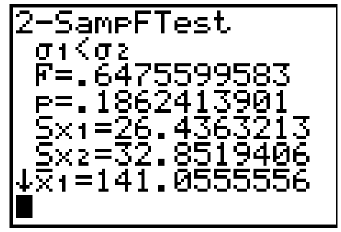
108	140	114	91	180	115	126	92	169	146
109	132	75	88	113	151	70	115	187	104

Solution. Let σ_1 be the standard deviation of all women's scores and let σ_2 be the standard deviation for all men's scores. We shall test H_0: $\sigma_1 = \sigma_2$ versus H_A: $\sigma_1 < \sigma_2$. To do so, first enter the data sets into lists, say `L1` and `L2`. Next, bring up the

`2-SampFTest` screen from the `STAT TESTS` menu, set `Inpt` to `Data`, enter the appropriate lists names and alternative, and calculate.





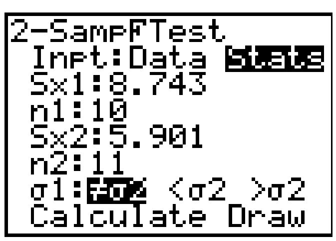


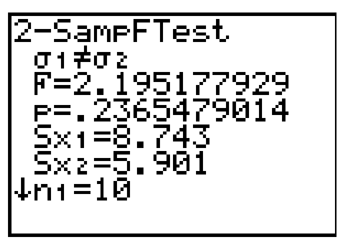
With a *p*-value of 0.18624, we do not have strong evidence to reject H_0. If σ_1 were equal to σ_2, then there would be an 18.6% chance of the women's sample deviation of `Sx1` = 26.4363 being so much lower than the men's sample deviation of `Sx2` = 32.8519. Thus, we cannot assert strongly that the women's scores are less variable.

Example 7.12 Calcium and Blood Pressure. Here are the summary statistics for the drop in blood pressure from two sample groups of patients undergoing treatment. Test to see if the groups in general have the same standard deviation.

Group	***n***	$\bar{x}$	***s***
Calcium	10	5.000	8.743
Placebo	11	–0.273	5.901

Solution. Let σ_1 be the standard deviation of all possible patients in the calcium group, and let σ_2 be the standard deviation of all possible patients in the placebo group. We will test the hypothesis H_0: $\sigma_1 = \sigma_2$ versus H_A: $\sigma_1 \neq \sigma_2$. Bring up the `2-SampFTest` screen and set `Inpt` to `Stats`. Enter the summary statistics and alternative, then calculate.

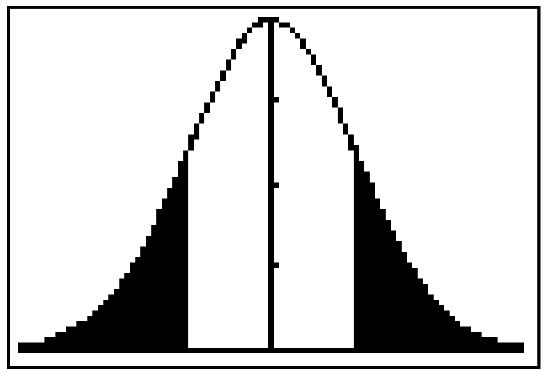




We obtain an *F*-statistic of 2.195 and a *p*-value of 0.2365. If σ_1 were equal to σ_2, then there would be about a 23.6% chance of `Sx1` and `Sx2` being so far apart with samples of these sizes. This evidence may not be significant enough to reject H_0 in favor of the alternative.

CHAPTER

8



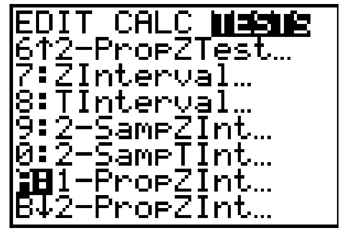
Inference for Proportions

Introduction

In this chapter, we discuss how to find confidence intervals and to conduct hypothesis tests for a single proportion and for the difference in two population proportions. We will also show how to adjust the counts for confidence intervals in which there is a small number of successes or failures.

8.1 Inference for a Single Proportion

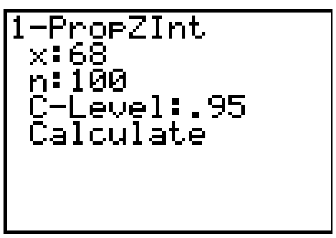
Both the large-sample and plus-four level *C* confidence intervals can be calculated using the `1-PropZInt` feature from the `STAT TESTS` menu. Significance tests can be worked using the `1-PropZTest`.

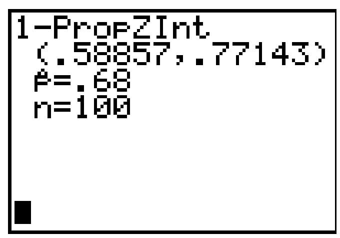


A Large-Sample Confidence Interval

Example 8.1 Stress in Restaurant Workers. In a restaurant worker survey, 68 of a sample of 100 employees agreed that work stress had a negative impact on their personal lives. Find a 95% confidence interval for the true proportion of restaurant employees who agree.

Solution. Bring up the `1-PropZInt` screen, enter 68 for `x`, enter 100 for `n`, and enter .95 for `C-Level`. Then press ⊆ on `Calculate` to obtain a 95% confidence interval of (0.58857, 0.77143). Based on this sample, with 95% confidence, between 58.9% and 77.1% of restaurant workers will agree that work stress has a negative impact on their personal lives.





A Plus-Four Confidence Interval

Example 8.2 Blind Medical Trials. Many medical trials randomly assign patients to either an active treatment or a placebo. The trials are supposed to be double-blind, but sometimes patients can tell whether or not they are getting the active treatment. Reports of medical research usually ignore the problem. Investigators looked at a random sample of 97 articles and found that only 7 discussed their success in blinding the trial. What proportion of all such studies discuss the success of blinding in their trials? Give a 95% confidence interval estimate.

Because we have a small number of studies that discussed the success of blinding in the trial (less than 10), we use the "plus-four" method: we simply add 4 to the number of studies (it becomes 101) and 2 to the number of "successes" (we'll consider that 9 discussed the success in blinding the trial). We use `1-PropZInt` as before, with the adjusted counts.

Solution. In the `1-PropZInt` screen, enter 9 for `x`, which is 2 more than the actual number and enter 101 for `n`, which is 4 more than the actual sample size. Enter the desired `C-level` (95%) and press ⊆on `Calculate` to obtain the plus-four estimate $\tilde{p}$ = 0.089 and the confidence interval.

```
1-PropZInt
 x:9
 n:101
 C-Level:.95
 Calculate
```

```
1-PropZInt
 (.03355,.14467)
 p̂=.0891089109
 n=101
```

Based on this sample, we are 95% confident that between 3.4% and 14.5% of published medical studies will discuss the success of blinding in their trials.

Choosing a Sample Size

As with confidence intervals for the mean, we often would like to know in advance what sample size would provide a certain maximum margin of error m with a certain level of confidence. The required sample size n satisfies $n \geq \left(\frac{z^*}{m}\right)^2 p^*(1-p^*)$, where z^* is the appropriate critical value depending on the level of confidence and p^* is a guessed value of the true proportion p. If $p^* = 0.50$, then the resulting sample size insures that the margin of error is no more than m, regardless of the true value of p.

While the formula is fairly easy to use, we supply a simple program **PSAMPSZE** that displays the required sample size, rounded up to the nearest integer, after one enters the desired error m, the confidence level, and the guess p^*.

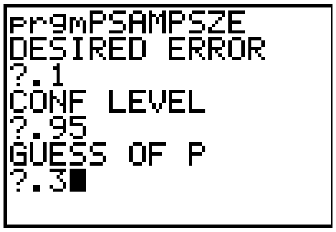
The PSAMPSZE Program

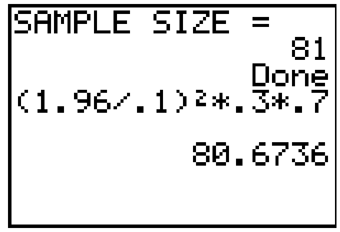
PROGRAM:PSAMPSZE	:If int(M)=M
:Disp "DESIRED ERROR"	:Then
:Input E	:M↵N
:Disp "CONF. LEVEL"	:Else
:Input R	:int(M+1) ↵N
:Disp "GUESS OF P"	:End
:Input P	:ClrHome
:invNorm((R+1)/2,0,1)↵Q	:Disp "SAMPLE SIZE="
:(Q/E)ϒ*P(1-P) ↵M	:Disp int(N)

Example 8.3 Alcohol Awareness — Find a Sample Size. Among students who completed an alcohol awareness program, you want to estimate the proportion who state that their behavior towards alcohol has changed since the program. Using the guessed

value of p^* = 0.30 from previous surveys, find the sample size required to obtain a 95% confidence interval with a maximum margin of error of m = 0.10.

Solution. Executing the **PSAMPSZE** program, we find that a sample of size 81 would be required. This value also can be obtained by $n = (1.96/.1)^2 * .3 * .7 = 80.6736$ which we round up (as always for sample sizes) to 81.



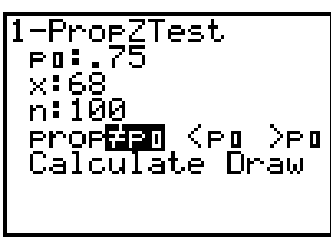


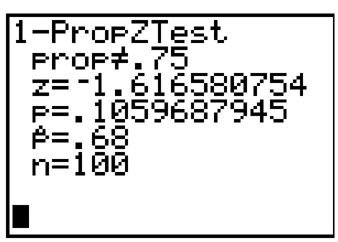
Significance Tests

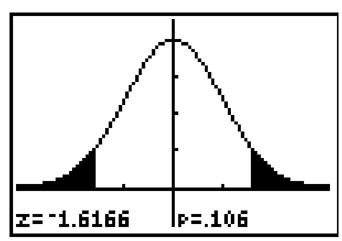
We now show how to conduct hypothesis tests for a single population proportion p using the `1-PropZTest` feature (item 5) in the `STAT TESTS` menu.

Example 8.4 More Stress in Restaurant Workers. In the restaurant worker survey, 68 of a sample of 100 employees agreed that work stress had a negative impact on their personal lives. Let p be the true proportion of restaurant employees who agree. Test the hypothesis H_0: $p = 0.75$ versus H_A: $p \neq 0.75$.

Solution. Enter the data and alternative into the `1-PropZTest` screen, then press ÷ on `Calculate` or `Draw`. We obtain a (two-sided) p-value of 0.106 from a z-statistic of –1.61658. If p were equal to 0.75, then there would be a 10.6% chance of obtaining $\hat{p}$ as far away as 0.68 with a sample of size 100. We therefore believe the proportion of restaurant workers who would agree that work stress has a negative impact on their personal lives is not significantly different from 75%.

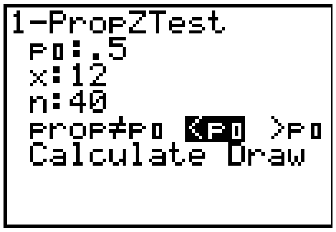


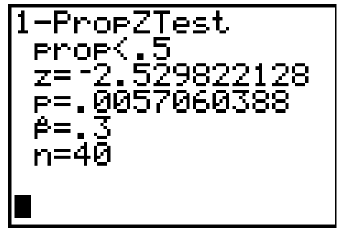




Example 8.5 Who Likes Instant? In a taste test of instant versus fresh-brewed coffee, only 12 out of 40 subjects preferred the instant coffee. Let p be the true probability that a random person prefers the instant coffee. Test the claim H_0: $p = 0.50$ versus H_A: $p < 0.50$ at the 5% level of significance.

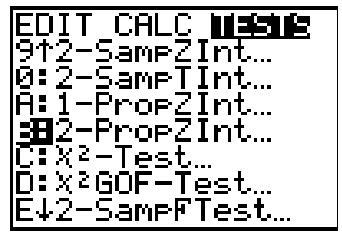
Solution. Enter the data and alternative into the `1-PropZTest` screen and calculate. We obtain a test statistic of –2.53 and a p-value of 0.0057. If p were 0.50, then there would be only a 0.0057 probability of $\hat{p}$ being as low as 0.3 with 40 subjects. There is strong evidence to reject H_0, and conclude that those who prefer instant coffee are a minority of the population.





8.2 Comparing Two Proportions

We now demonstrate confidence intervals and significance tests for the difference of two population proportions (p_1 and p_2). These calculations are done using the `2-PropZInt` and `2-PropZTest` (item 5) features from the `STAT TESTS` menu.

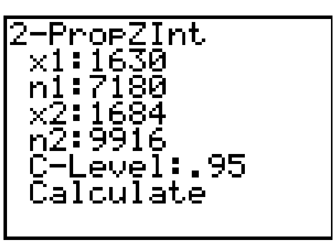


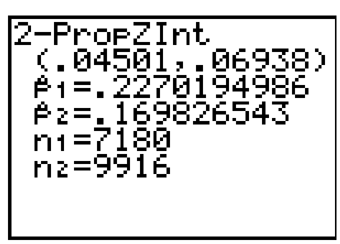
A Large-Sample Confidence Interval for Difference of Proportions

Example 8.6 Binge Drinking on Campus. The table below gives the sample sizes and numbers of men and women who responded "Yes" to being frequent binge drinkers in a survey of college students. Find a 95% confidence interval for the difference between the proportions of men and women who are frequent binge drinkers.

Population	n	X
Men	7180	1630
Women	9916	1684

Solution. Call the men group 1 and the women group 2. In the `2-PropZInt` screen, enter **1630** for `x1`, **7180** for `n1`, **1684** for `x2`, and **9916** for `n2`. Set the confidence level to **.95** and press ÷ on `Calculate`.





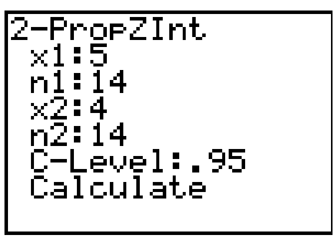
We obtain a confidence interval of (0.045, 0.069). That is, the proportion of male binge drinkers is from 4.5 percentage points higher to 6.9 percentage points higher than the proportion of female binge drinkers, with 95% confidence.

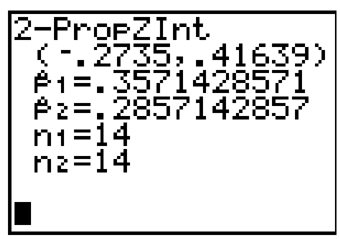
A Plus-Four Confidence Interval for Difference of Proportions

Example 8.7 Gender Differences. In studies that look for a difference between genders, a major concern is whether or not apparent differences are due to other variables that are associated with gender. A study of 12 boys and 12 girls found that 4 of the boys and 3 of the girls had a Tanner score (a measure of sexual maturity) of 4 or 5 (a high level of sexual maturity). Find a plus-four 95% confidence interval for the difference in proportions of all boys and all girls who would score a 4 or 5.

Solution. We can still use the `2-PropZInt` feature to find a plus-four confidence interval for $p_1 - p_2$. But for `x1` and `x2`, enter 1 more than the actual number of positive responses (we split the two successes added in the one sample case). For `n1` and `n2`, enter 2 more than the actual sample sizes (splitting the added four trials). Here, enter **5** for `x1`, **4** for `x2`, and enter **14** for both `n1` and `n2`. Then, set the confidence level to **.95** and calculate.

We see that $-0.2735 < p_1 - p_2 < 0.4164$. That is, the true proportion of boys who score 4 or 5 is from 27.35 percentage points lower to 41.64 percentage points higher than the true proportion of girls who score 4 or 5. Since 0 is included in this interval, there may be no difference in the population proportions of Tanner scores.





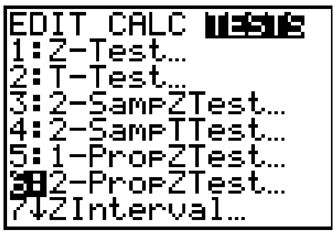
Significance Tests for Difference of Proportions

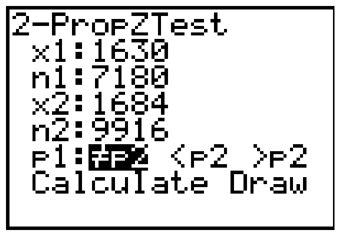
We now show how to conduct hypothesis tests about $p_1 - p_2$ using the `2-PropZTest` feature from the `STAT TESTS` menu.

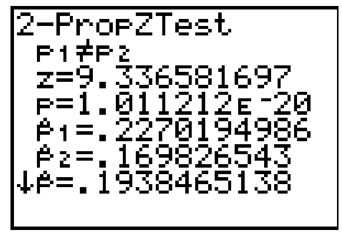
Example 8.11 More Binge Drinking. The table below again gives the sample sizes and numbers of men and women who responded “Yes” to being frequent binge drinkers in a survey of college students. Does the data give good evidence that the true proportions of male binge drinkers and female binge drinkers are different?

Population	n	X
Men	7180	1630
Women	9916	1684

Solution. Let p_1 be the true proportion of male students who are frequent binge drinkers, and let p_2 be the true proportion of female students who are frequent binge drinkers. We shall test the hypothesis H_0: $p_1 = p_2$ versus the alternative H_A: $p_1 \neq p_2$. Bring up the `2-PropZTest` screen, enter the actual data and the two-sided alternative, then calculate.





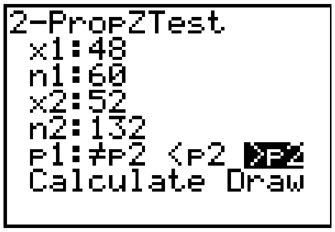


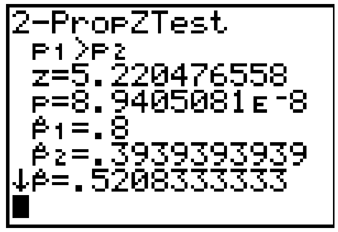
We obtain a *p*-value of 1×10^{-20} (essentially 0) from a test statistic of $z = 9.33658$. If $p_1 = p_2$ were true, then there would be no chance of obtaining sample proportions as far apart as $\hat{p}_1 = 0.227$ and $\hat{p}_2 = 0.1698$ with samples of these sizes. So, we can reject H_0 and conclude that not only are the genders different in (admitting to) binge drinking, but that males are also more likely to engage in binge drinking (or at least to admit to it). This reinforces our results obtained in Example 8.6; the confidence interval was wholly positive, so there would be no belief that the two proportions should be the same.

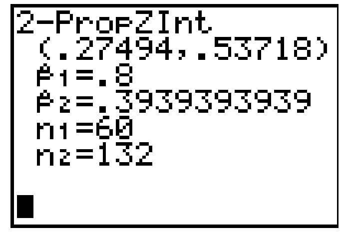
Example 8.12 Juvenile References. The table below gives the results of a gender bias analysis of a textbook. Do the data give evidence that the proportion of juvenile female references is higher than the proportion of juvenile male references? Compare the results with those of a 90% confidence interval.

Gender	n	X (juvenile)
Female	60	48
Male	132	52

Solution. Let p_1 be the true proportion of juvenile female references in all such texts, and let p_2 be the true proportion of juvenile male references. We will test the hypothesis H_0: $p_1 = p_2$ versus H_A: $p_1 > p_2$. Bring up the `2-PropZTest` screen, enter the data, set the alternative to `>p2`, and calculate. Then bring up the `2-PropZInt` screen, and calculate a 90% confidence interval.





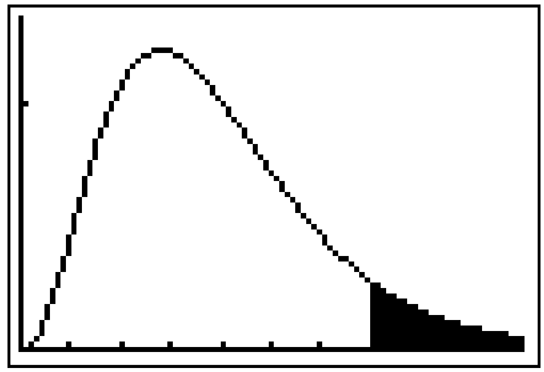


We obtain a very low p-value of 8.94×10^{-8}, which gives strong evidence to reject H_0 and conclude that there are more female juvenile references than male juvenile references, based on this sample. If $p_1 = p_2$ were true, then there would be almost no chance of obtaining a $\hat{p}_1$ that is so much higher than $\hat{p}_2$ with samples of these sizes.

The 90% confidence interval states that $0.275 < p_1 - p_2 < 0.537$. That is, among all such texts, the proportion of juvenile female references is from 27.5 percentage points higher to 53.7 percentage points higher than the proportion of juvenile male references.

CHAPTER

9



Inference for Two-Way Tables

Introduction

In this chapter, we describe how to perform a chi-square test on data from a two-way table. We shall be testing whether there is any association between the row variable traits and the column variable traits, or whether these row and column traits are independent. We will also perform a test to decide whether data agree with a proposed distribution.

9.1 Data Analysis for Two-Way Tables

Example 9.1 Business Nonresponse. Questionnaires were mailed to 300 randomly selected businesses in each of three categorical sizes. The following data show the number of responses.

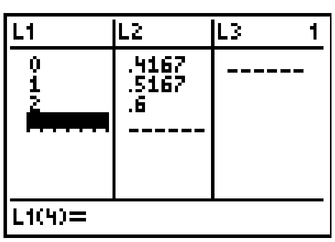
Size of company	Response	No response	Total
Small	175	125	300
Medium	145	155	300
Large	120	180	300

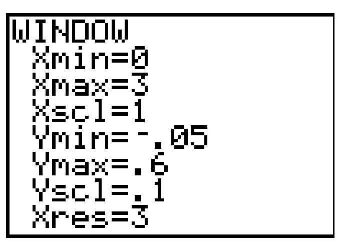
(a) What was the overall percent of nonresponse?
(b) For each size of company, compute the nonresponse percentage.
(c) Draw a bar graph of the nonresponse percents.
(d) Using the total number of responses as a base, compute the percentage of responses that came from each size of company.

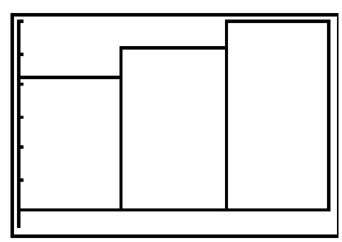
Solution. (a) To answer this question, we divide the total number of responses (125 + 155 + 180 = 460) by the total number of sent questionnaires (900). 460/900 = .5111. Overall, 51.1% of the businesses did not respond.

(b) The nonresponse rate for small companies is 125/300 = .4167 = 41.67%. For medium-sized companies, 155/300 = .5167 = 51.67% nonresponse. For large companies, 180/300 = .6 = 60% did not respond. It appears that as companies get larger, they are more likely not to respond to questionnaires.

(c) To make a (rough) bar graph of the nonresponse percents, we first enter the values **0**, **1**, and **2** into list `L1` in order to represent the three types of companies, and then enter the nonresponse proportions calculated in part (b) into list `L2`. Next, we adjust the `WINDOW` (to ensure that the bars have width 1) and `STAT PLOT` settings for a histogram of `L1` with frequencies `L2`, then σ. Remember that proper bar graphs should not have bars connected if you try to copy this down!







(d) There were 175 + 145 + 120 = 440 total responses. Of these, 175/440 = 39.77% came from small businesses, 145/440 = 32.95% from medium-sized businesses, and 120/440 = 27.27% came from large businesses. We again see that as business size increases, the businesses are more likely not to respond, since the response rates decline.

9.2 Inference for Two-Way Tables

We continue here with an example that shows how to compute the χ^2 test of no association between the row variable and column variable in a two-way table, and obtain the matrix of expected cell counts.

TI Calculators have a built-in χ^2–`Test` feature (on the `STAT TESTS` menu) that will compute the expected counts of a random sample under the assumption that the conditional distributions are the same for each category type. To use this feature, we first must enter data from a two-way table into a matrix in the `MATRX EDIT` screen.

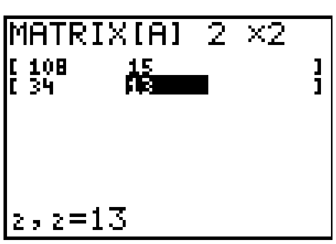
Example 9.2 Franchising Success. The following table shows the two-way relationship between whether a franchise succeeds and whether it has exclusive territory rights for a number of businesses.

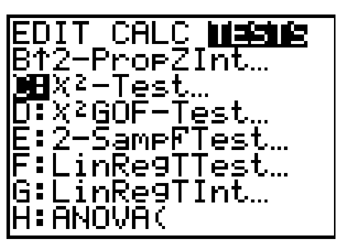
	Observed number of firms		
	Exclusive territory		
Success	Yes	No	Total
Yes	108	15	123
No	34	13	47
Total	142	28	170

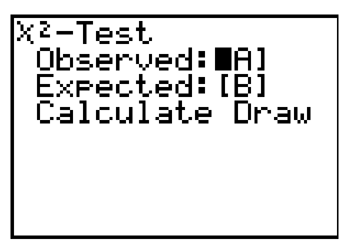
Under the assumption that there is no relationship between success and exclusive territory rights, find the expected number of successful franchises for each type of firm. And determine whether or not there is an association between franchise success and whether or not they have an exclusive territory.

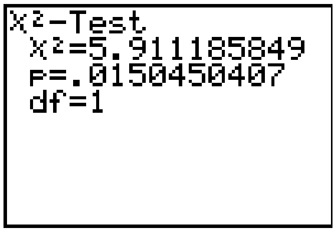
TI-83/84 Solution. First, enter the 2×2 table of data (excluding totals) into matrix `[A]` in the `MATRIX EDIT` screen. To get to this screen, press ψ□ for the `MATRIX` menu, then arrow to `EDIT`. Press ⊆ to select matrix `[A]`. Enter the size of the body of the matrix (rows × columns) pressing ÷ to advance the cursor. Enter the observed counts across the rows, pressing ÷ each time to advance the cursor. This menu is fussy so you must press ψ3 = `Quit` to exit the editor.

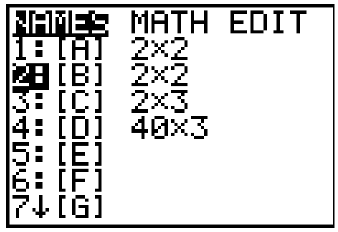
Next, bring up the χ^2–`Test` screen from the `STAT TESTS` menu, and adjust the `Observed` and `Expected` names if needed. The calculator defaults for this test are matrix `[A]` for `Observed`, and matrix `[B]` for `Expected`. If you have used another matrix, press ψ□ for the `MATRIX NAMES` screen and locate the matrix name you used; press ÷ to select it. Next, press ÷ on `Calculate`.

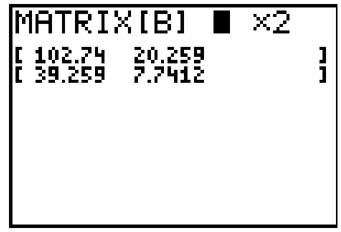




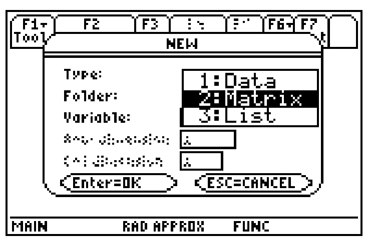


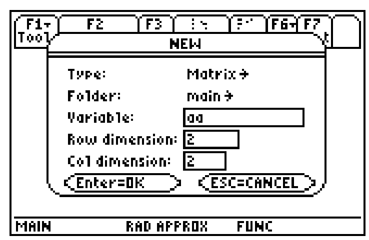


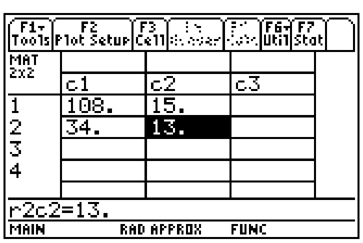


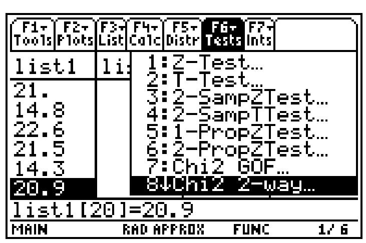


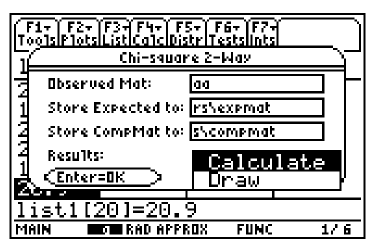
TI-89 Solution. On TI-89 calculators, the Matrix Editor is an application. Press O. Either press { or move the arrow down and press the right arrow to expand the options for what to edit. If you want to reuse a matrix name with a matrix of the same size, select option `2:Open` or select `3:New` here. You next expand the `Type:` box to select `2:Matrix`. In a similar way, select a folder (usually `Main`) and give the matrix a name (here `aa`) and its dimensions (ours is a 2×2). The matrix editor screen appears. Enter the observed counts across the rows, pressing ÷ each time to advance the cursor. This menu is fussy so you must press ψN = `Quit`5to exit the editor. Return to the `Stats/List Editor` application and press 2□ = □ for the `Tests` menu. The test we want is option `8:Chi2 2-way`. Select it and use 2| = ° to locate and enter your matrix name in the `Observed Mat` box. One can normally use the defaults for the `Expected` and `CompMat` matrices. As with the TI-83/84 series, there is a `Calculate` or `Draw` option for the output. Output is similar to the TI-83/84, but the first portion of the expected counts and components of the χ^2 statistic are shown. To observe these matrices fully, return to the Matrix Editor application and use the `Open` option to view them. Below, the contents of `Expmat` (the expected counts) and `Compmat` (contributions from each cell to the χ^2 statistic) are shown.

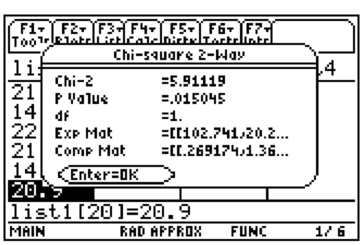


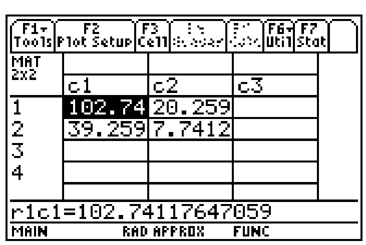


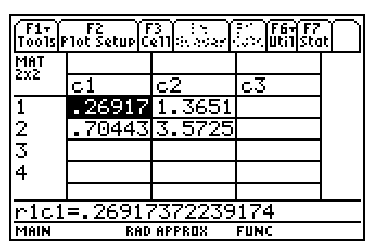










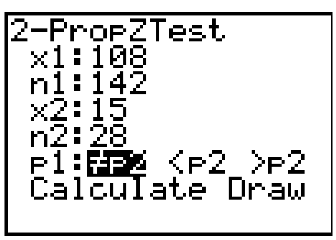


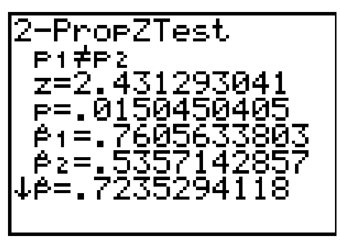
Conclusions. The χ^2–`Test` feature on all these calculators displays the chi-square test statistic, the *p*-value, and the degrees of freedom for the chi-square test of no association between the row and column variables. In this case, the low *p*-value of 0.015 gives strong evidence to reject the claim that there is no relationship between exclusive territory rights and franchise success. The TI-89 `Compmat` shows us that the largest components to the overall χ^2 statistic are from the franchises without exclusive territory. Those are most unlike what is expected under an assumption of no relationship.

If there were no relationship between success and exclusive territory rights, then we would expect 102.74 successful exclusive-territory franchises and 20.259 successful non-exclusive-territory franchises, as shown in the first row of matrix `[B]`. These values differ slightly from the observed values of 108 and 15 in the original data. The actual number of successful franchises in exclusive territory was somewhat higher than expected, and the actual number of successful franchises in nonexclusive territory somewhat lower than expected. It appears that having exclusive territory does help a franchise to be successful.

Comparison with the `2-PropZTest`

The $\chi 2$ test for a 2×2 table is equivalent to the two-sided *z* test for H_0: $p_1 = p_2$ versus H_A: $p_1 \neq p_2$. In Example 9.2, we could let p_1 be the true proportion of all successful exclusive-territory franchises and let p_2 be the true proportion of all successful non-exclusive-territory franchises. Then $\hat{p}_1$ = 108/142 and $\hat{p}_2$ = 15/28. If we enter these values into the `2-PropZTest` screen and use the alternative `≠p2`, then we obtain the same *p*-value of 0.015.





9.3 Formulas and Models for Two-Way Tables

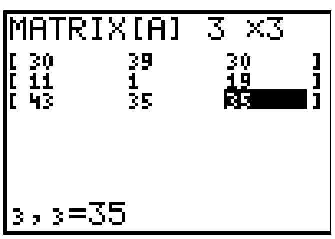
In this section, we further explain the test of no association among the row traits and the column traits versus the alternative that there is some relation between these traits.

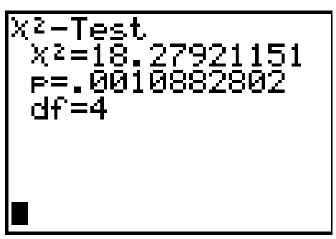
Example 9.3 Wine and Music. The two-way table that follows shows the types of wine purchased in a Northern Ireland supermarket while a certain type of music was being played. Test whether there is an association between the type of wine purchased and the type of music being played.

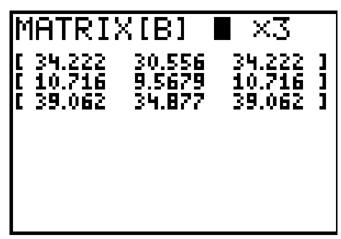
		Music		
Wine	None	French	Italian	Total
French	30	39	30	99
Italian	11	1	19	31
Other	43	35	35	113
Total	84	75	84	243

Solution. Looking at the data, it appears that music may affect the type of wine purchased. Notice that only one individual bought Italian wine while French music was playing. We first enter the data (excluding totals) into a 3×3 matrix `[A]` in the `MATRIX EDIT` screen.

To see if there really is an association, we shall test the null hypothesis H_0 that there is no relation between the music and the type of wine purchased. The alternative is that there is a relation, or that the type of wine purchased does depend on the music being played. After calculating, we obtain a *p*-value of 0.001 from a test statistic of 18.2792. If there were no association between the type of wine purchased and the type of music being played, then there would be only a 0.001 probability of obtaining observed cell counts that differ so much from the expected cell counts shown in matrix `[B]`. Due to this low *p*-value, we can reject the null hypothesis and say that the type of music *does* affect the type of wine purchased.







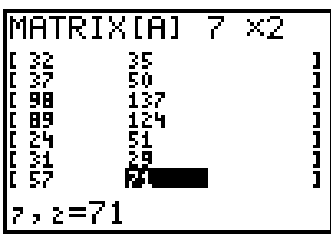
Looking further at the differences between the observed and expected counts, it appears that people are more willing to buy French wine when French music is playing (the observed count was 39 and the expected count 30.556), less likely to buy Italian wine when French music is playing (our observed count was 1 and we expect 9.568 if there were no relationship), and more willing to buy Italian wine when Italian music is playing (we observed 19 and expect 10.716).

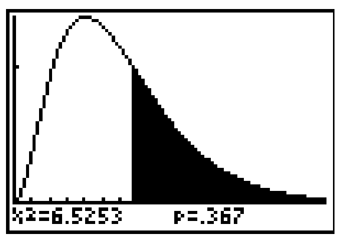
Example 9.4 Student Loans. A recent study of 865 college students found that 42.5% had student loans. The following table classifies the students by field of study and whether or not they have a loan.

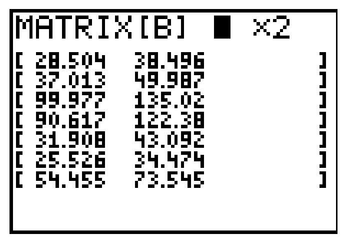
	Student loan	
Field of study	**Yes**	**No**
Agriculture	32	35
Child development and family studies	37	50
Engineering	98	137
Liberal arts and education	89	124
Management	24	51
Science	31	29
Technology	57	71

Carry out an analysis to see if there is a relationship between field of study and having a student loan.

Solution. We will test to see if the proportion of students having a student loan is the same regardless of field of study (i.e., if having a loan is independent of field). First, we enter the data into a 7×2 matrix `[A]` in the `MATRX EDIT` screen. Next, we bring up the χ^2`-Test` screen and designate `[A]` for `Observed`. This time selecting the `Draw` option, we obtain a *p*-value of 0.367 from a test statistic of 6.5253.







If having a loan were independent of field of study (i.e., if all fields had the same proportion of students with loans), then there would be a 0.367 probability of obtaining observed cell counts that differ as much from the expected cell counts. Notice that most of the expected counts in matrix `[B]` are close to the observed counts. Because of the high *p*-value, we can say that the observed differences are due to random chance and are not statistically significant. Thus, we will not reject the hypothesis that having a loan is independent of field of study. To further examine the data, we could compute the proportion of students in each major who have loans. For Agriculture majors, this is 32/67 = 47.8%; for Child development the proportion is 37/87 = 42.5%; for Engineering 98/235 = 41.7%; for Liberal arts 89/213 = 41.8%; Management 24/75 = 32%; Science 31/60 = 51.7%; and the proportion of Technology majors with a loan is 57/128 = 44.5%. All these proportions, with perhaps the exception of Management and Science, are very similar.

9.4 Goodness of Fit

TI-84 and -89 calculators have a built-in function to perform the goodness of fit for a discrete distribution. The commands for a TI-83 are simple enough, but we provide a small program below that also will accomplish the task.

Before executing the **FITTEST** program that follows, enter the specified proportions into list `L1` and enter the observed cell counts into list `L2`. The expected cell counts are computed and stored in list `L3`, and the individual contributions to the chi-square test statistic are stored in list `L4`. The program displays the test statistic and the *p*-value.

The FITTEST Program

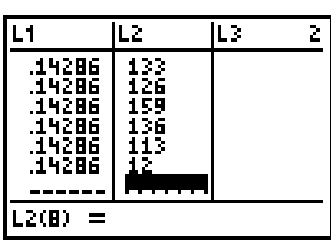
Program:FITTEST :sum(L2)↵N :N*L1↵L3 :(L2-L3)^2/L3↵L4 :sum(L4)↵X	:1-χ2cdf(0,X,dim(L2)-1)↵P :ClrHome :Disp "CHI SQ STAT" :Disp X :Disp "P VALUE" :Disp P

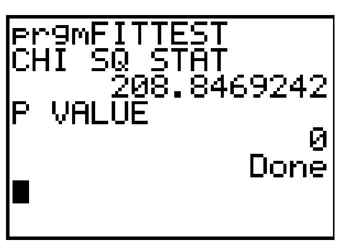
Example 9.5 Cell Phones and Accidents. The following table gives the number of motor vehicle collisions by drivers using a cell phone broken down by days of the week over a 14-month period. Are such accidents equally likely to occur on any day of the week? It appears from the table that weekends have fewer accidents of this type; is this real or just randomness at work?

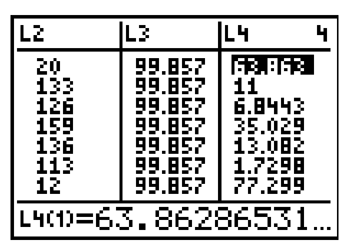
Number of collisions by day of the week

Sun.	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Total
20	133	126	159	136	113	12	699

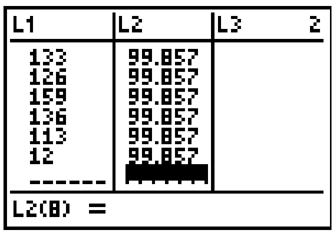
TI-83 Solution. If each day were equally likely, then 1/7 of all accidents should occur on each day. To test the fit of this distribution, we shall use the **FITTEST** program. We first enter **1/7** seven times into list `L1` (the calculator displays a decimal equivalent) to specify the expected distribution, and enter the given frequencies from the table into list `L2`. Next, we run the **FITTEST** program to obtain a *p*-value of 0 from a chi-square test statistic of 208.8469.

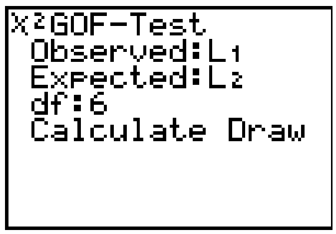


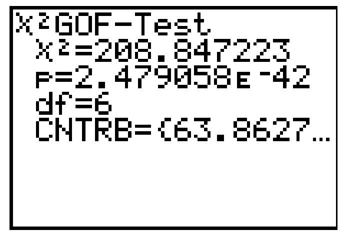




TI-84/89 Solution. Both of these calculators have built-in chi-square goodness-of-fit tests and they function similarly. In both cases, we enter the observed and expected counts into two lists. The expected count under the null hypothesis (accidents are equally likely each day of the week) is 1/7*699 = 99.857. Our observed counts have been entered into `L1` and the expected counts in `L2`. From the STAT TESTS menu select option `χ2GOF-Test` (`Chi2 GOF` on a TI-89). Enter the list names for your observed and expected counts and the degrees of freedom for the test which are $k - 1$, where k is the number of categories. Since there are 7 days per week, for this example $k = 7$, so we have `df` = 6. The TI-84 output seen below shows the contributions to the χ^2 statistic. To see all of them, use the right-arrow key. The TI-89 creates a new list of components which can be viewed in the List editor.



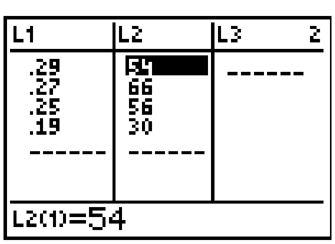


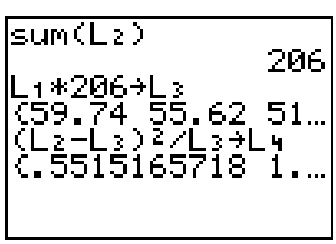


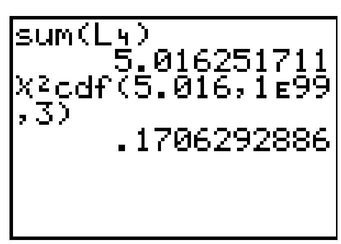
Conclusions. If these accidents were equally likely to occur on any day of the week, then there would be no chance of obtaining a sample distribution that differs so much from the expected counts of $1/7*699 = 99.857$ for each day. So we can reject the claim that accidents are equally likely on each day. Looking at the components, we clearly see that weekends are much lower than expected; Wednesday is also higher than expected.

Example 9.6 Sampling College Students. At a particular college, 29% of undergraduates are in their first year, 27% in their second, 25% in their third, and 19% are in their fourth year. But a random survey found that there were 54, 66, 56, and 30 students in the first, second, third, and fourth years, respectively. Use a goodness-of-fit test to examine how well the sample reflects the college's population.

Solution. Using commands on a TI-83,we shall to test the goodness of fit to examine how well the sample reflects the college's population. We have entered the stated proportions **0.29**, **0.27**, **0.25**, and **0.19** into list `L1` and the observed frequencies into list `L2`. On the home screen, we first sum `L2` to obtain the total number in the sample, then multiply the total by the proportions in L1 to obtain the expected counts which are stored in `L3`. The components of the χ^2 statistic, $(O - E)^2/E$ are stored in `L4`. We sum `L4` to obtain the statistic and use `χ2pdf` from the `DISTR` menu to find the p-value. We obtain a p-value of 0.17061 from a chi-square test statistic of 5.016. For the given distribution, there is about a 17% chance of obtaining observed sample counts that differ as much as these do from the expected counts. We conclude that the sample does not differ significantly from what was expected.

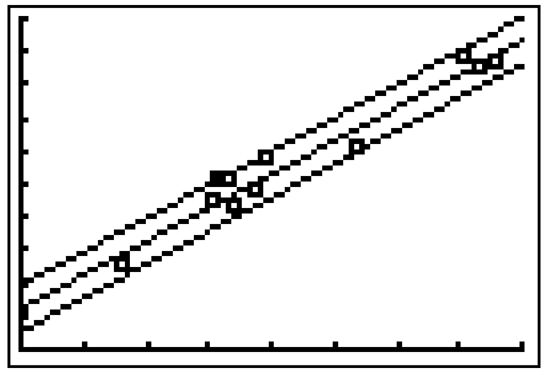






CHAPTER

10



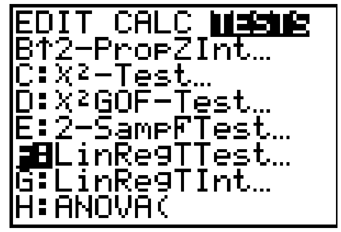
Inference for Regression

Introduction

In this chapter, we provide details on using TI calculators to perform the many difficult calculations for linear regression. In particular, we again find and graph the least-squares regression line and compute the correlation. We can then perform a *t* test to check the hypothesis that the correlation (or, equivalently, the regression slope) is equal to 0. We will also show how to compute a test statistic and find a *p*-value when the hypothesized slope is something other than 1. We also provide a program that computes confidence intervals for the regression slope and also computes a prediction interval for a future observation and a confidence interval for a mean response.

10.1 Simple Linear Regression

We begin by demonstrating the `LinRegTTest` feature (item E) from the `STAT TESTS` menu that will compute a least-squares regression line while simultaneously testing the null hypothesis that the regression slope equals 0. Notice that my TI-84 calculator also has a `LinRegTInt` feature. This is also available on a TI-89. We'll discuss that (and a TI-83 program to perform this) later.

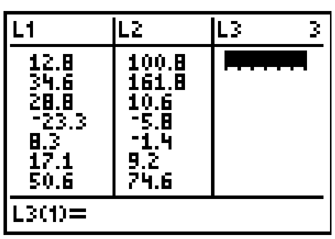


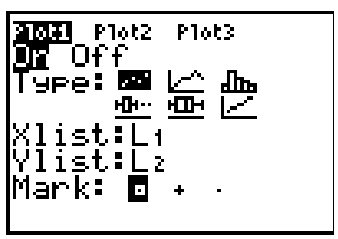
Example 10.1 Stock and Bond Investing. Here are data on the net new money (in billions of dollars) flowing into stock and bond mutual funds from 1985 to 2000. Negative values indicate more money went *out* of the market than went in. We'd like to find out whether or not the relationship is statistically significant. In other words, are the two related, and if so, how?

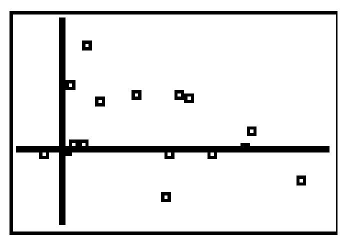
Year	1985	1986	1987	1988	1989	1990	1991	1992
Stocks	12.8	34.6	28.8	–23.3	8.3	17.1	50.6	97.0
Bonds	100.8	161.8	10.6	–5.8	–1.4	9.2	74.6	87.1

Year	1993	1994	1995	1996	1997	1998	1999	2000
Stocks	151.3	133.6	140.1	238.2	243.5	165.9	194.3	309.0
Bonds	84.6	–72.0	–6.8	3.3	30.0	79.2	–6.2	–48.0

Solution. We begin by making a scatterplot as explained in Section 2.1 by entering the data into lists and defining a scatterplot on the `STAT PLOT` screen that uses stocks as the predictor (*X*) variable and Bonds as the response (*Y*) variable. Press θ→ to display the plot. The relationship is fairly linear, but not extremely strong, since there is a good bit of scatter in the plot. The relationship is a decreasing one; if more money went to stocks, less went to bonds.







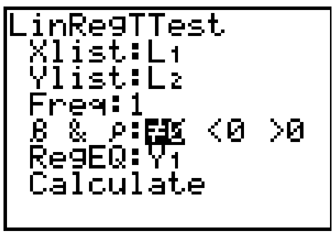
In Chapter 2, we showed how to compute and graph the least-squares line using the `LinReg(a+bx)` command from the `STAT CALC` menu. But now we shall use `LinRegTTest` feature from `STAT TESTS` menu. As before, we first should make sure that the `DiagnosticOn` command has been entered from the `CATALOG`.

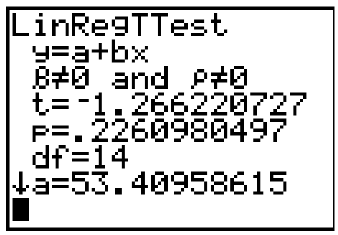
With `LinRegTTest`, we can find the least-squares line while at the same time testing the null hypothesis that the slope β_1 is equal to 0. We will use the alternative hypothesis that

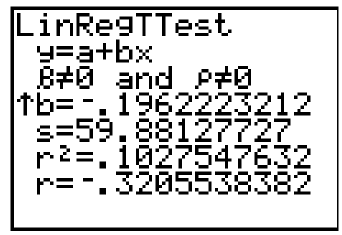
$\beta_1 \neq 0$. This test is equivalent to testing the null hypothesis that the population correlation ρ is equal to 0 with an alternative that $\rho \neq 0$.

After bringing up the `LinRegTTest` screen, adjust the settings for `L1` versus `L2` (or to whatever lists contain the data) and enter the alternative `≠0`. To enter `Y1` for `RegEQ`, scroll down to the right of `RegEQ`, press □, scroll right to `Y-VARS`, press ⊆ for `Function` and then press ⊆ for `Y1`. Then press ⊆ on `Calculate`.

We obtain the regression line *y* = a + b*x* (or $y = \hat{\beta}_0 + \hat{\beta}_1 x$), which is also stored into `Y1`. Here the equation rounds to *y* = 53.4096 – 0.1962*x*. Press σ to see the plot with the regression line.







The *p*-value for the *t* test is given as 0.226 from a *t* statistic of –1.26622. If β_1 were equal to 0, then there would be a 22.6% chance of obtaining a value for *b* as low as –0.1962, or of obtaining a correlation as low as *r* = –0.32, with a sample of this size. This rather high *p*-value means that we do *not* have statistically significant evidence that there is some straight-line relationship between the flows of cash into bond funds and stock funds. In other words, we do not have enough evidence to reject that β_1 = 0 or to reject that ρ = 0.

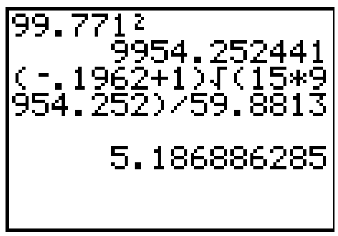
Hypotheses Other Than 0

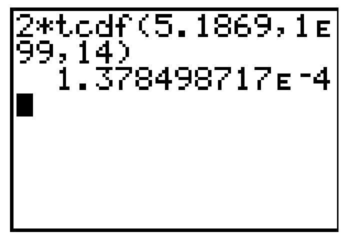
In certain circumstances, one might be interested in a null hypothesis different from $H_0 : \beta = 0$. The general *t* statistic for a test of this type is

$$t = \frac{b_1 - \beta_{10}}{s(b_1)} = \frac{b_1 - \beta_{10}}{s_e\sqrt{\dfrac{1}{\sum (x-\bar{x})^2}}} = \frac{(b_1 - \beta_{10})\sqrt{(n-1)s_x^2}}{s_e}$$

with *n* – 2 degrees of freedom. In this formula, β_{10} is the hypothesized slope, s_e is the standard deviation of the residuals around the regression line, and s_x^2 is the sample variance of the predictor (*x*) variable. This *t* statistic has *n* – 2 degrees of freedom.

Example 10.2 Another Hypothesis About the Slope. In Example 10.1, we might be concerned that in any given month there are only so many dollars available to invest. Let's call that *D*. If investors were to split their money between stocks and bonds, we would have an equation *S* + *B* = *D*. This would argue that the regression relationship should be of the form $B = D - S$, and we should have slope $\beta_1 = -1$. We will test this as H_0 against H_A: $\beta_1 \neq -1$. We had $b_1 = -0.1962$, $s_e = 59.8813$ (given as s on the calculator output), and using `1-Var Stats`, we find $s_x^2 = 99.771^2 = 9954.252$. Putting this together with the fact that we had $n = 16$ data points, we find a t statistic of 5.1869. To find the p-value for this two-sided test, we will double the area above $t = 5.1869$ under the t distribution curve with 14 degrees of freedom. Our p-value is 0.00014, so we have sufficient evidence based on this sample that there is not an inverse relationship between net investments in stocks and bonds (the true slope is not –1.)



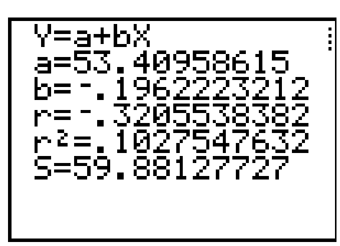


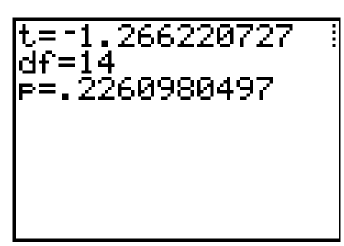
Confidence Intervals in Regression Inference

We provide a program called **REGINF** (the listing is at the end of this chapter) that will compute confidence intervals for the slope and for the intercept of the linear regression model $y = \beta_0 + \beta_1 x$, as well as confidence intervals for the mean value of y at a given x and prediction intervals y given a new value of x. TI-84 calculators have a built-in function to compute confidence intervals for the slope (it is shown on the first line of the output, the rest is similar to `LinRegTTest`); TI-89 calculators will find all these intervals (but only at 95% confidence, since that is the most common value) given the right input. The program below can be used on TI-83/84 calculators for more general confidence and prediction intervals.

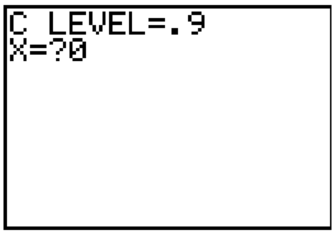
Example 10.3 Stocks and Bonds Continued. Using the data from Example 10.1, find the 90% confidence intervals for the slope β_1, and intercept β_0 of the linear regression model.

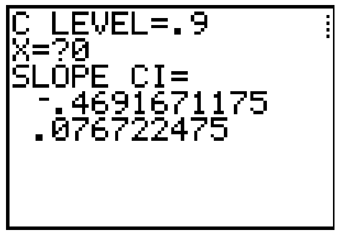
TI-83/84 Program Solution. With the data already entered into lists `L1` and `L2`, bring up the **REGINF** program. The program first asks for the list names containing your data. It then displays the regression output we have seen before from `LinRegTTest`. Notice the scrolling dots at the upper right of the screen. Press ⊆ to continue on.

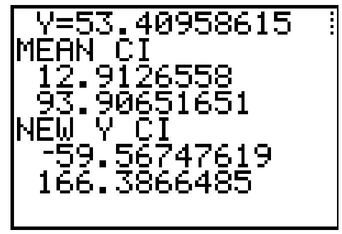




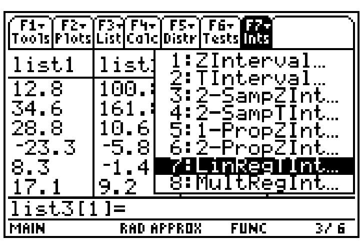
You are now asked for the confidence level. Here we want 90%, so enter .9. You are now asked what *x* value we are interested in. Since we are interested in the intercept, enter 0. Press ⊆ to continue. We first find a 90% confidence interval for the slope. Based on these data we are 90% confident the true slope for predicting net bond dollars based on net stock dollars is between –0.469 and 0.077. Notice that our original null hypothesis H_0: $\beta_1 = 0$ is contained in the interval. Press ⊆ again and we will find the value predicted from the line at our desired *x* value (notice it is exactly the intercept here) and a confidence interval for the mean response *y* at that *x* and a prediction interval for a new observation at *x*. Press ⊆ again and the program will display a "`Done`" message.

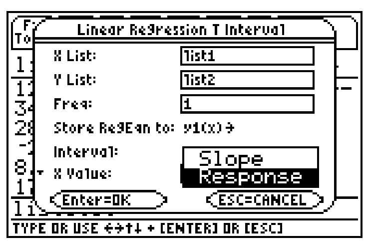


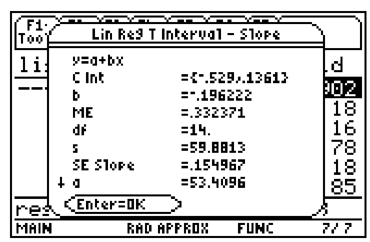




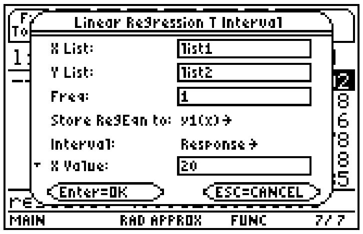
TI-89 Solution. The data have been entered in `list1` and `list2`. From the 2□=□ `Ints` menu, select option `7:LinRegTInt`. Use 2| = ° to enter the list names. Store the regression equation (if desired). We can use the right arrow to expand the `Interval` box to select the type of interval: `Slope` or `Response`. If we select for a Slope, press ⊆ to perform the calculations; if for a Response, you can enter the particular *x* value of interest (for the intercept this would be 0). Here we see that a 95% confidence interval for the slope is (–0.529, 0.126). We also see the standard error of the slope, $s(b_1)$, labeled as `SE Slope`. Press the down arrow to find r and r^2.

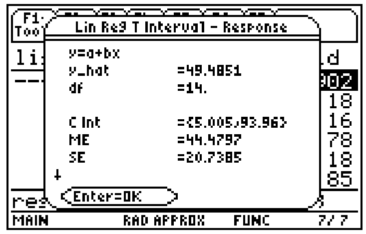


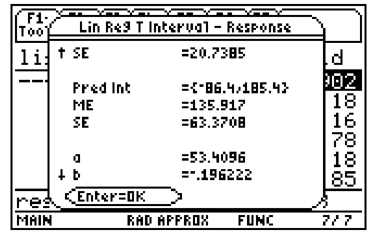




Suppose we wanted to find confidence and prediction intervals for the amount of money going into bonds when 20 billion dollars is going to stocks. Use the same as above, but select `Response` as the `Interval` type and type 20 in the `X Value` box. The first portion of the output gives the $\hat{y}$ prediction from the line. Further down, we see that, with 95% confidence, on average there should be between 5.005 and 93.96 billion dollars going into bonds when 20 billion dollars is going into stocks. We are also explicitly told the margin of error and the standard error of the confidence interval. Press the down arrow to find the prediction interval information. For a particular month with 20 billion dollars going into stocks, with 95% confidence, somewhere between –86.4 and 185.4 billion dollars will go into bonds.





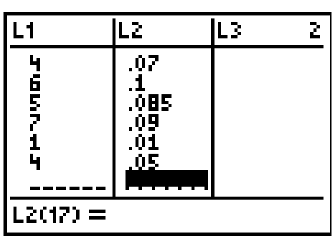


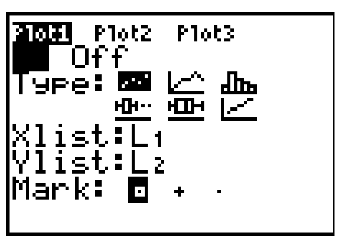
Example 10.4 Beer and Blood Alcohol. Several years ago a study was conducted at The Ohio State University in which 16 student volunteers were randomly assigned to drink a number of cans of beer. The students were equally divided between men and women and varied in weight and normal drinking habits. Thirty minutes after they finished drinking, a police officer measured their blood alcohol content (BAC) in grams of alcohol per deciliter of blood. The data are displayed below. Investigate the relationship between number of beers drunk and blood alcohol level.

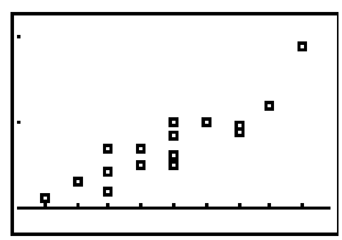
Student	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Beers	5	2	9	8	3	7	3	5
BAC	0.10	0.03	0.19	0.12	0.04	0.095	0.07	0.06

Student	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Beers	3	5	4	6	5	7	1	4
BAC	0.02	0.05	0.07	0.10	0.085	0.09	0.01	0.05

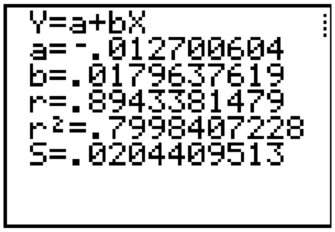
Solution. First, plot the data. Is the relationship linear enough for our regression to make sense? The data have been entered into lists L1 and L2. We define the scatterplot and press θ→ to display it. The plot is very linear, but the individual represented at the upper right (9 beers, BAC 0.19) might be influential to this regression.

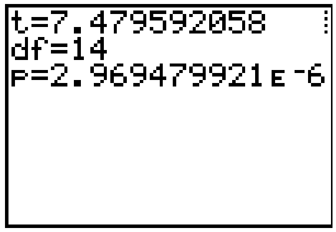


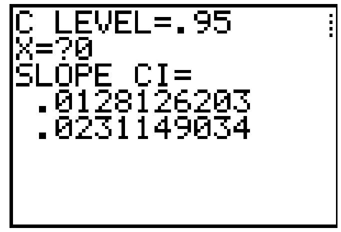




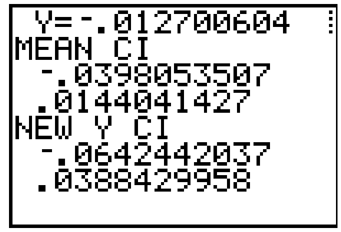
Using program **REGINF**, we find the following: The regression equation is BAC = –0.127 + 0.180*Beers. About (r^2) 80% of the overall variation in BAC is accounted for by the number of beers consumed. The standard deviation of the residuals around the line is s = 0.0204. The relationship is highly significant. With a p-value of 0.00000297, we're convinced based on this sample that there is a linear relationship between the number of beers consumed and BAC. Further, we are 95% confident the mean blood alcohol level increases between 0.0128 and 0.0231 per beer.







Last, the intercept is *not* significantly different from 0, since 0 is included in the confidence interval for the value when $x = 0$. This makes sense; when a person hasn't had anything alcoholic to drink, they should have a BAC of 0.

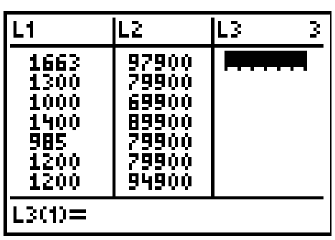


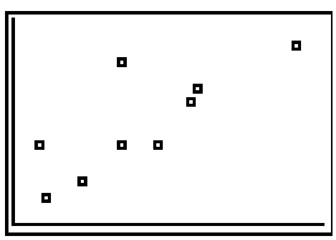
Example 10.5 Augusta House Prices. Not too long ago, the author of this manual was involved in the purchase of a house for her daughter in Augusta, GA. The sizes of the houses in the first group we looked at (all older homes) and their asking prices are:

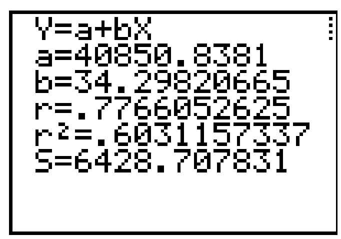
Sq. Ft.	1663	1300	1000	1400	985	1200	1200	1100	1384
Price ($)	97,900	79,900	69,900	89,900	79,900	79,900	94,900	73,000	87,595

Find a linear regression model to relate the size of the house to the asking price, and a confidence interval for the slope. Find an interval estimate for the average price of a house with 1200 square feet of space.

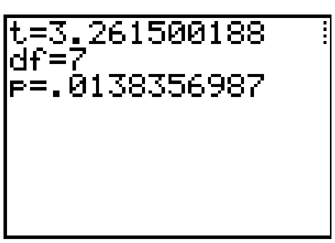
Solution: Enter the data into two lists, plot the data, then use program **REGINF** to find the regression equation and the confidence interval for the slope. First, we notice that there appears to be a strong relationship between the size of the house and the asking price; however, there also seem to be two outliers which may affect our results.

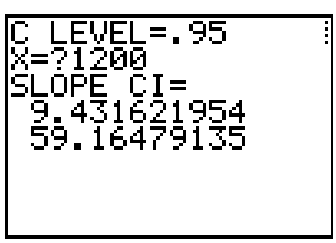


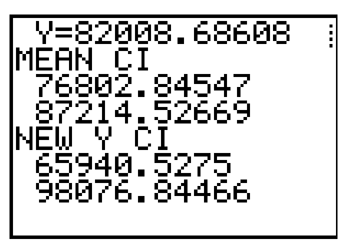




The regression is $Asking\$ = 40850.84 + 34.30 * sqft$. Further, we see that the size of the house explains about 60% of the variation in asking price (other lurking variables like age, materials, location, will explain more).







From confidence interval for the slope, we see that we are 95% confident the average cost per square foot in Augusta at that time (for houses of this type) was somewhere between \$9.43 and \$59.16. Based on this sample, the average asking price of older homes in Augusta at this time with 1200 square feet should have been between \$76,802.85 and \$87,214.53 with 95% confidence.

Intervals Without a Program

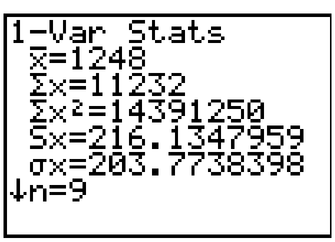
Keying in a program can be tedious, if you do not have access to a TI-connect cable. Here, we demonstrate how to use quantities available with standard calculator output to create confidence or prediction intervals of your own. We will use two basic formulas:

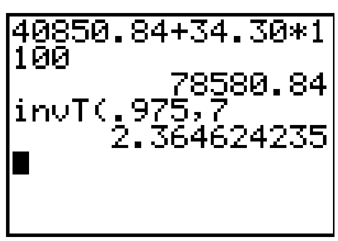
Confidence interval for the mean response at x: $\hat{y} \pm t^* s_e \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum (x-\bar{x})^2}}$

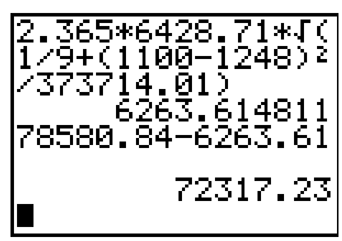
Prediction interval for a new response at x: $\hat{y} \pm t^* s_e \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum (x-\bar{x})^2}}$

Example 10.6. More on House Prices. Based on the model found in Example 10.5, what should be the average price for an older home in Augusta with 1100 square feet? Use a 95% confidence interval estimate.

Solution. As demonstrated in Example 10.2, we can use `1-Var Stats` on the list containing our x variable to find both `Sx` and $\bar{x}$. Since $\sum (x-\bar{x})^2 = (n-1)s_x^2$ we have $\sum (x-\bar{x})^2 = 8 * 216.1348^2 = 373714.01$. We also see n = 9. From calculating the regression, we have $s_e = 6428.71$. Using `invT` from the `DISTR` menu (or a table), we find t^* = 2.365. Also, using our equation for the regression, we have a point estimate of $\hat{y} = 40850.84 + 34.20 * 1100 = \$78,580.84$. Putting it all together, we find a margin of error for this confidence interval is 6263.61. The lower end of the confidence interval is \$72,317.23, and we similarly find the upper end is \$84,844.45.

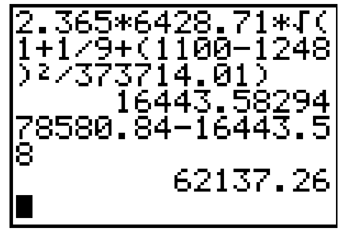






Example 10.7 Still More House Prices. Suppose the real estate agent told us about a "really nice" home with 1100 square feet of space. What should its asking price be, based on this model with 95% confidence?

Solution. One particular home is not an average. In this case, we need a prediction interval for a new observation. We found all the needed pieces in Example 10.5. All we need to do is add another 1 under the square root to account for the fact that individuals are more variable than means. The new margin of error is $16,443.58. The lower end of the 95% prediction interval is $62,137.26. Changing the minus sign to a plus, we find the upper end of the prediction interval is $95,024.42.



10.2 More Detail about Simple Linear Regression

Analysis of variance (ANOVA) is another method to test the null hypothesis H_0: $\beta_1 = 0$, with an alternative H_A: $\beta_1 \neq 0$. Although really more useful in a true ANOVA or multiple regression setting, in the case of simple linear regression it has a use and relation to a *t* test similar to that of using a χ^2 test for two proportions. The **REGANV** program on page 118 performs such a test, stores an ANOVA table into lists L4, L5, and L6, and displays the associated *F*-statistic and *p*-value. Before executing the **REGANV** program, we must enter paired data into lists L1 (*x*) and L2 (*y*).

Note: The ANOVA table generated by the **REGANV** program is not the same one that will be generated by the calculator's built-in ANOVA(command that is used for testing whether or not several populations have identical means. That command will be explained in Chapter 12.

The REGANV Program for Linear Regression ANOVA

```
PROGRAM:REG3
:LinReg(a+bx) L1,L2
:sum(seq((a+bL1(I)-□)²,
  I,1,dim(L1)))↵A
:sum(seq((RESID(I))²,I,1,
  dim(L1)))↵B
:sum(seq((L2(I)- □)²,
  I,1,dim(L1)))↵C
:(n-2)A/B↵F
:1-Fcdf(0,F,1,n-2) ↵P
:ClrList L4,L5,L6
:1↵L4(1):A↵L5(1):A↵L6(1)
:n-2↵L4(2):B↵L5(2):B/(n-2) ↵L6(2)
:n-1↵↵L4(3):C↵L5(3):C/(n-1) ↵L6(3)
:ClrHome
:Disp "F-STAT, P-VAL"
:Output(2,2,{round(F,3),round(P,4)})
:Output(4,1,"SEE L4, L5, L6")
:Output(5,5,"DF, SS, MS")
:Output(6,3,"M")
:Output(7,3,"E")
:Output(8,3,"T")
:Stop
```

Example 10.8 More Stocks and Bonds. Consider the data from Example 10.1 on the net new money (in billions of dollars) flowing into stock and bond mutual funds from 1985 to 2000.

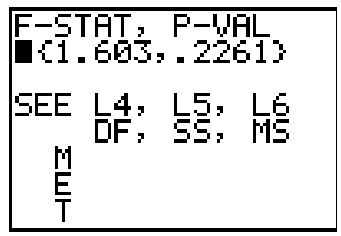
Year	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**	**1992**
Stocks	12.8	34.6	28.8	–23.3	8.3	17.1	50.6	97.0

Bonds	100.8	161.8	10.6	–5.8	–1.4	9.2	74.6	87.1

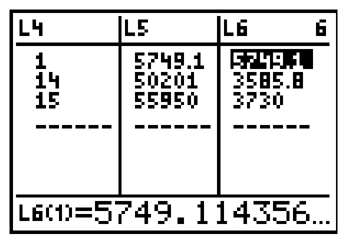
Year	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Stocks	151.3	133.6	140.1	238.2	243.5	165.9	194.3	309.0
Bonds	84.6	–72.0	–6.8	3.3	30.0	79.2	–6.2	–48.0

(a) Construct the ANOVA table.
(b) State and test the hypotheses using the ANOVA *F*-statistic.
(c) Give the degrees of freedom for the *F*-statistic for the test of H_0.
(d) Verify that the square of the *t* statistic for the equivalent *t* test is equal to the *F*-statistic in the ANOVA table.

Solution. (a) To construct an ANOVA table with the **REGANV** program, we must first enter our data into lists `L1` (for stocks) and `L2` (for bonds). After doing so, bring up and execute the program. (There is nothing to input, so just press ⊆.)



The program displays the *F*-statistic and *p*-value. The ANOVA table is stored in lists `L4`, `L5`, and `L6` and contains the degrees of freedom (DF), the sum of squares (SS), and the mean square (MS) for each of the model (M), error (E), and total (T). Notice that the *p*-value for the *F* test is exactly what we found in Example 10.1 with `LinRegTTest` (page 111).



The ANOVA table is displayed at right. The model has 1 degree of freedom here since we had one predictor variable. Total degrees of freedom in the last line is $n - 1$. Its sum of squares would be the numerator in computing the variance of the *y* variable in the regression, and Mean Square is the variance of *y*. If we divide SSM by SST we have 5749.1/55950 = 0.10275 which was the r^2 found in Example 10.1.

(b) The ANOVA test is about the linear regression slope β_1. We test the null hypothesis H_0: $\beta_1 = 0$ with the alternative H_A: $\beta_1 \neq 0$. With the above *p*-value of 0.2261, we do not have strong evidence in this case to reject H_0 in favor of the alternative.

(c) The degrees of freedom for the *F*-statistic are given by 1, $n - 2$, which in this case is 14. This number is the same as the degrees of freedom of the error (E) displayed in the ANOVA table.

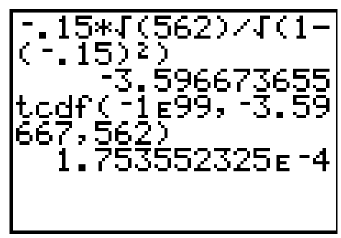
(d) The `LinRegTTest` was used in Example 10.1. The *t* statistic was computed as $t = -1.266220727$. If we square this value, then we obtain 1.603314929, which is the actual value of the displayed (rounded) *F*-statistic from the ANOVA test.

Sample Correlation and the *t* Test

One may be required to perform a correlation *t* test without an actual data set, but instead by using only the values of the sample correlation *r* and the sample size *n*. Although TI calculators do not have a built-in procedure for this type of test, the *t* statistic and *p*-value are easily calculated in this case. Here the test statistic, which follows a $t(n–2)$ distribution, is given by

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Example 10.9 Avoidance of New Foods. In a study of 564 children who were 2 to 6 years of age, the relationship of food neophobia (avoidance of unfamiliar foods) and the frequency of consumption gave a correlation of $r = –0.15$ for meat. Perform a significance test about the correlation of meat neophobia and the frequency of meat consumption among children 2 to 6 years of age.



Solution. We shall test H_0: $\rho = 0$ versus H_A: $\rho < 0$. The rejection region is a left tail; thus the *p*-value will be the left-tail probability of the $t(n–2) = t(562)$ distribution. We compute the *t* statistic and *p*-value "manually" by entering $-.15 * \sqrt{562} / \sqrt{1-(-.15)^2}$ to obtain $t = –3.59667$. Next, we compute the *p*-value $P(t(562) \le -3.59667)$ with the `tcdf(` command from the `DISTR` menu. We enter `tcdf(-1E99,-3.59667,562)` to obtain a *p*-value of 1.75×10^{-4}.

If the true correlation were 0, then there would be only a 0.000175 probability of obtaining a sample correlation as low as $r = –0.15$ with a random sample of size 564. We have statistical evidence to reject H_0 in favor of the alternative that $\rho < 0$. We note that for a two-sided alternative, then our *p*-value would be $2*(1.75\times10^{-4}) = 3.5\times10^{-4}$, which would still be small enough for us to reject H_0.

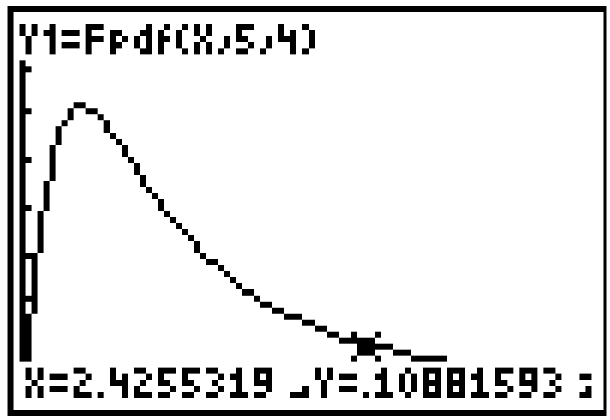
The REGINF Program for Linear Regression Inference

```
:Input "X LIST=",LX
:Input "Y LIST=",LY
:FnOff
:LinRegTTest LX,LY,0,Y1:
:σx^2n↲V:s↲S:⌡↲M:n↲N
:ClrHome
:Output(1,2,"Y=a+bX"
:Output(2,2,"a="
:Output(2,4,a
:Output(3,2,"b="
:Output(3,4,b
:Output(4,2,"r="
:Output(4,4,r
:Output(5,2,"r^2="
:Output(5,5,r^2
:Output(6,2,"S="
:Output(6,4,S
:Pause
:df↲K
:ClrHome
:Output(1,1,"t="
:Output(1,3,t
:Output(2,1,"df="
:Output(2,4,K
:Output(3,1,"p="
:Output(3,3,p
:Pause
```

```
:ClrHome
:Input "C LEVEL=",C
:Prompt X
:K+1↲K:b↲A
:TInterval 0, √ (K),K,C
:upper↲T
:A-T*S/√ (V)↲P
:A+T*S/√ (V)↲Q
:Output(3,1,"SLOPE CI="
:Output(4,2,P
:Output(5,2,Q
:Pause
:Y1-ST√(N^-1+(X-M)^2/V)↲P
:Y1+ST√(N^-1+(X-M)^2/V)↲Q
:Y1-ST√(1+N^-1+(X-M)^2/V)↲J
:Y1+ST√(1+N^-1+(X-M)^2/V)↲U
:ClrHome
:Output(1,2,"Y="
:Output(1,4,Y1
:Output(2,1,"MEAN CI"
:Output(3,2,P
:Output(4,2,Q
:Output(5,1,"NEW Y CI"
:Output(6,2,J
:Output(7,2,U
:Pause
:ClrHome
```

CHAPTER

11



Multiple Regression

Introduction

In this chapter, we demonstrate how to use a program for TI-83 and -84 calculators to calculate a multiple linear regression model. We also show how to do this with built-in functions on a TI-89.

11.1 Inference for Multiple Regression

The **MULREG** program below computes the regression coefficients and an ANOVA table for the multiple linear regression model $\mu_Y = \beta_0 + \beta_x x_1 + \ldots + \beta_k x_k$. The program stores the regression coefficients in `L1`, their standard errors in `L2`, individual *t* statistics in `L3`, and *p*-values in `L4`. The squared multiple correlation coefficient (r^2), the *F*-statistic, the *p*-value, and the standard deviation of the residuals are also displayed along with the entries in the ANOVA table. The program can also be used to find prediction and confidence intervals for the response given a set of predictor values.

The MULREG Program

```
Disp "DATA IN COLS","OF [A]"
dim([A]) ↵θ1:Lθ1(1)↵N:Lθ1(2) ↵L:L-1↵K
{N,1}↵dim([B]):Fill(1,[B])
augment([B],[A]) ↵ [B]
[B]^T[B] ↵ [D]
seq([D](I,1),I,2,L+1)/N↵Lθ1
{L,L}↵dim([C]):Ans↵dim([E])
{N,1}↵dim([C])
Input "RESPONSE COL=",R
For(I,1,N):[A](I,R) ↵ [C](I,1):End
If R≤K:Then
For(J,R+1,L):For(I,1,N)
[B](I,J+1) ↵ [B](I,J)
End:End
[B]^T[B] ↵ [D]
End
{L,L}↵dim([D]):[D]□↵ [D]
{N,L}↵dim([B])
[B]^T [C]:[D]Ans↵ [E]
Matr▶list([E],Lθ1)
Disp "COEF Lθ1=":Pause Lθ1
[B][E] ↵ [B]
Matr▶list([C],LY)
Matr▶list([B],LYP)
mean(LY) ↵Y
sum((LY-Y)^2)↵T
sum((LY-LYP)^2↵E
LYP↵L5
LY-LYP↵L6
DelVar LY:DelVar LYP
N-L↵M:T-E↵R:R/K↵Q:E/M↵D:√(D)
↵S:Q/D↵F
S√(seq([D](I,I),I,1,L)) ↵Lθ2
Disp "STDEV Lθ2=":Pause Lθ2
Menu("PREDICT Y?","YES",A,"QUIT",B)
Lbl A
Input "X{}=",Lθ5
Input "CONF. LEVEL=",C
{L,1}↵dim([C])
1↵ [C](1,1)
Lθ1/Lθ2↵Lθ3
Disp "T-RATIO Lθ3=":Pause Lθ3
Disp "COEF P Lθ4="
1-2seq(tcdf(0,abs(Lθ3(I)),M),I,1,L) ↵Lθ4
Pause Lθ4
1-(N-1)*D/T↵A
ClrHome
Output(1,1,"S="
Output(1,9,S
Output(2,1,"R^2="
Output(2,9,R/T
Output(3,1,"R^2ADJ="
Output(3,9,A
Output(4,1,"REG DF="
Output(4,9,K
Output(5,1,"ERR DF="
Output(5,9,M
Output(6,1,"TOT DF="
Output(6,9,N-1
Pause
ClrHome
Output(1,1,"SS REG="
Output(1,9,R
Output(2,1,"SS ERR="
Output(2,9,E
Output(3,1,"SS TOT="
Output(3,9,T
Output(4,1,"MS REG="
Output(4,9,Q
Output(5,1,"MS ERR="
Output(5,9,D
Output(6,1,"F=    "
Output(6,9,F
Output(7,1,"P-VAL="
1-Fcdf(0,F,K,M) ↵P
round(P,5) ↵P
Output(7,9,P
Pause
ClrHome
Lbl C
```

```
For(I,1,K):Lθ5(I) ↲ [C](I+1,1):End
[E]^T[C]
Ans(1,1) ↲Y
ClrHome
round(Y,5) ↲Y
Output(1,1,"Y HAT="
Output(1,8,Y
[C]^T[D][C]
```

```
Ans(1,1) ↲V
round(S√(V),5) ↲T
Output(2,1,"S(FIT)="
Output(2,8,T
TInterval 0,√(M+1),M+1,C
upper↲T
Output(3,1,"C.I.="
Y+TS√(V) ↲D
Y-TS√(V) ↲E
Output(4,3,E
Output(5,3,D
Output(6,1,"P.I.="
Y-TS√(1+V) ↲E
Y+TS√(1+V) ↲D
```

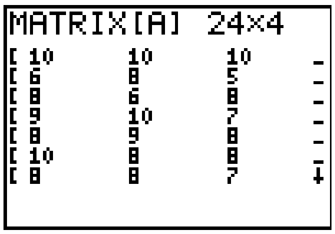
To execute the program, we first must enter our sample data into matrix `[A]` with the dependent (*Y*) variable in the last column.

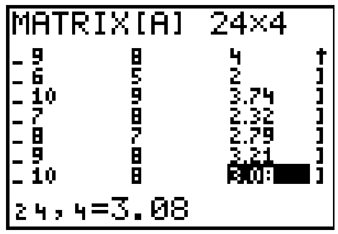
Example 11.1 Predicting Grades. Consider this data from a sample of 24 students at a large university that uses a 4.0 GPA grade scale. Run a multiple regression analysis for predicting the GPA from the three high school grade variables (Math, Science, and English where 10 = 'A', 9 = 'A–', 8='B+', etc).

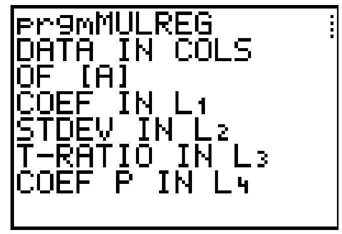
OBS	GPA	HSM	HSS	HSE	SATM
1	3.32	10	10	10	670
2	2.26	6	8	5	700
3	2.35	8	6	8	640
4	2.08	9	10	7	670
5	3.38	8	9	8	540
6	3.29	10	8	8	760
7	3.21	8	8	7	600
8	2.00	3	7	6	460
9	3.18	9	10	8	670
10	2.34	7	7	6	570
11	3.08	9	10	6	491
12	3.34	5	9	7	600
13	1.40	6	8	8	510
14	1.43	10	9	9	750
15	2.48	8	9	6	650
16	3.73	10	10	9	720
17	3.80	10	10	9	760
18	4.00	9	9	8	800
19	2.00	9	6	5	640
20	3.74	9	10	9	750
21	2.32	9	7	8	520
22	2.79	8	8	7	610
23	3.21	7	9	8	505
24	3.08	9	10	8	559

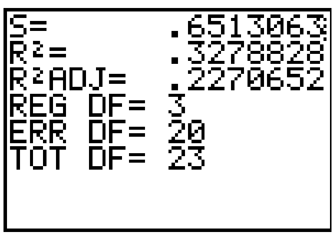
TI-83/84 Solution. We use the model $\mu_{GPA} = \beta_0 + \beta_1 x_{HSM} + \beta_2 x_{HSS} + \beta_3 x_{HSE}$, which is of the form $\mu_Y = \beta_0 + \beta_x x_1 + \beta_2 x_2 + \beta_3 x_3$. In this case, there are 24 sample points and four variables. We must first enter the data into matrix [A], much as we have been entering data into lists. Matrix [A] should have 24 rows and 4 columns. However, to execute the program correctly, the dependent variable must be in the *last* column.

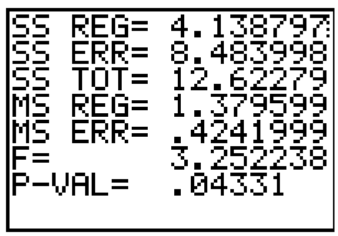
In the MATRIX EDIT screen, enter the dimensions of matrix [A] as 24 x 4, then enter the data. The calculator enters data into a matrix by going sequentially across the rows. So the first row will be 10 10 10 3.32, and the second row will be 6 8 5 2.26. After the data is entered into matrix [A] press ψζ to exit the Matrix editor and then execute the **MULREG** program.

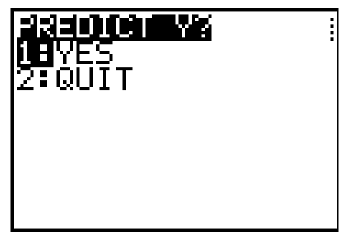










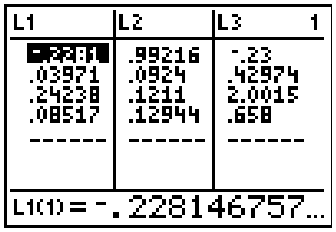


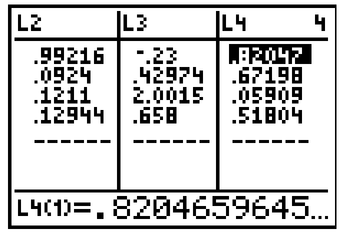
The program first reminds you that it is using matrix A as input. It then reminds you where the coefficients and their associated statistics are. Press ÷ to continue. Next, we see the standard deviation of the residuals, R^2 and R^2 adjusted (a penalized R^2 for having predictors in the model that are not useful), and the degrees of freedom for the ANOVA table. Press ÷ to continue. Next we see the sums of squares for the ANOVA table, the F statistic and its p-value. Press ÷ to continue. We are now asked if we want to use the model to predict a response given a set of predictors. For now, press ♥ to quit.

The displayed ANOVA table gives the degrees of freedom (DF), the sum of squares (SS), and the mean square (MS) for each of the regression model, error, and total. The squared correlation coefficient, the F-statistic and p-value of the F-test, and the standard deviation S are also displayed. The value of R^2 here is telling us that only 32.8% of the variation in GPAs is explained by the high school variables.

The null hypothesis for the F-test is that each variable coefficient in the linear regression model is equal to 0 or H_0: $\beta_1 = \beta_2 = \beta_3 = 0$. In this case, the p-value is 0.0433, which is a value usually considered low enough to be statistically significant. Thus, even with this small sample of size 24, we have enough evidence to reject H_0.

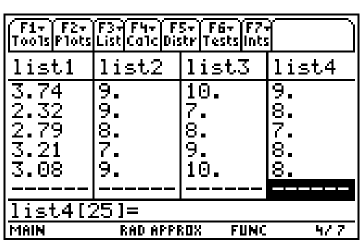
The program stores the computed regression coefficients and their associated statistics in `L1` through `L4`. In this case, our model becomes μ_{GPA} = –0.228 + 0.0397*HSM + 0.2424*HSS + 0.0852*HSE. The Standard errors are in `L2`, with the associated t statistics in `L3`. Finally, the p-values for testing individually H_0: $\beta_i = 0$ are in `L4`. In this case, the only coefficient statistically different from 0 is for the second predictor, high school Science.

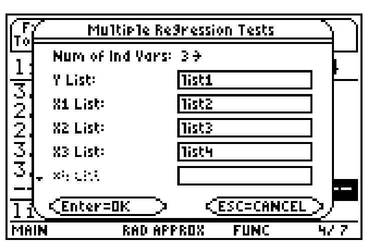


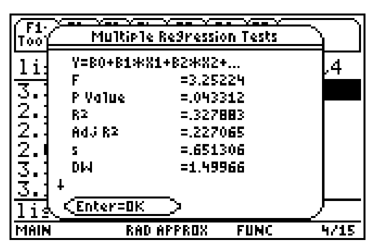


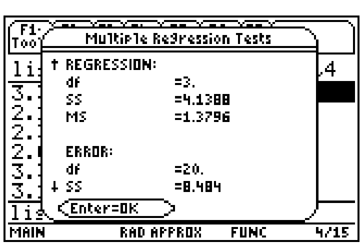
TI-89 Solution. Enter the data into four lists. From the `Tests` menu, select option `B:MultRegTests`. The first input item asks for the number of independent variables. Press the right arrow to expand the box. Here, select 3. Then use ° to input the name of the list containing the dependent (Y) variable and the list names for the independent variables.

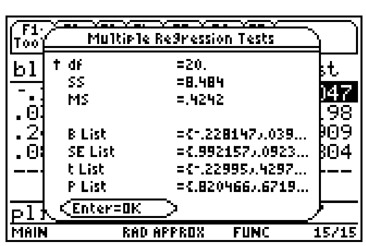
The first portion of the output shows the overall F-statistic for the model and its p-value, the multiple coefficient of determination (R^2) and the adjusted R^2 (this penalizes R^2 for adding predictor variables into a model that aren't helpful), the standard deviation of the residuals around the regression model and the Durbin-Watson statistic (a measure of autocorrelation among the residuals — since order of entry in this model has no real meaning, we ignore this here). Scrolling down the output, we find the ANOVA table for the model. We also see the first part of B list (list of coefficients for the model), SE list (the standard errors of the coefficients), T list (list of t statistics for each coefficient), and P list (list of p-values for each t statistic). Press ÷ to see these lists have been added into the List Editor. Scrolling left through the List Editor, we also see that a list of predicted values for each observation (`yhat1`), residuals (`resid`), standardized residuals (`sresid`), and leverage (`lever`), a measure of the influence of each observation on the regression, have also been added.

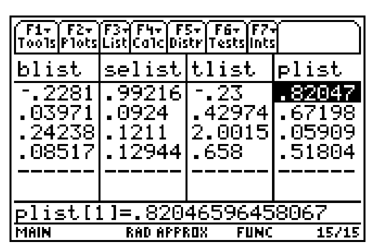






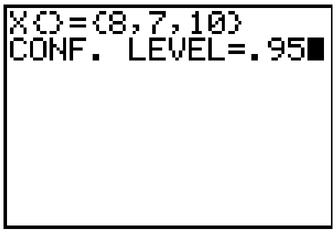


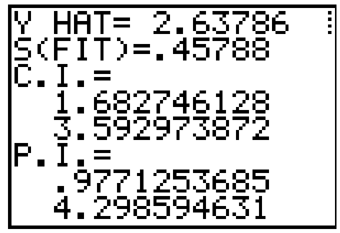




Example 11.2 More Grade Predictions. Use the preceding regression model to predict the GPA of a student with HSM = 8, HSS = 7, and HSE = 10.

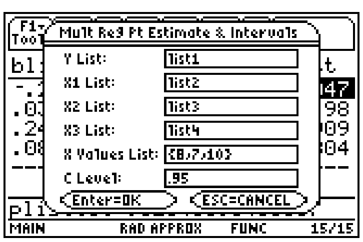
TI-83/84 Solution. Rerun program MULREG, but this time press ÷ to answer `YES` to the `Predict Y` question. You are next prompted to enter the X values. Enter them in curly braces (⍓χ and ⍓δ) separated by commas. Press ÷. You are now asked for the confidence level. Enter it as a decimal. Press ÷. The next screen gives the actual prediction from the model along with the standard error of that prediction, the confidence interval, and the prediction interval.

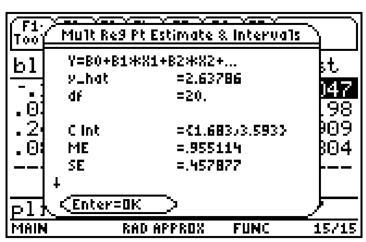


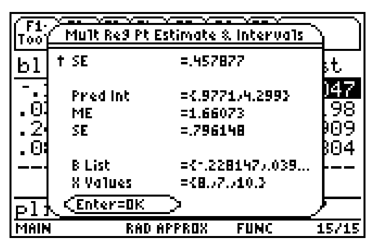


We see that the model predicts 2.64 as the GPA for students with these high school grades. We are 95% confident the *average* GPA for *all* students with these grades will be between 1.68 and 3.59. For an *individual* student with these grades, we would predict a GPA between 0.98 and 4.30 (4.0) with 95% confidence. The intervals are wide due to the low r^2 for the model.

TI-89 Solution. From the □ `Ints` menu, select option `8:MultRegInt`. You will first be asked for the number of independent variables (which should already be set to 3 from the prior regression). The list of values to predict for (these must be entered within curly braces which are ⍓χ or ⍓δ) and the confidence level. Pressing ÷, we see the predicted value from the equation, a confidence interval for the average GPA of all students with this combination of high school grades, a prediction interval for a particular student with these grades, and the beginning of the coefficient list.

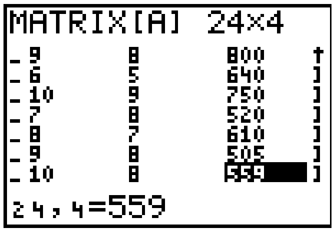


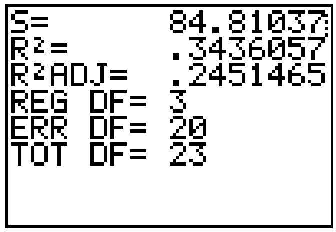


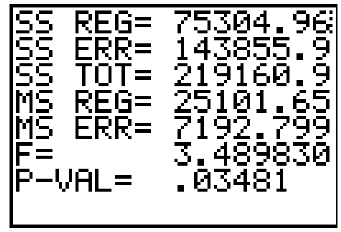


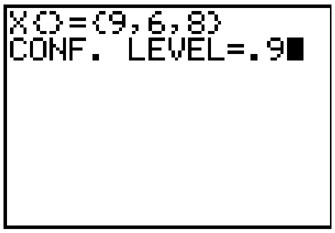
Example 11.3 Predicting SAT Math Scores. With the above 24 sample points, perform a regression analysis to predict SATM from the three high school grade variables. Then find the predicted SATM for a student with HSM = 9, HSS = 6, and HSE = 8, using 90% confidence.

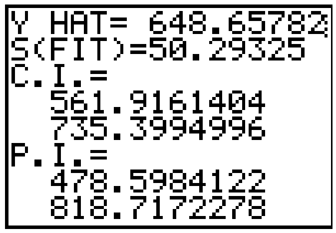
Solution. First, we must edit the last column of matrix [A] so that it contains the SATM scores. To enter these easily, use the down arrow after each entry rather than pressing ÷. Then rerun the **MULREG** program to obtain a new regression model.

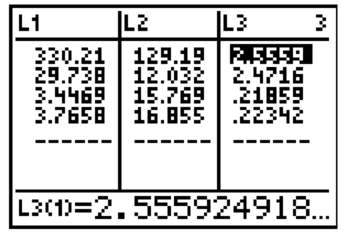












The new model is μ_{SATM} = 330.21 + 29.738*HSM + 3.4469*HSS + 3.7658*HSE. The displayed *p*-value of 0.0348 gives us statistical evidence to reject that all regression coefficients of HSM, HSS, and HSE are 0. Thus, at least one coefficient is nonzero; observing the *t* statistics in L3, the only one that is large is for the first predictor variable (high school Math grade). The model predicts a SATM score for a student with these high school grades of about 649 (practically 650); however, the standard error of this prediction is large.

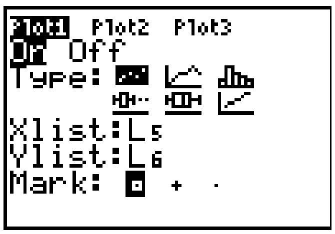
11.2 A Case Study

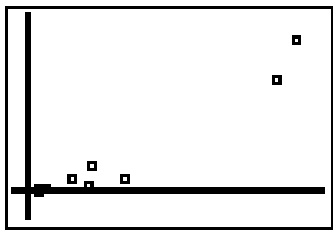
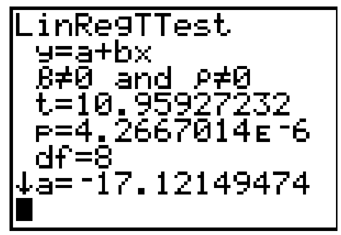
Example 11.4 On-line Brokerages. Below are some data for the top ten internet brokerages. The variables are the market share of the firm, the number of internet accounts in thousands, and the total assets of the firm in billions of dollars.

ID	Broker	Mshare	Accounts	Assets
1	Charles Schwab	27.5	2500	219.0
2	E*Trade	12.9	909	21.1
3	TD Waterhouse	11.6	615	38.8
4	Datek	10.0	205	5.5
5	Fidelity	9.3	2300	160.0
6	Ameritrade	8.4	428	19.5
7	DLJ Direct	3.6	590	11.2
8	Discover	2.8	134	5.9
9	Suretrade	2.2	130	1.3
10	National Discount Brokers	1.3	125	6.8

(a) Use a simple linear regression to predict assets using the number of accounts. Give the regression equation and the results of the significance test for the regression coefficient.
(b) Do the same using market share to predict assets.
(c) Run a multiple regression using both the number of accounts and market share to predict assets. Give the multiple regression equation and the results of the significance test for the two regression coefficients.
(d) Compare the results of parts (a), (b), and (c).

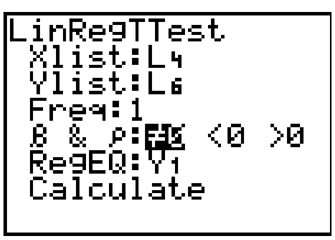
Solution. (a) We will enter the data into lists L4, L5, and L6. Then we define a scatterplot using Assets (in L6) on the *Y* axis and Accounts (in L5) on the *X* axis. The plot is very roughly linear; Fidelity and Charles Schwab clearly will influence the regression. We shall use LinRegTTest from the STAT TESTS menu. Then we execute the LinRegTTest on lists L5 and L6 to obtain a linear regression equation predicting assets using the number of accounts.

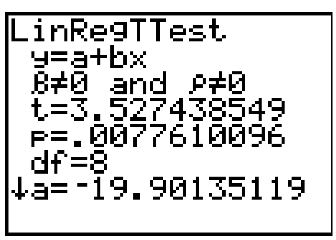


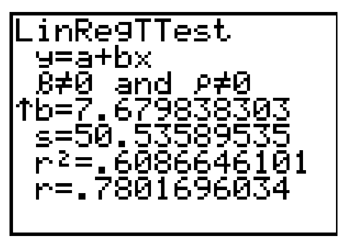



The *p*-value for a two-sided significance test is 4.2667×10^{-6}. If the regression coefficient β_1 were equal to 0, then there would be only a very small chance of obtaining a value of *b* as large as 0.0832, or a correlation as high as 0.96827, even with such a small sample. Thus, we have strong evidence to reject the hypothesis that $\beta_1 = 0$.

(b) Now we execute the LinRegTTest on lists L4 and L6 to obtain a linear regression equation $Assets = a + bx = -19.901 + 7.680 MarketShare$.

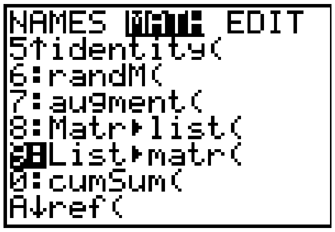


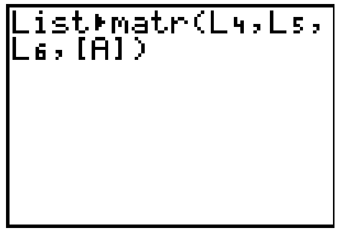


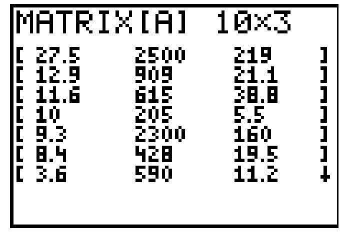


The *p*-value for the two-sided significance test is 0.007761. If the regression coefficient β_1 were equal to 0, then there would be less than a 1% chance of obtaining a value of *b* as high as 7.68 or a correlation as high as 0.78. Thus, we again can reject the hypothesis that $\beta_1 = 0$.

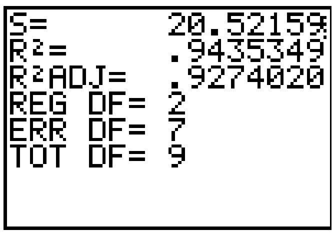
(c) To execute the **MULREG** program, we must first move the data to matrix [A]. To do so, we can use the List▸matr(command from the MATRX MATH menu. After bringing this command to the `Home` screen, enter the command `List |matr(L4,L5,L6,[A])`, where `[A]` is retrieved from the `MATRIX NAMES` menu.

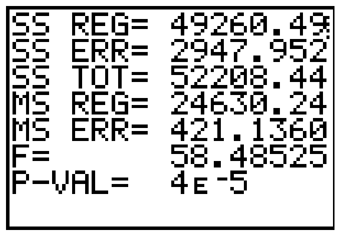


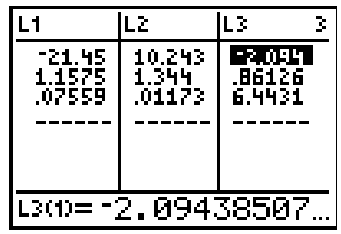




Last, execute the **MULREG** program.





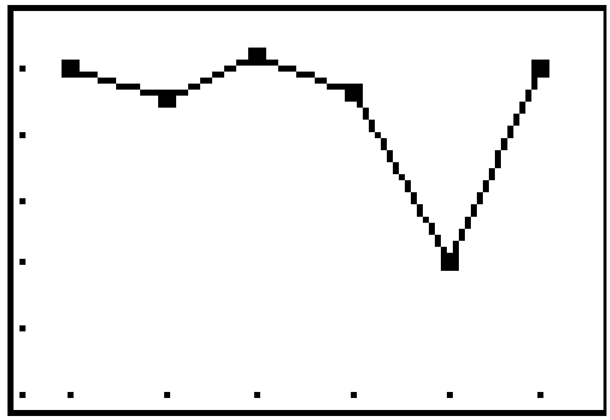


The regression equation is $\mu_{ASSETS} = -21.45 + 1.1575\,MarketShare + 0.0756\,Accts.$ The *p*-value of (approximately) 0 gives significant evidence to reject H$_0$: $\beta_1 = \beta_2 = 0$.

(d) Comparing the r^2 values from each part, we see that the multiple linear regression r^2 of 94.35% is not really different from the 93.8% obtained from the simple model using only the number of accounts. It seems reasonable to believe that market share (which has a very small *t* statistic) and number of accounts might be collinear (related to each other). The simple regression using only the number of accounts is most likely a “best” model.

CHAPTER

12



One-Way Analysis of Variance

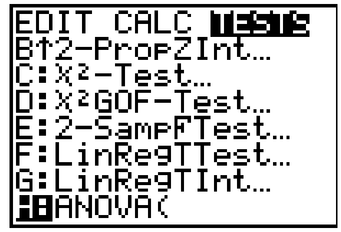
Introduction

Just as a chi-squared test can be seen as an extension of the two-proportion test, One-way Analysis of Variance (ANOVA) is an extension of the independent samples *t*-test to more than two groups.

In this chapter, we perform one-way analysis of variance (ANOVA) to test whether several normal populations, assumed to have the same variance, also have the same mean. The procedure is so similar on all models, that no additional explanation for the TI-89 is necessary.

12.1 Inference for One-Way Analysis of Variance

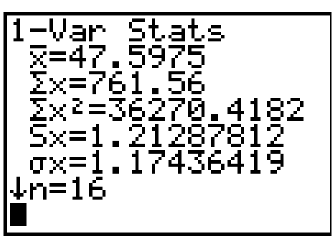
We begin with an exercise that demonstrates built-in analysis of variance capabilities using the `ANOVA(` command from the `STAT TESTS` menu.

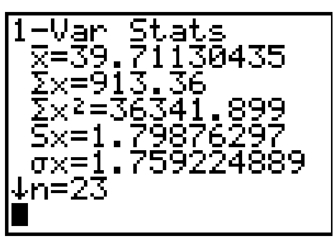


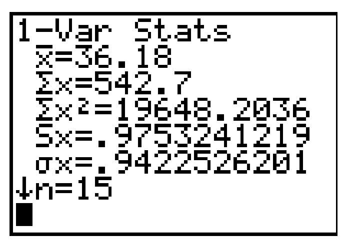
Example 12.1 Comparing Tropical Flowers. The data below give the lengths in millimeters of three varieties of the tropical flower *Heliconia*, which are fertilized by different species of hummingbird on the island of Dominica. Perform an ANOVA test to compare the mean lengths of the flowers for the three species.

H. bihai							
47.12	46.75	46.81	47.12	46.67	47.43	46.44	46.64
48.07	48.34	48.15	50.26	50.12	46.34	46.94	48.36
H. caribaea red							
41.90	42.01	41.93	43.09	41.47	41.69	39.78	40.57
39.63	42.18	40.66	37.87	39.16	37.40	38.20	38.07
38.10	37.97	38.79	38.23	38.87	37.78	38.01	
H. caribaea yellow							
36.78	37.02	36.52	36.11	36.03	35.45	38.13	37.1
35.17	36.82	36.66	35.68	36.03	34.57	34.63	

Solution. The `ANOVA(` command requires the data to be in lists. Here we shall enter the data into lists `L1`, `L2`, and `L3`. After doing so, we evaluate the basic statistics of each list.



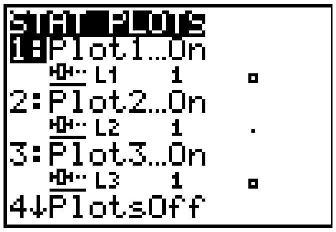


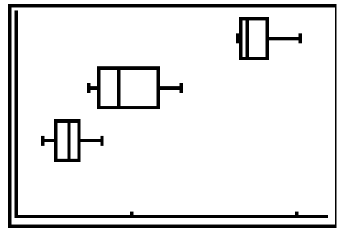


These sample means range from slightly larger than 36 to more than 47.5. Is the difference real (statistically significant) or just due to randomness?

Before proceeding to the ANOVA, we must consider the assumptions of the test. The first is that all groups have the same *population* standard deviation. In practice, the results will be approximately correct if the largest standard deviation is no more than twice as large as the smallest. Since $2*.975 = 1.95 > 1.799$, we are safe to continue based on this criterion.

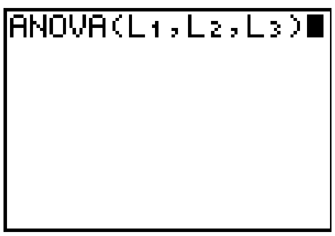
The second assumption is that the data come from Normal populations. We could do a Normal plot for each type of flower, but in this case side-by-side boxplots are a good idea. We're looking for indications of skewness or outliers. If you've forgotten how to make a boxplot, refer to Chapter 1 of this manual. Once all three plots are defined, press θ→ to display the graphs.

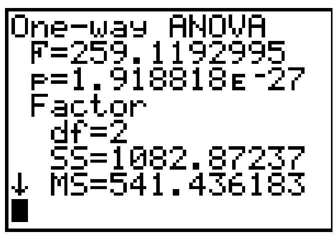


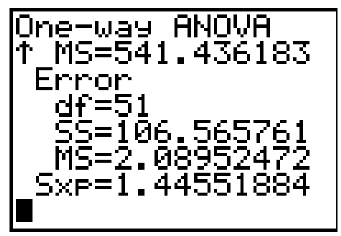


These data are not perfectly symmetric around the medians; also, there may be some skew in the distribution of *H. bihai*. However, with no outliers, we'll rely on the robustness of the ANOVA *F* test in continuing, as we have reasonable sample sizes.

Next, we test the hypothesis that the mean lengths of the three species are equal: $H_0 : \mu_1 = \mu_2 = \mu_3$. To do so, we bring the `ANOVA(` command to the `Home` screen and enter the command `ANOVA(L1,L2,L3)`.



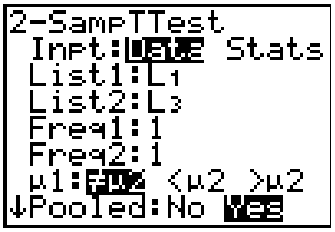


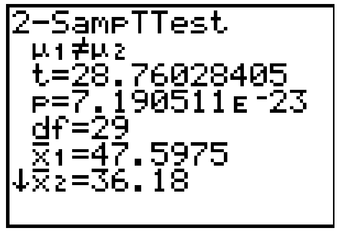


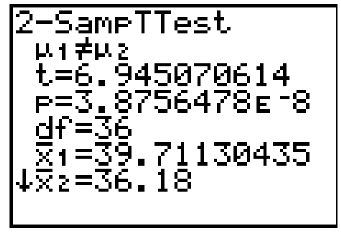
We receive a *p*-value of 1.92×10^{-27} from an *F*-statistic of 259.1193. The pooled standard deviation value is also displayed as `Sxp`=1.4455. If the true means were equal, then there would be almost no chance of the sample means varying by as much as they do with samples of these sizes. Thus, we have significant evidence to reject the claim that the mean lengths of these species are equal. We note that the r^2 value is not displayed, but it can be computed from the two displayed SS values. Here we can use r^2 = SSF/(SSF + SSE) = 1082.87237/(1082.87237 + 106.565761) = 0.9104.

Because we have rejected the hypothesis that the mean lengths are all equal, we can say that there is at least one pair of species that have different means. From the summary statistics, it appears that the species *H. bihai* and *H. caribaea yellow* have different mean lengths. But the sample means of *H. caribaea red* and *H. caribaea yellow* are close enough so that one might hypothesize that these species have the same mean length.

We can test any pair of species for equality of mean very quickly using `2-SampTTest`, but need to pay attention to *p*-values that might cause problems due to multiple comparisons. Since one of the assumptions of ANOVA is that the population standard deviations are the same we will answer `Yes` to `Pooled`. We demonstrate below with the pairs (`L1`, `L3`) and (`L2`, `L3`). Due to the extremely low *p*-values (7×10^{-23} for `L1` and `L3` and 4×10^{-8} for `L2` and `L3`), we see that we can reject both $\mu_1 = \mu_3$ and $\mu_2 = \mu_3$.







ANOVA for Summary Statistics

When raw data are given, then we can enter the data into lists and use the built-in `ANOVA(` command to test for equality of means. However, sometimes the summary statistics are given instead. In this case, we can use the **ANOVA1** program below to perform the analysis of variance on a TI-83/84 calculator; this is a built-in capability on a TI-89 — simply tell the calculator the data method is `Stats` and you will be prompted for the sample statistics instead of list names.

The ANOVA1 Program

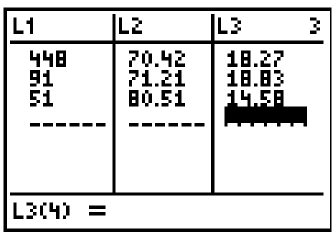
```
PROGRAM:ANOVA1
:1-Var Stats L2,L1
:sum(seq(L1(I)(L2(I)-⌋)²,I,1,dim(L1)))↵A
:sum(seq((L1(I)-1)(L3(I)²),I,1,dim(L1)))↵B
:dim(L/)↵I
:(A/(I-1))/(B/(n-I)) ↵F
:1-Fcdf(0,F,I-1,n-I) ↵P
:√(sum(seq((L/(J)-1)*L‹(J)²,J,1,dim(L1)))
 /(n-I)) ↵S
:ClrHome
:Output(2,1,"MEAN")
:Output(2,6,⌋)
:Output(3,1,"SP")
:Output(3,6,S)
:Output(4,1,"MSG")
:Output(4,6,A/(I-1))
:Output(5,1,"MSE")
:Output(5,6,B/(n-I))
:Output(6,1,"R^2")
:Output(6,6,A/(A+B))
:Output(7,1,"F")
:Output(7,6,F)
:Output(8,1,"P")
:Output(8,6,round(P,8))
```

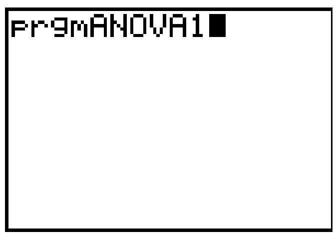
The program computes and displays the overall sample mean, the pooled sample deviation s_p, the mean square for groups MSG, the mean square for error MSE, the coefficient of determination, the *F*-statistic, and the *p*-value. Before executing the program, enter the sample sizes into list `L1`, the sample means into list `L2`, and the sample deviations into list `L3`.

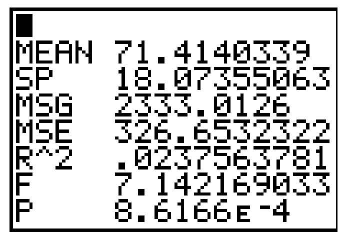
Example 12.2 Workplace Safety. The table below gives the summary data on safety climate index (SCI) as rated by workers during a study on workplace safety. Apply the ANOVA test and give the pooled deviation, the value of r^2, the F-statistic, and the p-value.

Job category	n	$\bar{x}$	s
Unskilled workers	448	70.42	18.27
Skilled workers	91	71.21	18.83
Supervisors	51	80.51	14.58

Solution. Enter the summary statistics into lists L1, L2, and L3, and execute the **ANOVA1** program. The desired statistics are then displayed. In particular, we obtain a pooled deviation of $s_p = 18.07355$ and a coefficient of determination of $r^2 = 0.023756$.







Due to the low p-value of 0.00086, we see that there is strong evidence to reject the claim that each group of workers has the same mean SCI. Looking at the given statistics, it is clear that supervisors rate the safety climate higher than the other two types of workers.

12.2 Comparing the Means

In this section, we provide a program for analyzing population contrasts based on summary statistics. Before executing the **CONTRAST** program, enter the sample sizes into list L1, the sample means into list L2, the sample deviations into list L3 (these are the same as for program **ANOVA1**), and the contrast equation coefficients into list L4. When prompted during the program, enter either **1** or **2** for a one-sided or two-sided alternative. The program then displays the p-value for a significance test and a confidence interval for mean contrast.

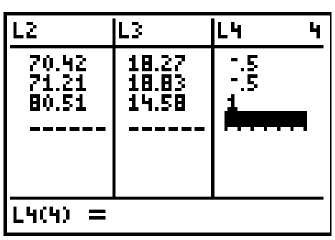
The CONTRAST Program

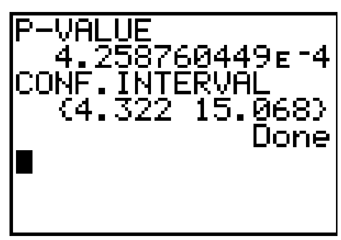
```
PROGRAM:CONTRAST
:Disp "ALTERNATIVE"
:Input A
:Disp "CONF. LEVEL"
:Input R
:1-Var Stats L2,L1
:√(sum(seq((L1(I)-1)*L3(I)²,
  I,1,dim(L1)))/(n-dim(L1)))↵S
:sum(seq(L2(I)L4(I),I,1,dim(L1)))↵C
:S√ (sum(seq(L4(I)²/L1(I),I,1,
  dim(L1))))↵E
:C/E↵T
:1-Var Stats L1
:Σx-dim(L1)↵D
:0.5-tcdf(0,abs(T),D) ↵P
:ClrHome
:Disp "P-VALUE"
:If A=1
:Then
:Disp P
:Else
:Disp 2*P
:End
:"tcdf(0,X,D)" ↵Y1
:solve(Y1-R/2,X,2) ↵Q
:Disp "CONF.INTERVAL"
:Disp {round(C-QE,3),
 round(C+QE,3)}
```

Example 12.3 More Workplace Safety. Using the data from Example 12.2 above, perform a two-sided significance test for the contrast between the average SCI of supervisors compared with the other two groups of workers. Also, find a 95% confidence interval for the contrast.

Solution. We use the contrast equation $\Psi = -.5\mu_{UN} - .5\mu_{SK} + 1\mu_{SU}$, and test the hypothesis H$_0$: $\mu_{SU} = .5\mu_{UN} + .5\mu_{SK}$ with the alternative H$_A$: $\mu_{SU} \neq .5\mu_{UN} + .5\mu_{SK}$.

If we still have the sample sizes, sample means, and sample deviations entered into lists L1, L2, and L3, then we just need to add the coefficients of Ψ into list L4. Then we execute the **CONTRAST** program by entering **2** for two-sided alternative and entering the desired confidence level.

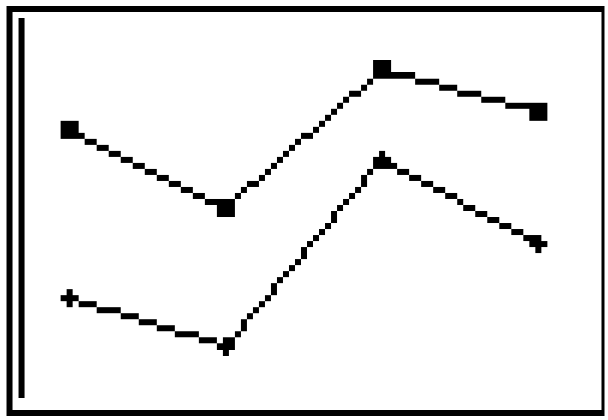




With the low *p*-value of 0.00043, we see that there is a significant contrast. According to the displayed confidence interval, the mean SCI rating by supervisors could be from 4.322 to 15.068 points higher than the average of the mean SCI ratings by unskilled workers and skilled workers.

CHAPTER

13



Two-Way Analysis of Variance

Introduction

Two-way analysis of variance is really best performed with a full statistical package on a computer. However, with a fairly lengthy program, TI-83 and -84 calculators can do a good job of this. TI-89 calculators easily generate the ANOVA table.

In this chapter, we provide a program that performs two-way analysis of variance to test for equality of means simultaneously among populations and traits in a two-factor experiment, and detail how a TI-89 might be used.

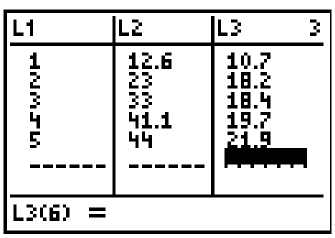
13.1 Plotting Means

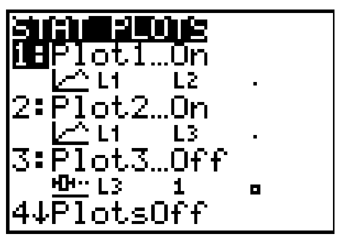
Example 13.1 Time Spent Eating. The table below gives the mean length of time (in minutes) that various groups of people spent eating lunch in various settings. Plot the group means for this example.

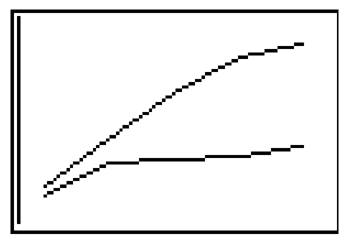
	Number of people eating				
Lunch setting	1	2	3	4	5
Workplace	12.6	23.0	33.0	41.1	44.0
Fast-food restaurant	10.7	18.2	18.4	19.7	21.9

Solution. We note that we cannot use ANOVA to analyze this data because each cell contains the *mean* of an unknown number of subjects, and we cannot assume that there are a fixed number of measurements per cell. In order to perform two-way ANOVA, we also would need to know the sample sizes and the variances of the data that produced the mean of each cell. However, we can plot the means with a time plot in order to observe a possible interaction.

First, enter the integers 1 through 5 into list `L1`, the five means for workplace diners into list `L2`, and the five means for the fast foods into list `L3`. Next set `Plot1` for a time plot of `L1` and `L2`, and set `Plot2` for a time plot of `L1` and `L3`.







We see that the patterns are not parallel, so it appears that we have an interaction.

13.2 Inference for Two-Way ANOVA

In this section, we give an **ANOVA2W** program that can be used on TI-83 and -84 calculators to perform two-way analysis of variance given the data. Two-way ANOVA for balanced designs (equal observations per cell) is a built-in function on the TI-89. This is a long program; we recommend it be downloaded from the text's website using a TI-Connect cable. Once your instructor (or a fellow student) has downloaded the program it is easily copied from one calculator to another (see Chapter 0 for more details on this).

The ANOVA2W Program

```
PROGRAM:ANOVA2W
Disp "DATA IN [A]"
Disp "","FACTOR A =
ROWS?","1=YES","2=NO",""
Repeat F≠0:getKey↵F:End
If F=92:[A]^T↵ [A]
Input "NO. B LEVELS=",B
dim([A]) ↵LθA: LθA (2) ↵A: LθA (1)/B↵N
{0}↵LθB
{B,A}↵dim([B])
For(I,1,B):For(J,1,A)
sum(seq([A]((I-1)N+K,J),K,1,N)) ↵ [B](I,J)
End:End
For(J,1,A)
sum(seq([B](I,J),I,1,B)) ↵LθA(J)
End
For(I,1,B)
sum(seq([B](I,J),J,1,A)) ↵LθB(I)
End
sum(LθA)^2/(ABN) ↵C
For(I,1,BN):For(J,1,A)
Ans+[A](I,J) ○
End:End
Ans-C↵T
If F=92:[A]^T↵ [A]
sum(LθA^2)/(BN)-C↵U
sum(LθB^2)/(AN)-C↵V
For(I,1,B):For(J,1,A)
Ans+[B](I,J) ^2
End:End
Ans/N-U-V-C↵W
If N>1
Menu("FIT INTERACTION?","YES",G,"NO",H)
Else
Goto H
End
Lbl G
T-U-V-W↵Z
ClrHome
Output(1,1,"SSA="
Output(1,9,U
Output(2,1,"SSB="
Output(2,9,V
Output(3,1,"SSAB="
Output(3,9,W
Output(4,1,"SSE="
Output(4,9,Z
Output(5,1,"SSTO="
Output(5,9,T
Output(6,1,"     "
Pause
ClrHome
```

```
Output(5,1,"       "
Pause
ClrHome
Output(1,1,"F A="
Output(1,9,U/Z
Output(2,1," p A="
1-Fcdf(0,U/Z,A,K) ↵P
round(P,6) ↵P
Output(2,9,P
Output(3,1,"F B="
Output(3,9,V/Z
Output(4,1," p B="
1-Fcdf(0,V/Z,B,K) ↵P
round(P,6) ↵P
Output(4,9,P
Output(5,1,"F AB="
Output(5,9,W/Z
Output(6,1," p AB="
1-Fcdf(0,W/Z,J,K) ↵P
round(P,6) ↵P
Output(6,9,P
Output(7,1,"      "
Pause
Goto R
Lbl H
ClrHome
T-U-V↵Z
Output(1,1,"ASSUME NO"
Output(2,1,"INTERACTION"
Output(3,1,"SSA="
Output(3,9,U
Output(4,1,"SSB="
Output(4,9,V
Output(5,1,"SSE="
Output(5,9,Z
Output(6,1,"SST="
Output(6,9,T
Output(7,1,"   "
Pause
ClrHome
A-1↵A
Output(1,1,"DF A="
Output(1,9,A
U/A↵U
B-1↵B
Output(2,1,"DF B="
Output(2,9,B
V/B↵V
(A+1)(B+1)N-A-B-1↵K
Output(3,1,"DF E="
Output(3,9,K
Z/K↵Z
```

```
A-1↲A
Output(1,1,"DF A="
Output(1,9,A
U/A↲U
B-1↲B
Output(2,1,"DF B="
Output(2,9,B
V/B↲V
(A)(B) ↲J
Output(3,1,"DF AB="
Output(3,9,J
W/J↲W
(A+1)(B+1)(N-1) ↲K
Output(4,1,"DF E="
Output(4,9,K
Z/K↲Z
Output(5,1,"DF T="
(A+1)(B+1)N-1↲G
Output(5,9,G
Pause
ClrHome
Output(1,1,"MSA="
Output(1,9,U
Output(2,1,"MSB="
Output(2,9,V
Output(3,1,"MSAB="
Output(3,9,W
Output(4,1,"MSE="
Output(4,9,Z
```

```
Output(4,1,"DF T="
Output(4,9,(A+1)(B+1)N-1
Output(5,1,"MSA="
Output(5,9,U
Output(6,1,"MSB="
Output(6,9,V
Output(7,1,"MSE="
round(Z,6) ↲Z
Output(7,9,Z
Pause
ClrHome
Output(1,1,"F A="
Output(1,9,U/Z
Output(2,1," p A="
1-Fcdf(0,U/Z,A,K) ↲P
round(P,6) ↲P
Output(2,9,P
Output(3,1,"F B="
Output(3,9,V/Z
Output(4,1," p B="
1-Fcdf(0,V/Z,B,K) ↲P
round(P,6) ↲P
Output(4,9,P
Pause
Lbl R
ClrHome
1/N[B] ↲ [B]
Output(1,1,"MEANS IN MATRX B"
Output(2,1,"COLS FACTOR A"
Pause
ClrHome
Stop
```

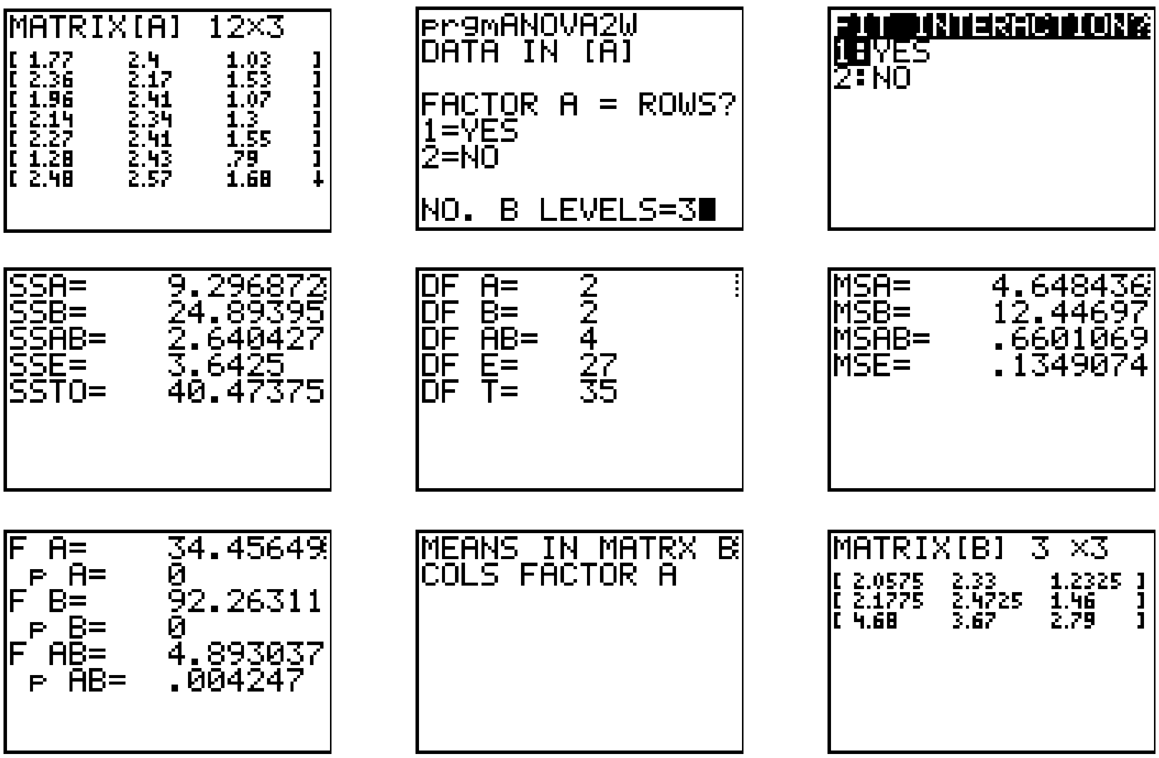
To execute the program, enter the data into matrix `[A]`. The program displays the *F*-statistics and *p*-values for the two factors, and for the interaction when there is more than one observation per cell (or interaction was not specified in the model).

Example 13.2 Iron in Food. Does the type of cooking pot affect the iron content in food, and does the type of food cooked matter? In many parts of the world where people suffer from anemia, this could be important information. The table below gives the amount of iron in certain foods, measured in milligrams of iron per 100 grams of cooked food, after samples of each food were cooked in each type of pot.

IRON		**Food**	
Type of pot	Meat	Legumes	Vegetables
Aluminum	1.77 2.36 1.96 2.14	2.40 2.17 2.41 2.34	1.03 1.53 1.07 1.30
Clay	2.27 1.28 2.48 2.68	2.41 2.43 2.57 2.48	1.55 0.79 1.68 1.82
Iron	5.27 5.17 4.06 4.22	3.69 3.43 3.84 3.72	2.45 2.99 2.80 2.92

Perform two-way ANOVA on the data regarding the main effects and interaction, then plot the means and examine the plot.

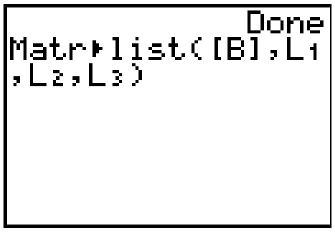
TI-83/84 Solution. Enter the data into matrix [A]. (Hint: it may sometimes be easier to enter data into lists and then use `List▸Matrix` from the `List Math` menu to create matrix [A].) It makes no difference what is considered factor A or B; here the data have been entered so that all meat observations are in the first column, legumes in the second, and vegetables in the third. All the aluminum pot observations come first, then clay pots, and finally iron pots. You are asked if factor A is in the rows. Since here we consider the type of food factor A, press ♥ for `No` (otherwise the program will look for multiples of this across the rows). If you had entered the data just as it was in the table above (a 3 × 12 matrix) the answer here would be No. You are then asked for the number of levels of Factor B. Since there were three types of pots, there are 3 levels. Type 3 followed by ÷. The next question is self-explanatory. One should always fit the interaction term unless prior examination shows this is not significant.

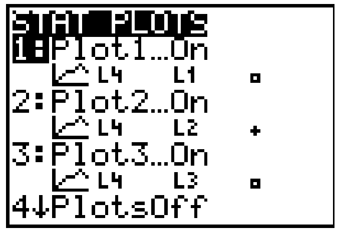


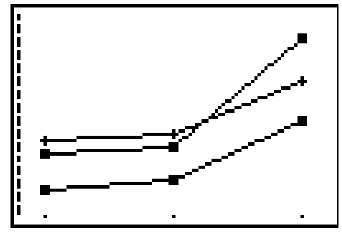
The program then displays the entries in the ANOVA table and pauses after each display. Press ÷ to proceed to the next. The program next reminds you that "cell" means are stored in matrix [B] and the columns of this matrix represent the levels of factor A (food here).

We obtain very low *p*-values for both factors and for the interaction. Therefore, we can conclude that, even with these small samples, there is a significant difference in the mean due to type of pot as well as type of food. Also, there is a statistically significant interaction.

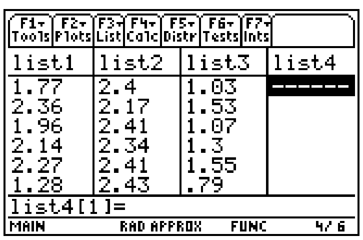
To create the means plot, use `Matr▸List` from the `Matrix Math` menu to create lists to plot. You need to specify as many lists as there are columns in [B]. Next, input a column containing values 1, 2, and 3 in `L4`. Define connected plots (time plots) using L4 as the `Xlist` and each list of means as the `Ylist`; in this example you will have three plots. Pressing θ→ displays the plots. Here, we can clearly see the interaction since the lines cross. Vegetables have consistently the lowest amount of iron but cooking them in an iron pot increases the amount. Meat cooked in an iron pot has the largest amount of iron due to the interaction.

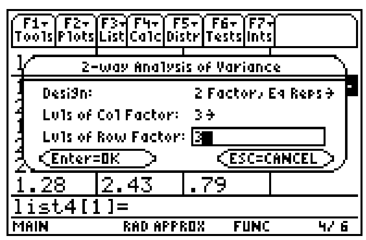


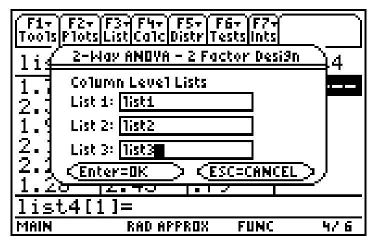


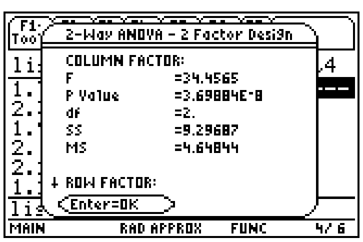


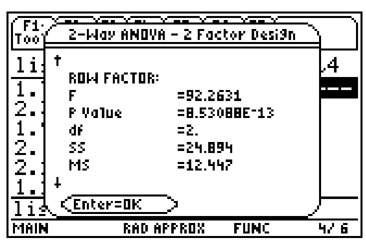
TI-89 Solution. Enter the data. Here, all meat observations are in `list1`, all legumes in `list2,` and all vegetables in `list3`. Select option `D:ANOVA2-Way` from the `Tests` menu. Use the right arrow key to expand the menu options on the first screen. We have a 2 Factor Equal Replications set of data with three levels of the column factor (the three lists) and three levels of the row factor. Press ÷ to continue. Next, indicate which lists contain your data. The calculator displays the entries of the ANOVA table.

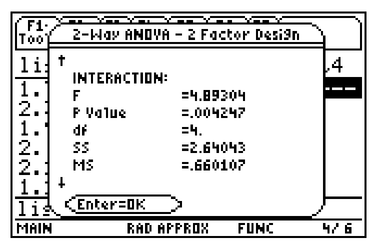


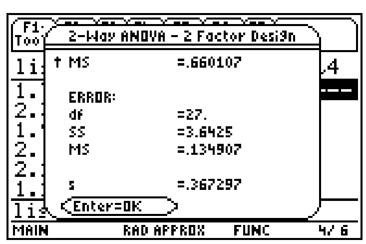




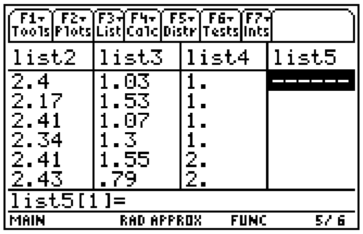


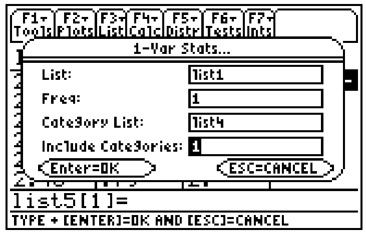


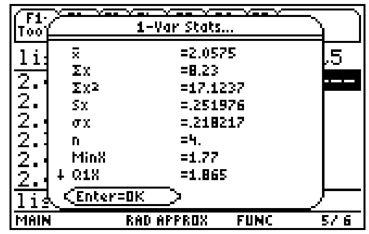




Unfortunately, the TI-89 does not create the cell means for you to create the means plot. You can calculate these fairly easily, though. Enter four 1s in list 4, then four 2s, then four 3 s. These will correspond to the type of pot (1 = Aluminum, 2 = Clay, 3 = Iron). Now, from the `Calc` menu, use `1-Var Stats` and the Category list to get statistics for each category for each list. Below, we see the mean of meat cooked in aluminum pots is 2.0575.

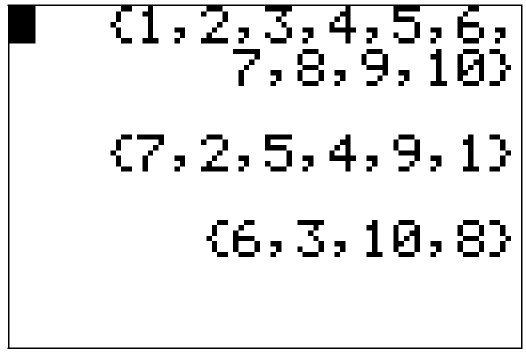






CHAPTER

14



Bootstrap Methods and Permutation Tests

Introduction

In this chapter, we demonstrate the bootstrap methods for estimating population parameters. Throughout, we use TI-83 Plus programs to perform the necessary resampling procedures.

To truly perform these resampling methods correctly requires more processing power than a TI calculator possesses (we'd want several hundred resamples). The methods given in this chapter will approximate a resampling distribution only. Some of these programs will run for several minutes, depending on the size of the original sample and the number of resamples.

Bear in mind that since these methods are based on random resamples, your output will *not* duplicate ours.

14.1 The Bootstrap Idea

We first provide a TI calculator program that performs resampling on a previously obtained random sample. Before executing the **BOOT** program, enter the random sample into list `L1`. Then bring up the program and enter the desired number of resamples. If you also want a confidence interval for the statistic, enter **1** for `CONF. INTERVAL?`; otherwise, enter **0**. The program takes resamples from the entered random sample, enters their means into list `L2`, and displays the mean and bootstrap standard error of the resamples.

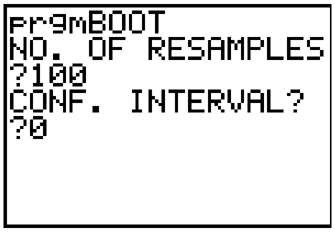
The BOOT Program

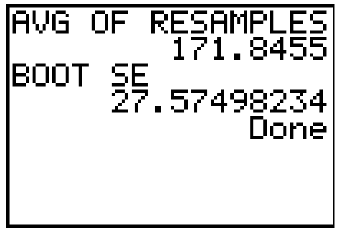
```
Program:BOOT
:dim(L1)↵N
:Disp "NO. OF RESAMPLES"
:Input B
:Disp "CONF. INTERVAL?"
:Input Y
:If Y=1
:Then
:Disp "CONF. LEVEL"
:Input R
:"tcdf(0,X,N-1)" ↵Y1
:solve(Y1-R/2,X,2) ↵Q
:1-Var Stats L1
:⌡↵X
:End
:ClrList L2
:For(I,1,B)
:randInt(1,N,N) ↵L3
:sum(seq(L1(L3(J)),J,1,N))/N↵L2(I)
:End
:ClrList L3
:1-Var Stats L2
:ClrHome
:Disp "AVG OF RESAMPLES"
:Disp ⌡
:Disp "BOOT SE"
:Disp Sx
:If Y=1
:Then
:Output(6,2,"CONF. INTERVAL")
:Output(7,2,X-Q*Sx)
:Output(8,2,X+Q*Sx)
:End
```

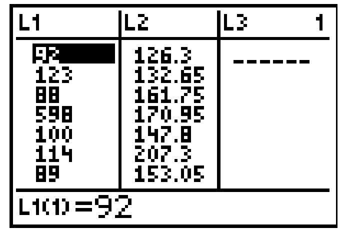
Example 14.1 Guinea Pig Survival. Here is an SRS of 20 guinea pig survival times (in days) after they were injected with *tubercle bacillus* during a medical trial. Create and inspect the bootstrap distribution of the sample mean for these data.

92	123	88	598	100	114	89	522	58	191
137	100	403	144	184	102	83	126	53	79

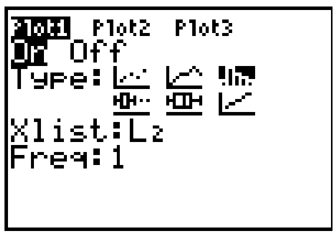
Solution. We will use the **BOOT** program with 100 resamples to create a bootstrap distribution. But first, we must enter the data into list `L1`. In this running of the program, we will not find a confidence interval, so we enter **0** for `CONF. INTERVAL?` when prompted. After taking a few minutes to execute, the program stores the 100 means from the resamples in list `L2`.

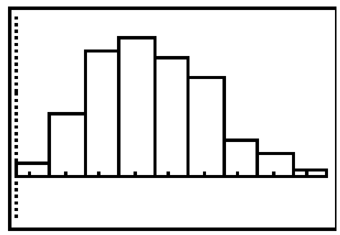






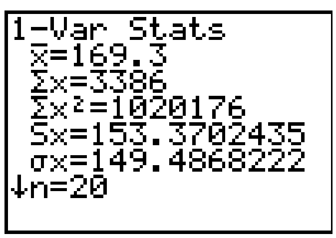
To assess the normality of the distribution of the resample means, we shall make a histogram of the data in list L2. We adjust the STAT PLOT settings and press θ→ to display the graph; if desired, the default settings can be adjusted using WINDOW. We see that the bootstrap distribution appears close to normal.

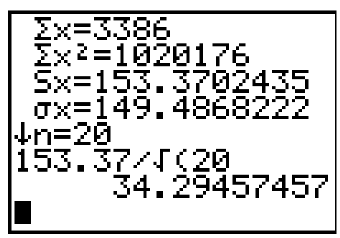




Example 14.2 Bootstrap Standard Error and the Original. Compute the standard error $s/\sqrt{n}$ for the 20 survival times in Example 14.1 and compare it with the bootstrap standard error from the resampling done in that exercise.

Solution. With the data entered into list L1, we can use the command 1-Var Stats L1 to compute the sample deviation *s*. We then divide Sx by $\sqrt{20}$ to compute the standard error.





We see that the standard error of 34.294629 is somewhat higher than the bootstrap standard error of 27.575 obtained from our 100 resamples in Example 14.1. This is due to the skew and outliers in the original distribution.

14.2 First Steps in Using the Bootstrap

In this section, we demonstrate how to compute a bootstrap *t* confidence interval for a population parameter. We begin though with a quick estimation of the bias obtained using the preceding exercises.

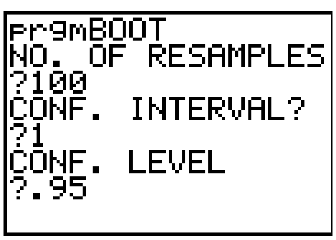
Example 14.3 More Guinea Pig Survival. Let's return to the 20 guinea pig survival times from Example 14.1.
(a) What is the bootstrap estimate of the bias?
(b) Give the 95% bootstrap *t* confidence interval for μ.
(c) Give the usual 95% one-sample *t* confidence interval.

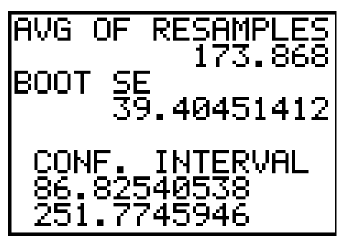
Solution. (a) In Example 14.2, we found the sample mean of the data, along with the standard deviation, using the `1-Var Stats` command. From the output display, we see that $\overline{x}$ = 169.3. The average of all 100 resample means obtained in Example 14.1 was 171.846. Thus, the bootstrap estimate of bias is 171.846 – 169.3 = 2.546.

(b) We now apply the formula $\overline{x} \pm t^* SE_{BOOT}$, where $\overline{x}$ is the mean of the *original* sample, the critical value *t** is from the $t(n-1)$ distribution and the sample size is *n* = 20. To find this critical value from the *t*(19) distribution, we can use the **TSCORE** program (page 82), Table C, or `InvT` on a TI-84 or -89. Doing so, we obtain *t** = 2.093.

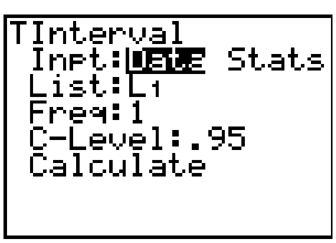
The 95% bootstrap *t* confidence interval for mean guinea pig survival (in days) after being injected with *tubercle bacillus* becomes $169.3 \pm 2.093 * 27.575$, or (111.586, 227.014).

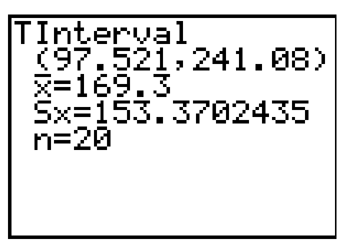
Note: This confidence interval would have been computed and displayed when executing the **BOOT** program in Example 14.1 had we entered **1** for `CONF. INTERVAL?`. A different bootstrap confidence interval is displayed below from another execution of the **BOOT** program. Remember, since this is a different random resample; we do not expect the two to agree.





(c) To compute the standard one-sample *t* confidence interval, we use the `TInterval` from the `STAT TESTS` menu. After adjusting the `List` to `L1`, we calculate a 95% confidence interval of (97.521, 241.08).





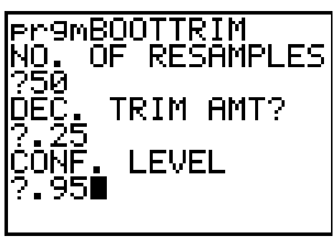
Confidence Interval for a Trimmed Mean

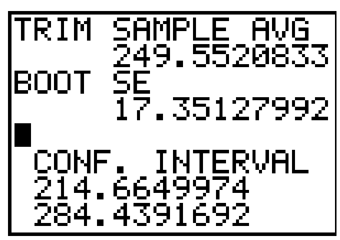
Next, we demonstrate a program that will compute a bootstrap *t* confidence interval for a trimmed mean (where x% of data at the top and bottom are deleted or trimmed, and a mean calculated). Before executing the **BOOTTRIM** program given on the next page, enter the random sample into list `L1`. Then enter the desired number of resamples, the decimal amount to be trimmed at each end, and the desired confidence level when prompted. The program takes resamples from the entered random sample and enters their trimmed mean into list `L2`. Then the trimmed mean of the original sample, the trimmed bootstrap standard error, and the confidence interval are displayed.

Example 14.4 Seattle Real Estate. The table below gives an SRS of 50 real estate sale prices in Seattle (in thousands of dollars) during 2002. Use the bootstrap *t* method to give a 95% confidence interval for the 25% trimmed mean sale price.

142	175	197.5	149.4	705	232	50	146.5	155	1850
132.5	215	116.7	244.9	290	200	260	449.9	66.407	164.95
362	307	266	166	375	244.95	210.95	265	296	335
335	1370	256	148.5	987.5	324.5	215.5	684.5	270	330
222	179.8	257	252.95	149.95	225	217	570	507	190

Solution. After entering the data into list `L1`, we simply execute the **BOOTTRIM** program. Here we used 50 resamples. The trimmed mean of the original sample is given as 249.552, and the 25% trimmed bootstrap confidence interval for the mean price of a house in Seattle in 2002 is 214.665 to 284.439 thousand dollars.





Note: Due to programming differences, the trimmed mean given by the **BOOTTRIM** program may not agree with the value given by software. Because the sample size of 50 is not a multiple of 4, we cannot trim precisely 25% of the measurements from the high and low ends. In the case of Example 14.4 above, the **BOOTTRIM** program trimmed the lowest 12 and the highest 13 measurements before computing the trimmed sample mean.

The BOOTTRIM Program

```
Program:BOOTTRIM
:dim(L1)↲N
:Disp "NO. OF RESAMPLES"
:Input B
:Disp "DEC. TRIM AMT?
:Input M
:Disp "CONF. LEVEL"
:Input R
:"tcdf(0,X,N-1)" ↲Y1
:solve(Y1-R/2,X,2) ↲Q
:int(M*N)+1↲L
:int((1-M)*N) ↲U
:SortA(L1)
:sum(seq(L1(I),I,L,U))/(U+1-L) ↲X
:ClrList L2,L4
:For(I,1,B)
:randInt(1,N,N) ↲L3
:For(J,1,N)
:L1(L3(J)) ↲L4(J)
:End
:SortA(L4)
:sum(seq(L4(I),I,L,U))/(U+1-L) ↲L2(I)
:End
:ClrList L3,L4
:1-Var Stats L2
:ClrHome
:Disp "TRIM SAMPLE AVG"
:Disp X
:Disp "BOOT SE"
:Disp Sx
:Output(6,2,"CONF. INTERVAL")
:Output(7,2,X-Q*Sx)
:Output(8,2,X+Q*Sx)
:Stop
```

Difference in Means

We conclude this section with a program that computes a bootstrap *t* confidence interval for the difference in means. Before executing the **BOOTPAIR** program, enter a random sample from the first population into list `L1` and enter a random sample from the second population into list `L2`. Then enter the desired number of resamples and the confidence level. The resampled differences in mean are stored in list `L3`. The difference of the original sample averages is displayed along with the bootstrap standard error and the confidence interval.

The BOOTPAIR Program

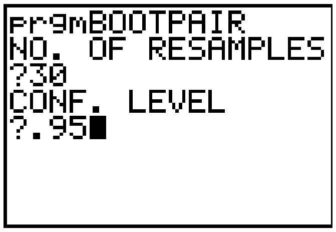
```
Program:BOOTPAIR
:dim(L1)↵N
:dim(L2) ↵M
:Disp "NO. OF RESAMPLES"
:Input B
:Disp "CONF. LEVEL"
:Input R
:1-Var Stats L1
:⌡ ↵X
:1-Var Stats L2
:⌡ ↵Y
:ClrList L3,L4
:For(I,1,B)
:randInt(1,N,N) ↵L5
:randInt(1,M,M) ↵L6
:sum(seq(L1(L5(J)),J,1,N))/N↵L3(I)
:sum(seq(L2(L6(J)),J,1,M))/M↵L4(I)
:End
:1-Var Stats L3
:Sx↵S
:1-Var Stats L4
:Sx↵T
:int((S²/N+T²/M) ²/((S²/N) ²/
 (N-1)+(T²M) ²/(M-1))) ↵P
:√(S²/N+T²/M) ↵D
:"tcdf(0,X,P)" ↵Y1
:solve(Y1-R/2,X,2) ↵B
:B*D↵E
:ClrHome
:Disp "DIFF OF AVGS"
:Disp X-Y
:Disp "BOOT SE"
:Disp D
:Output(6,2,"CONF. INTERVAL")
:Output(7,2,X-Y-E)
:Output(8,2,X-Y+E)
```

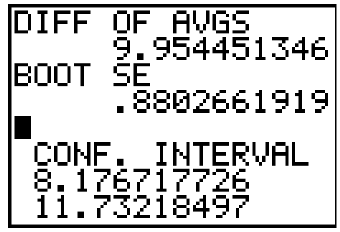
Example 14.5 Reading Scores. Following are the scores on a test of reading ability for two groups of third-grade students. The Treatment group used a set of "directed reading activitites" and the Control group followed the same curriculum without the activities.

Treatment group				**Control group**			
24	61	59	46	42	33	46	37
43	44	52	43	43	41	10	42
58	67	62	57	55	19	17	55
71	49	54		26	54	60	28
43	53	57		62	20	53	48
49	56	33		37	85	42	

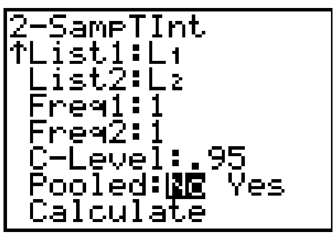
(a) Bootstrap the difference in means $\bar{x}_1 - \bar{x}_2$ and give the bootstrap standard error and a 95% bootstrap *t* confidence interval.
(b) Compare the bootstrap results with a two-sample *t* confidence interval.

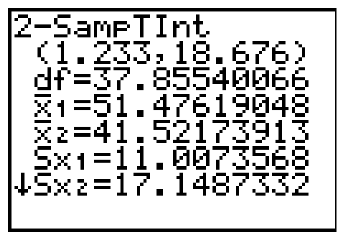
Solution. (a) We first enter the treatment scores into list `L1` and the control scores into list `L2`. We shall use 30 resamples to bootstrap the difference in means and to obtain a 95% confidence interval. Upon executing the **BOOTPAIR** program, we obtain a bootstrap standard error of about 0.880 and a 95% confidence interval for the difference in mean reading scores for the two teaching methods of (8.177, 11.732). Since the interval does not contain 0, we have evidence that the directed reading activities help students reading ability.





(b) To find a traditional two-sample *t* confidence interval, we can use `2-SampTInt` from the `STAT TESTS` menu. Because we are not assuming normal populations with the same variance, we do not pool the sample variances. Upon calculating, we obtain a 95% confidence interval of (1.233, 18.676). This interval is much wider than the bootstrap interval due to the much larger standard error obtained when using the original sample deviations s_x and s_y as opposed to the sample deviations from the collection of resample means.



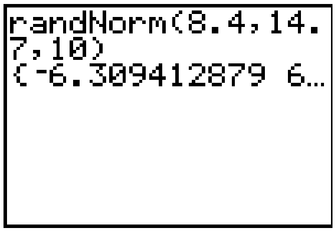


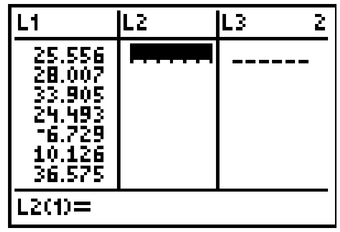
14.3 How Accurate Is a Bootstrap Distribution?

Below we work an exercise to demonstrate how the bootstrap distribution varies with the sample size.

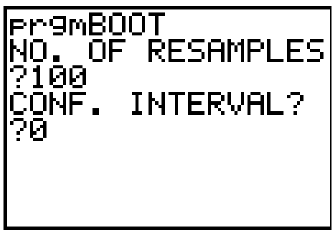
Example 14.6 Bootstrapping a Normal Distribution. Draw an SRS of size 10 from a *N*(8.4, 14.7) population. What is the exact distribution of the sampling mean $\bar{x}$ for this sample size? Bootstrap the sample mean using 100 resamples. Give a histogram of the bootstrap distribution and the bootstrap standard error. Repeat the process for a sample of size 40.

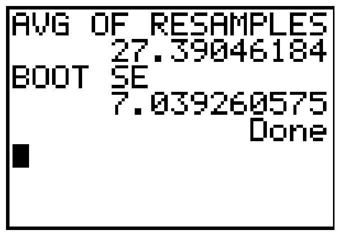
Solution. For an SRS of size $n = 10$ from a *N*(8.4, 14.7) population, $\bar{x} \sim N(8.4, 14.7/\sqrt{10}) = N(8.4, 4.6845)$. To draw such an SRS, we will use the `randNorm(` command from the `MATH PRB` menu. Enter the command `randNorm(8.4,14.7,10)↵L1` to store the SRS in list `L1`.

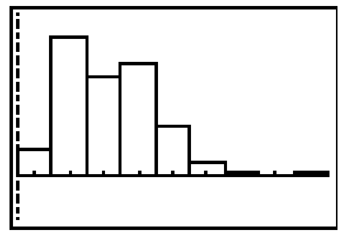




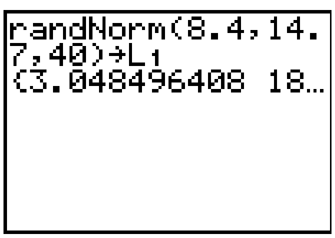
With the sample in list `L1`, we can execute the **BOOT** program to bootstrap the sample mean. In this case, we enter **0** for `CONF. INTERVAL?` because we are only interested here in the bootstrap distribution. Due to the small sample size, we obtain statistics of this bootstrap distribution that differ noticeably from those of the $N(8.4, 4.6845)$ distribution. A histogram of the resample means stored in list `L2` is shown; this distribution does not look Normally distributed.

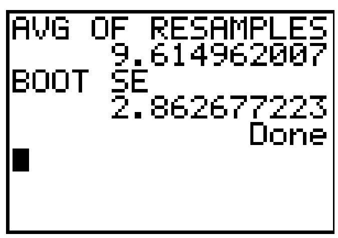


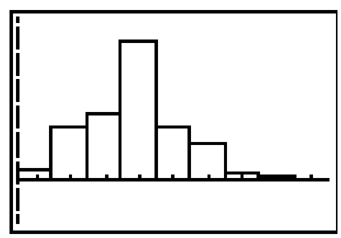




(c) We now repeat the process for an SRS of size 40. With this sample size, the bootstrap distribution statistics become much closer to the actual $N(8.4, 14.7/\sqrt{40}) = N(8.4, 2.324)$ distribution of $\bar{x}$.





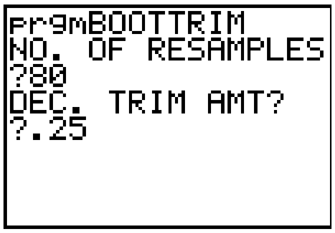


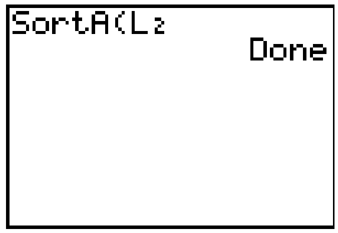
14.4 Bootstrap Confidence Intervals

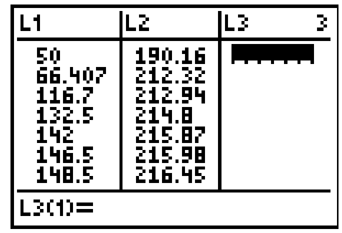
In this section, we demonstrate how to find a confidence interval based solely on the bootstrap distribution. We also provide another program that will bootstrap the correlation coefficient or the regression slope for two related variables.

Example 14.7 More Seattle Real Estate. The Seattle house prices in Example 14.4 are strongly right-skewed with outliers. In Example 14.4, we calculated a bootstrap interval for a 25% trimmed mean, using a *t* critical value. How does this interval compare to one based on the bootstrap distribution?

Solution. We run the **BOOTTRIM** program using 80 resamples since the endpoints of our interval will be the 2.5th and 97.5th percentiles of the bootstrap distribution. Since .025*80 = 2, the endpoints will be the 3rd and 78th observations in a sorted bootstrap distribution. From the sorted list, we see the lower end of the bootstrap confidence interval will be 212.94. Moving to the upper end, we find the 78th observation in this distribution is 274.95. Our bootstrap confidence interval for the mean 2002 Seattle house price is then $212,940 to $274,950. This is slightly narrower than the interval found in Example 14.4, which was $214,665 to $284,439.







The BOOTCORR Program

```
Program:BOOTCORR
:dim(L1)↵N
:ClrList L3,L4,L5
:Disp "1=CORRELATION","2=REG. SLOPE"
:Input W
:Disp "NO. OF RESAMPLES"
:Input B
:Disp "CONF. LEVEL"
:Input R
:For(I,1,B)
:randInt(1,N,N) ↵L6
:For(J,1,N)
:L1(L6(J)) ↵L4(J)
:L2(L6(J)) ↵L5(J)
:End
:LinReg(ax+b) L4,L5
:If W=1
:Then
:r↵L3(I)
:Else
:a↵L3(I)
:End
:End
:LinReg(ax+b) L1,L2
:If W=1
:Then
:r↵X
:Else
:a↵X
:End
:ClrList L4,L5,L6
:1-Var Stats L3
:"tcdf(0,X,N-1)" ↵Y1
:solve(Y1-R/2,X,2) ↵Q
:ClrHome
:If W=1
:Then
:Disp "SAMPLE CORR."
:Else
:Disp "REG. SLOPE"
:End
:Disp X
:Disp "BOOT SE"
:Disp Sx
:Output(6,2,"CONF. INTERVAL")
:If W=1
:Then
:Output(7,2,max(–1,X-Q*Sx))
:Output(8,2,min(X+Q*Sx,1))
:Else
:Output(7,2,X-Q*Sx)
:Output(8,2,X+Q*Sx)
:End
```

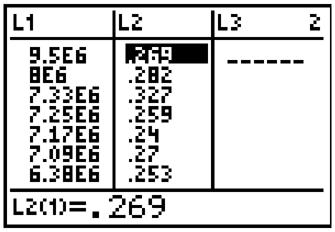
The **BOOTCORR** program can be used to bootstrap the correlation coefficient or the regression slope for paired sample data. Before executing the program, enter the random sample data into lists `L1` and `L2`. When prompted, enter **1** to bootstrap the correlation coefficient or enter **2** to bootstrap the regression slope. Then enter the desired number of resamples and the confidence level. The resampled statistics are stored in list `L3`, and the statistic of the original sample data is displayed along with the bootstrap standard error and the confidence interval.

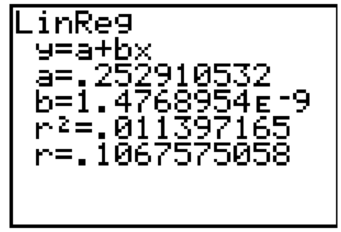
Example 14.8 Baseball Salaries and Batting Averages. The following table gives the 2002 salaries and career batting averages for 50 randomly selected MLB players (excluding pitchers).

Player	Salary	Avg	Player	Salary	Avg	Player	Salary	Avg
1.	$9,500,000	0.269	18.	$3,450,000	0.242	35.	$630,000	0.324
2.	$8,000,000	0.282	19.	$3,150,000	0.273	36.	$600,000	0.200
3.	$7,333,333	0.327	20.	$3,000,000	0.250	37.	$500,000	0.214
4.	$7,250,000	0.259	21.	$2,500,000	0.208	38.	$325,000	0.262
5.	$7,166,667	0.240	22.	$2,400,000	0.306	39.	$320,000	0.207
6.	$7,086,668	0.270	23.	$2,250,000	0.235	40.	$305,000	0.233
7.	$6,375,000	0.253	24.	$2,125,000	0.277	41.	$285,000	0.259
8.	$6,250,000	0.238	25.	$2,100,000	0.227	42.	$232,500	0.250
9.	$6,200,000	0.300	26.	$1,800,000	0.307	43.	$227,500	0.278
10.	$6,000,000	0.247	27.	$1,500,000	0.276	44.	$221,000	0.237
11.	$5,825,000	0.213	28.	$1,087,500	0.216	45.	$220,650	0.235
12.	$5,625,000	0.238	29.	$1,000,000	0.289	46.	$220,000	0.243
13.	$5,000,000	0.245	30.	$950,000	0.237	47.	$217,500	0.297
14.	$4,900,000	0.276	31.	$800,000	0.202	48.	$202,000	0.333
15.	$4,500,000	0.268	32.	$750,000	0.344	49.	$202,000	0.301
16.	$4,000,000	0.221	33.	$720,000	0.185	50.	$200,000	0.224
17.	$3,625,000	0.301	34.	$675,000	0.234			

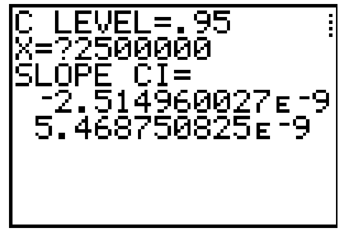
(a) Calculate the sample correlation between salary and average.
(b) Bootstrap the correlation and give a 95% confidence interval for the correlation.
(c) Calculate the least-squares regression line to predict batting average from salary. Give the traditional 95% *t* confidence interval for the slope of the regression line.
(d) Bootstrap the regression model. Give a 95% bootstrap *t* confidence interval and a 95% bootstrap percentile confidence interval for the regression slope.

Solution. (a) and (c) We can find the sample correlation, least-squares regression line, and traditional confidence interval for the slope using techniques from Chapter 10. First, we enter the data into lists, say lists `L1` and `L2`. Next, we use `LinReg(a+bx)` (item 8 from the `STAT CALC` menu) and enter the command `LinReg(a+bx) L1,L2`. We obtain a sample correlation of r = 0.1067575. The regression line is given as $BatAvg = .253 + 1.477 * 10^{-9}\,Salary$. Thus, we see that the slope appears to be 0.

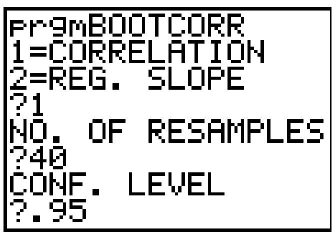


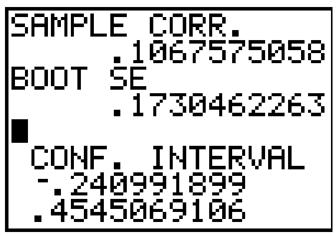


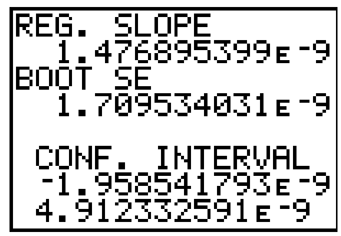
We can find a confidence interval for the slope using the **REGINF** program. After running the program, we find a 95% t confidence interval for the slope of the regression line to be $(2.515\times10^{-9}, 5.469\times10^{-9})$.



(b) and (d) We now execute the **BOOTCORR** program twice using 40 resamples each time. In the first running, we enter **1** to bootstrap the correlation. In the second running, we enter **2** to bootstrap the regression slope.

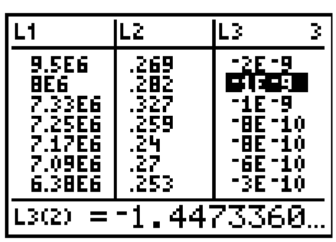


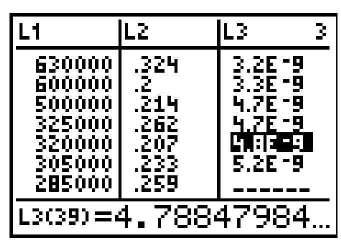




In the first running, we obtain a 95% bootstrap interval for the correlation of (–0.241, 0.456). In the second running, we obtain a 95% bootstrap interval for the regression slope of $(-1.956\times10^{-9}, 4.912\times10^{-9})$.

The 40 resample regression slopes are stored in list `L3` after option 2 of the **BOOTCORR** program ends. So next, we enter the command `SortA(L3)` to resort the list in increasing order. Because .025*40 = 1, the endpoints of the 95% percentile confidence interval are the 2nd and 39th elements of the sorted list. The interval here is $(-1.447\times10^{-9}, 4.788\times10^{-9})$.





14.5 Significance Testing Using Permutation Tests

We conclude this chapter with two programs that will simulate some of the bootstrap permutation tests. The **BOOTTEST** program performs a permutation test for the difference in means, and the **BTPRTEST** program performs a permutation test for either the difference in paired means or for the correlation.

Before executing the **BOOTTEST** program, enter data from the first population into list `L1` and enter data from the second population into list `L2`. When prompted, enter **1**, **2**, or **3** to designate the desired alternative $\mu_1 < \mu_2$, $\mu_1 > \mu_2$, or $\mu_1 \neq \mu_2$. The resampled differences in permuted means are ordered and then stored in list `L3`. The program displays the difference between the original sample means $\bar{x}_1 - \bar{x}_2$ and the p-value.

The BOOTTEST Program

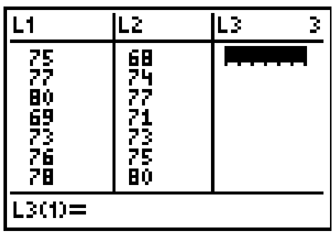
```
Program:BOOTTEST
:Disp "1 = ALT. <","2 = ALT. >","3 = ALT. _"
:Input C
:Disp "NO. OF RESAMPLES"
:Input B
:dim(L1)↵N
:dim(L2) ↵M
:augment(L1,L2)↵L1
:ClrList L3
:For(S,1,B)
:ClrList L6
:seq(J,J,1,N+M) ↵L4
:For(I,1,M)
:ClrList L5
:randInt(1,N+M-I+1) ↵A
:L4(A) ↵L6(I)
:1↵K
:While K<A
:L4(K) ↵L5(K)
:1+K↵K
:End
:A↵K
:While K≤(N+M-I)
:L4(K+1) ↵L5(K)
:1+K↵K
:End
:L5↵L4
:End
:sum(seq(L1(L5(J)),J,1,N))/N↵X
:sum(seq(L1(L6(J)),J,1,M))/M↵Y
:X-Y↵L3(S)
:End
:ClrList L4
:seq(L1(I),I,1,N) ↵L1
:mean(L1)↵X
:mean(L2)↵Y
:SortA(L3)
:X-Y+1↵L3(B+1)
:0↵K
:While L3(K+1) <(X-Y)
:K+1↵K
:End
:K/B↵P
:While L3(K+1) ≤ (X-Y)
:K+1↵K
:End
:(B-K)/B↵Q
:seq(L3(I),I,1,B) ↵L3
:ClrHome
:Disp "DIFF IN MEANS"
:Disp X-Y
:Disp "P VALUE"
:If C=1
:Then
:Disp P
:Else
:If C=2
:Then
:Disp Q
:Else
:Disp 2*min(P,Q)
:End
:End
:Stop
```

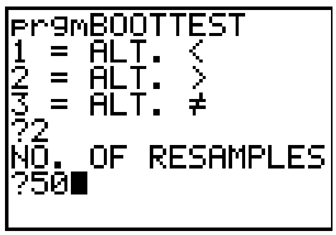
Example 14.9 French Fry Discrimination? A fast-food-restaurant customer complains that people 60 years old or older are given fewer french fries than people under 60. The owner responds by gathering data without knowledge of the employees. Below are the data on the weight of french fries (in grams) from random samples of the two groups of customers. Perform a permutation test using an appropriate alternative hypothesis and give the *p*-value.

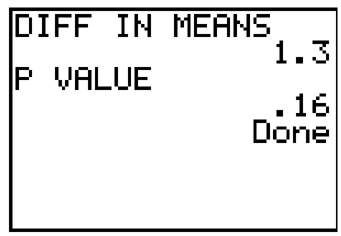
Age < 60:	75	77	80	69	73	76	78	74	75	81
Age ≥ 60:	68	74	77	71	73	75	80	77	78	72

Solution. We let μ_1 be the average weight of french fries served to customers under age 60, and let μ_2 be the average weight served to customers age 60 or older. We shall test H_0: $\mu_1 = \mu_2$ with the alternative H_A: $\mu_1 > \mu_2$.

First, we enter the sample weights for the under age 60 customers into list `L1` and the weights for the other group into list `L2`. Then we execute the **BOOTTEST** program by entering **2** to designate the alternative $\mu_1 > \mu_2$. Below are the results from 50 resamples.

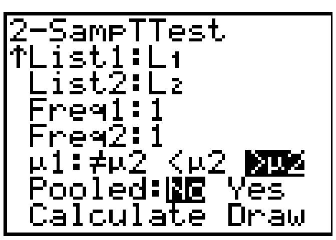


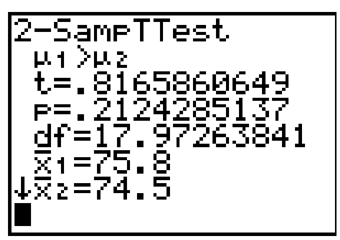




The 50 resampled differences in permuted mean are stored in list `L3`. The difference in the original sample means is $\bar{x}_1 - \bar{x}_2$ = 1.3, and 16% of the means in `L3` are greater than 1.3. Based on this sample of permutations, if $\mu_1 = \mu_2$, then there would be about a 16% chance of $\bar{x}_1 - \bar{x}_2$ being as high as 1.3 with samples of these sizes. This *p*-value of 0.16 does not provide enough evidence to reject H_0; it does not appear that the older individuals are getting fewer french fries.

Using the `2-SampTTest` feature, we see that the traditional two-sample *t* test gives a *p*-value of 0.21243 and leads us to the same conclusion.





Permutation Test for Paired Data

When we have paired sample data, such as "before and after" measurements, then we often consider the differences in measurements as one sample. The hypothesis test H_0: $\mu_1 = \mu_2$ then becomes H_0: $\mu_1 - \mu_2 = 0$ and may be tested with this one sample using a traditional *t* test or with many resamples using the permutation test. The **BTPRTEST** program may be used for this permutation test. It also may be used to test whether the correlation equals 0.

The BTPRTEST Program

```
Program:BTPRTEST
:Disp "1 = PAIRED MEAN","2 = CORRELATION"
:Input T
:ClrHome
:Disp "1 = ALT. <","2 = ALT. >","3 = ALT. _"
:Input C
:Disp "NO. OF RESAMPLES"
:Input B
:dim(L1)↲N
:ClrList L3
:For(I,1,B)
:ClrList L5,L6
:If T=2
:Then
:L1↲L5
:randInt(1,N,N) ↲L4
:For(J,1,N)
:L2(L4(J)) ↲L6(J)
:End
:LinReg(ax+b) L5,L6
:r↲L3(I)
:Else
:For(J,1,N)
:randInt(0,1) ↲A
:If A=0
:Then
:L1(J) ↲L5(J)
:L2(J) ↲L6(J)
:Else
:L1(J) ↲L6(J)
:L2(J) ↲L5(J)
:End
:End
:If T=1
:Then
:mean(L5)-mean(L6)↲L3(I)
:End
:End
:If T=1
:Then
:mean(L1)-mean(L2)↲X
:Else
:LinReg(ax+b) L1,L2
:r↲X
:End
:SortA(L3)
:X+1↲L3(B+1)
:0↲K
:While L3(K+1)<X
:K+1↲K
:End
:K/B↲P
:While L3(K+1)≤X
:K+1↲K
:End
:(B-K)/B↲Q
:seq(L3(I),I,1,B) ↲L3
:ClrHome
:If T=1
:Then
:Disp "DIFF. IN MEANS"
:Else
:Disp "SAMPLE CORR."
:End
:Disp X
:Disp "P VALUE"
:If C=1
:Then
:Disp P
:Else
:If C=2
:Then
:Disp Q
:Else
:Disp 2*min(P,Q)
:End
:End
:Stop
```

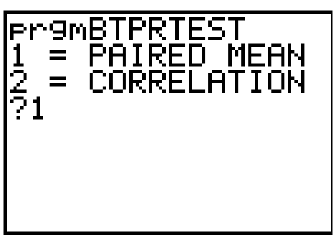
Before executing the **BTPRTEST** program, enter the data into lists `L1` and `L2`. When prompted, enter **1** or **2** to designate the desired test, then enter **1**, **2**, or **3** to designate the desired alternative. The resampled permuted pair differences in mean (or correlation) are ordered and stored in list L3. The statistic from the original paired sample is displayed along with the *p*-value of the permutation test.

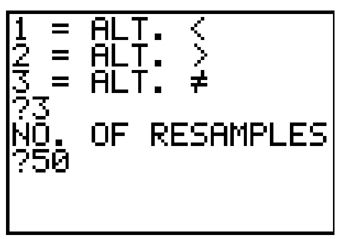
Example 14.10 Are the Technicians Consistent? Below is the data from Example 7.6 (page 85) that gives the total body bone mineral content of eight subjects as measured by two different X-ray machine operators. Perform a matched pairs permutation test to decide whether or not the two operators have the same mean.

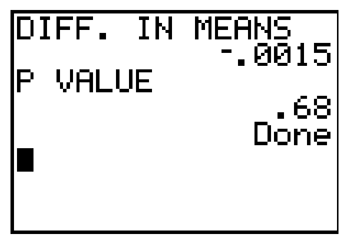
	Subject							
Operator	1	2	3	4	5	6	7	8
1	1.328	1.342	1.075	1.228	0.939	1.004	1.178	1.286
2	1.323	1.322	1.073	1.233	0.934	1.019	1.184	1.304

Solution. We let μ_1 be the average measurement from Operator 1 and let μ_2 be the average measurement from Operator 2. We shall test H_0: $\mu_1 = \mu_2$ with the alternative H_A: $\mu_1 \neq \mu_2$.

First, we enter the measurements into lists L1 and L2. Then we execute the **BTPRTEST** program by entering **1** to designate a paired mean, and then entering **3** for the alternative $\mu_1 \neq \mu_2$. Next, we display the results from 50 resamples.



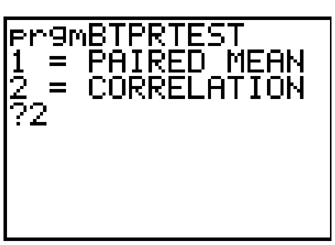


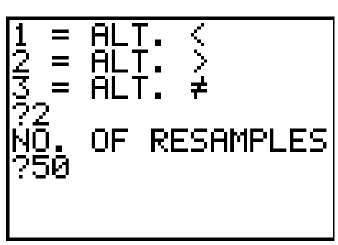


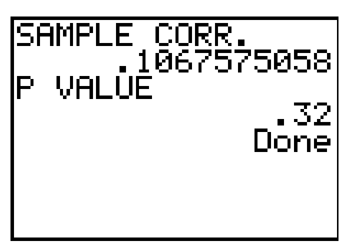
The 50 resampled permuted pair differences in mean are stored in list L3. The difference in the original sample means is $\bar{x}_1 - \bar{x}_2 = -0.0015$ and 68%/2 = 34% of the means in L3 are less than –0.0015. This large *p*-value means that we do not have significant evidence to argue that there is a difference in the operator's mean measurement.

Example 14.11 More on Baseball Salaries. Use the data from Example 14.8 to test whether the correlation between salary and batting average is greater than 0.

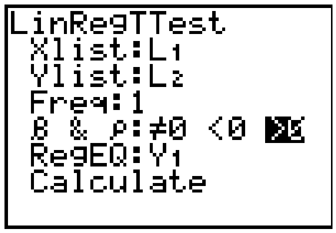
Solution. First, we enter the data into lists L1 and L2. Then to test the hypothesis H_0: $\rho = 0$ with alternative H_A: $\rho > 0$, we execute option 2 of the **BTPRTEST** program using the second alternative with 50 resamples.

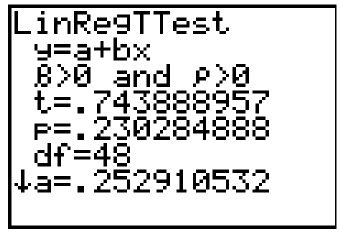






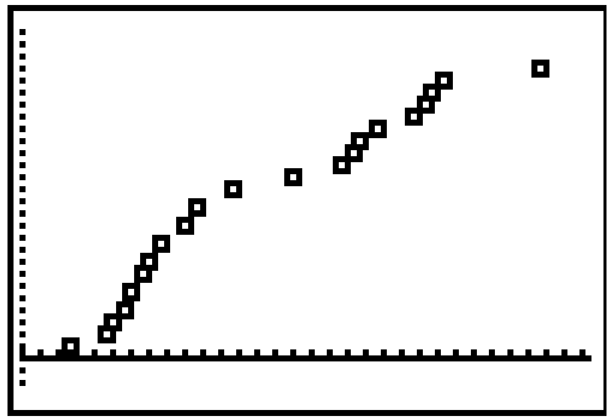
Upon running the program, the 50 resample correlations are stored in list L3. The sample correlation is r = 0.1067575. In this case, 16 of the resample correlations were greater than r, which gives us a one-sided p-value of 0.32. Thus, we do not have significant evidence to reject H_0 with this sample based upon 50 resamples. Our p-value of 0.32 compares favorably with the p-value of 0.23 obtained with the traditional linear regression t test.





CHAPTER

15



Nonparametric Tests

Introduction

In this chapter, we provide some supplementary programs for performing several nonparametric hypothesis tests.

These tests relax the assumption of Normal distribution of the data (or sample mean). They are less powerful than parametric tests, since they do not use all the information in the data but only the ranks (sorted order statistics); one has information on the smallest, next smallest, and so forth, but not the size of the differences.

The Rank Sum test is a stand-in for a two sample *t* test; the signed rank for a paired samples test, and the Kruskal-Wallis for one-way ANOVA.

15.1 The Wilcoxon Rank Sum Test

We first provide the program **RANKSUM** to perform the Wilcoxon rank sum test on data from two populations. To execute the program, we must enter the data into lists `L1` and `L2`. The program sorts each list, then merges and sorts the lists into list `L3`. Then it stores the rank of each measurement in `L3` next to it in list `L4`. All sequences of ties are assigned an average rank.

The Wilcoxon test statistic W is the sum of the ranks from `L1`. Assuming that the two populations have the same continuous distribution (and no ties occur), then W has a mean and standard deviation given by

$$\mu = \frac{n_1(N+1)}{2} \quad \text{and} \quad \sigma = \sqrt{\frac{n_1 n_2 (N+1)}{12}}$$

where n_1 is the sample size from `L1`, n_2 is the sample size from `L2`, and $N = n_1 + n_2$.

We test the null hypothesis H_0: there is no difference in distributions. A one-sided alternative is H_A: the first population yields higher measurements. We use this alternative if we expect or see that W is a much higher sum than its expected sum of ranks μ. In this case, the p-value is given by a normal approximation. We let W ~ $N(\mu, \sigma)$ and compute the right-tail $P(X \geq W)$ (using the continuity correction if W is an integer).

If we expect or see that W is the much lower sum than its expected sum of ranks μ, then we should use the alternative H_A: first population yields lower measurements. In this case, the p-value is given by the left-tail $P(X \leq W)$, again using continuity correction if needed. If the two sums of ranks are close, then we could use a two-sided alternative H_A: there is a difference in distributions. In this case, the p-value is given by twice the smallest tail value: $2 * P(X \geq W)$ if $W > \mu$, or $2 * P(X \leq W)$ if $W < \mu$.

The **RANKSUM** program displays the expected sum of ranks from the first list and the actual sums of the ranks from `L1` and `L2`. It also displays the smallest tail value created by the test statistic. That is, it displays $P(X \geq W)$ if $W > \mu$ and it displays $P(X \leq W)$ if $W < \mu$; for a two-tail alternate hypothesis, multiply the given p-value by 2. Conclusions for any alternative then can be drawn from this value. We note that if there are ties, then the validity of this test is questionable.

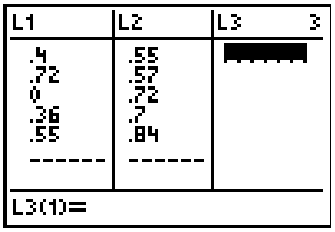
Example 15.1 Retelling Stories. Below are language usage scores of kindergarten students who were classified as high-progress readers or low-progress readers when asked to retell a story that had been read to them. Is there evidence that the scores of high- progress readers are higher than those of low-progress readers? Carry out a two-sample *t* test. Then carry out the Wilcoxon rank sum test and compare the conclusions for each test.

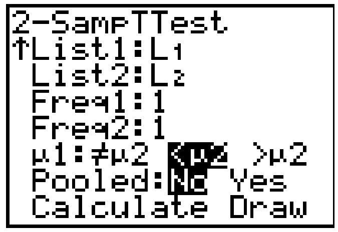
Child	Progress	Score
1	high	0.55
2	high	0.57
3	high	0.72
4	high	0.70
5	high	0.84
6	low	0.40
7	low	0.72
8	low	0.00
9	low	0.36
10	low	0.55

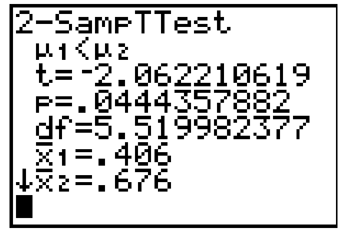
The RANKSUM Program

```
PROGRAM:RANKSUM
:ClrList L4
:SortA(L1):SortA(L2)
:augment(L1,L2) ↵L3
:SortA(L3)
:dim(L1) ↵M:dim(L2) ↵N
:L3 ↵L6
:L3(1)-1 ↵L3(M+N+1)
:1 ↵B
:1↵I
:Lbl 1
:While I<(M+N)
:If L3(I)<L3(I+1)
:Then
:B↵L4(I)
:1+I↵I
:1+B↵B
:Goto 1
:Else
:1↵J
:B↵S
:Lbl 2
:While L3(I)=L3(I+J)
:S+B+J↵S
:1+J↵J
:Goto 2
:End
:S/J↵T
:For(K,0,J-1)
:T↵L4(I+K)
:End
:I+J↵I:B+J↵B
:Goto 1
:End
:End
:If I=M+N
:Then
:M+N↵L4(I)
:End
:1↵I:0↵S:0↵J
:Lbl 3
:While I≤M
:Lbl 4
:If L1(I)=L3(I+J)
:Then
:S+L4(I+J) ↵S
:Else
:1+J↵J
:Goto 4
:End
:I+1↵I
:Goto 3
:End
:L6↵L3
:ClrList L6
:(M+N)(M+N+1)/2↵R
:M*(M+N+1)/2↵U
:√(M*N*(M+N+1)/12) ↵D
:(abs(S-U)-.5)/D↵Z
:If int(S)≠S:(abs(S-U))/D↵Z:End
:0.50-normalcdf(0,Z,0,1) ↵P
:Disp "EXPECTED 1ST SUM"
:Disp .5*M*(M+N+1)
:Disp "SUMS OF RANKS"
:Disp {S,R-S}
:If S=U:Disp "NO DIFFERENCE"
:If S<U:Disp "LEFT TAIL",round(P,4)
:If S>U:Disp "RIGHT TAIL",round(P,4)
:Stop
```

Solution. First, we enter the five low-progress scores into `L1` and the five high-progress scores into `L2`. Then we use the `2-SampTTest` from the `STAT TESTS` menu to test the hypothesis H_0: $\mu_1 = \mu_2$ versus the alternative H_A: $\mu_1 < \mu_2$.

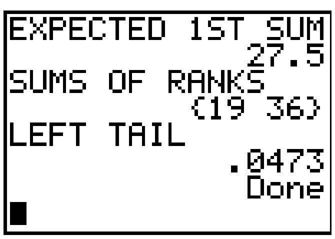


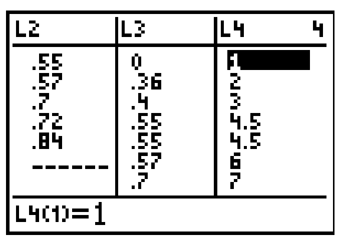


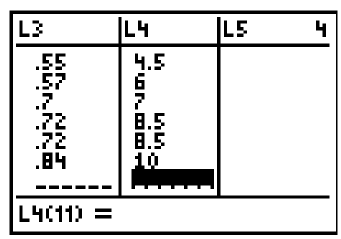


We obtain a *t* statistic of –2.06221 and a *p*-value of 0.0444 for the one-sided alternative. With the rather small *p*-value, we have significant evidence to reject H_0 and say that the average score of all high-progress readers is higher than the average score of all low-progress readers. For if H_0 were true, then there would be only a 0.0444 probability of obtaining a high-progress sample mean that is so much larger than the low-progress sample mean (0.676 compared to 0.406).

Now for the Wilcoxon rank sum test, we use H_0: the distribution is the same for both groups versus H_A: high-progress readers score higher when retelling the story. The Wilcoxon test statistic is the sum of ranks from `L1` in which we entered the low-progress scores. With the data entered into lists `L1` and `L2`, we now execute the **RANKSUM** program, and then observe the sorted data and ranks in lists `L3` and `L4`.







The sum of the ranks from the low-progress readers is 19, which is lower than the expected average of μ = (5*11)/2 = 27.5. According to the Wilcoxon test, if the distributions were the same, then there would be only be a 0.0473 probability (from the left-tail value) of the low- progress sum of ranks being so much smaller than the expected average of 27.5. Therefore, we should reject H_0 in favor of the alternative; high-progress readers are better at retelling stories told to them.

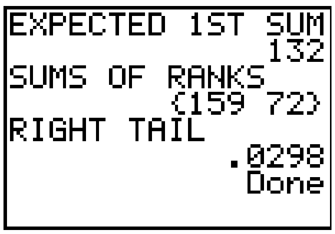
In this case, the Wilcoxon *p*-value is slightly higher than the *t* test *p*-value; however, both are low enough to result in the same conclusion.

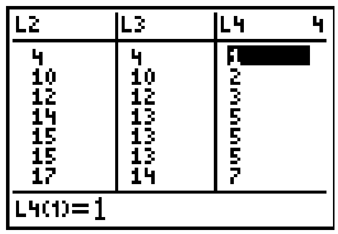
Example 15.2 Logging in the Rain Forest. Below is a comparison of the number of tree species in unlogged plots in the rain forest of Borneo with the number of species in plots logged eight years earlier.

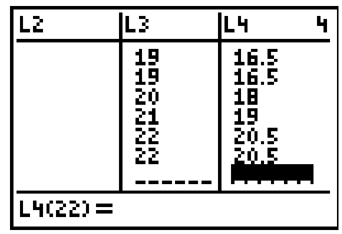
Unlogged	22	18	22	20	15	21	13	13	19	13	19	15
Logged	17	4	18	14	18	15	15	10	12			

Does logging significantly reduce the mean (median) number of species in a plot after eight years? State the hypotheses, do a Wilcoxon rank sum test, and state your conclusion.

Solution. We will test the hypothesis H_0: there is no difference in medians (or distributions) versus the alternative H_A: the unlogged median is higher. To do so, we first enter the unlogged measurements into list `L1` and enter the logged measurements into list `L2`. Then we execute the **RANKSUM** program which produces the following results:







We note that there are 21 total measurements with 12 unlogged measurements. If there were no difference in distributions or medians, then we would expect the sum of ranks from `L1` to be $\mu = 12 * 22/2 = 132$. But if there were no difference in medians, then there would be only a 2.98% chance of the sum of ranks from `L1` being as high as 159. This low *p*-value gives significant evidence to reject H_0 in favor of the alternative; there are more species of trees in the unlogged plots.

15.2 The Wilcoxon Signed Rank Test

Here we provide the **SIGNRANK** program to perform the Wilcoxon signed rank test on data sets of size *n* from two populations (this can also be used to test a single median). To execute the program, we must enter the data into lists `L1` and `L2`. The program will sort the absolute value of the differences `L2 – L1` into list `L3`, but it will disregard any zero differences. The sample size *n* is decreased so as to count only the nonzero differences. Then the program puts the rank of each measurement in `L3` next to it in `L4`. All sequences of ties are assigned an average rank.

The Wilcoxon test statistic *W* is the sum of the ranks from the positive differences. Assuming that the two populations have the same continuous distribution (and no ties occur), then *W* has a mean and standard deviation given by

$$\mu = \frac{n(n+1)}{4} \quad \text{and} \quad \sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

The SIGNRANK Program

```
PROGRAM:SIGNRANK
:ClrList L3,L4,L5
:0↵L3(1):0↵L4(1):0↵L5(1)
:0↵S:0↵R:0↵U:0↵V:1↵J
:For(I,1,dim(L1),1)
:If L2(I)-L1(I)≠0
:Then
:abs(L2(I)-L1(I)) ↵L3(J)
:1+J↵J
:End:End
:If L3(1)>0
:Then
:SortA(L3):dim(L3) ↵N
:L3↵L6:L3(1)-1↵L3(N+1)
:1↵B:1↵I
:Lbl 1
:If I<N
:Then
:If L3(I)<L3(I+1)
:Then
:B↵L4(I)
:1+I↵I:1+B↵B
:Goto 1
:Else
:1↵J:B↵S
:Lbl 2
:If L3(I)=L3(I+J)
:Then
:S+B+J↵S:1+J↵J
:Goto 2
:End
:S/J↵T
:For(K,0,J-1)
:T↵L4(I+K)
:End
:I+J↵I:B+J↵B
:End
:Goto 1
:End
:If I=N
:Then
:N↵L4(I)
:End
:0↵J
:For(I,1,dim(L1))
:If (L2(I)-L1(I))>0
:Then
:1+J↵J:abs((L2(I)-L`(I)) ↵L5(J)
:End:End
:SortA(L5)
:1↵I:0↵J
:If L5(1)>0
:Then
:Lbl 3
:While I≤dim(L5)
:Lbl 4
:If L5(I)=L3(I+J)
:Then
:V+L4(I+J) ↵V
:Else
:1+J↵J
:Goto 4
:End
:I+1↵I
:Goto 3
:End:End
:L6↵L3:ClrList L6
:N(N+1)/2↵R:N(N+1)/4↵U
:√(N(N+1)(2N+1)/24) ↵D
:(abs(V-U)-.5)/D↵Z
:If int(V)≠V:(abs(V-U))/D↵Z:End
:0.50-normalcdf(0,Z,0,1) ↵P
:Disp "EXP. + SUM"
:Disp N*(N+1)/4
:Disp "SUMS -,+ RANKS"
:Disp {R-V,V}
:If V=U:Disp "NO DIFFERENCE"
:If V<U:Disp "LEFT TAIL",round(P,4)
:If V>U:Disp "RIGHT TAIL",round(P,4)
:Stop
```

We test the null hypothesis H_0: there is no difference in distributions. A one-sided alternative may be H_A: the second population yields higher measurements. We use this alternative if we expect or see that W is a much higher sum, which means that there were more positive differences in `L2 – L1`. In this case, the p-value is given by a normal approximation. We let $X \sim N(\mu, \sigma)$ and compute the right-tail $P(X \geq W)$ (using a continuity correction if W is an integer).

If we expect or see that W is the much lower sum, then there were more negative differences. Now we should use the alternative H_A: the second population yields lower measurements. In this case, the p-value is given by the left-tail $P(X \leq W)$, again using continuity correction if needed. If the two sums of ranks are close, we could use a two-sided alternative H_A: there is a difference in distributions. In this case, the p-value is given by twice the smallest tail value.

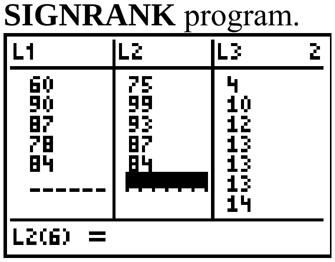
The **SIGNRANK** program displays the sums of the ranks of the negative differences and of the positive differences as well as the smallest tail value created by the test statistic. That is, it displays $P(X \geq W)$ if $W > \mu$, or $P(X \leq W)$ if $W < \mu$. Conclusions for any alternative can then be drawn from this value. Again, we note that if there were ties, then the validity of this test is questionable.

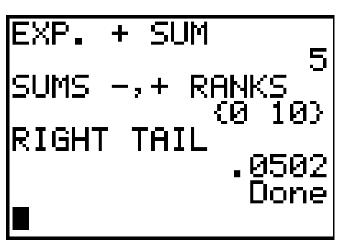
Example 15.3 Stepping and Heart Rates. A student project asked subjects to step up and down for three minutes. There were two treatments: stepping at a low rate (14 steps per minute) and a medium rate (21 steps per minute). Here are data for heart rates for five subjects and the two treatments. For each subject we have the initial resting heart rate and the heart rate at the end of the exercise. Does exercise at the low rate raise the heart rate significantly? State hypotheses in terms of the median increase in heart rate and apply the Wilcoxon signed rank test.

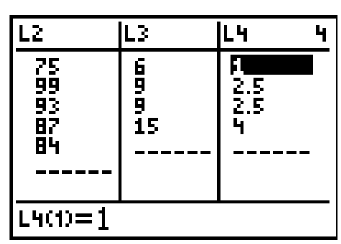
	Low rate		**Medium rate**	
Subject	**Resting**	**Final**	**Resting**	**Final**
1	60	75	63	84
2	90	99	69	93
3	87	93	81	96
4	78	87	75	90
5	84	84	90	108

Solution. We will test the hypothesis H_0: For the low stepping rate, resting and final heart rates have the same median versus H_A: final heart rates are higher.

Enter the five low rate resting heart rates into list `L1` and the five low rate final heart rates into list `L2`. The alternative means that there should be more positive differences, so that the sum of the positive ranks should be higher. Therefore, the p-value comes from the right-tail probability created by the test statistic. After the data is entered, execute the **SIGNRANK** program.







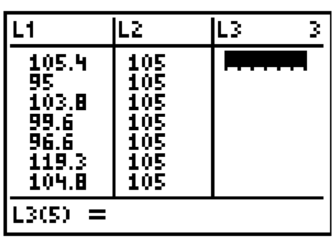
List L3 now contains the ordered absolute values of the four nonzero differences. Their corresponding (averaged) ranks are adjacent in list L4. List L5 contains only the positive differences, which in this case are all four of the differences.

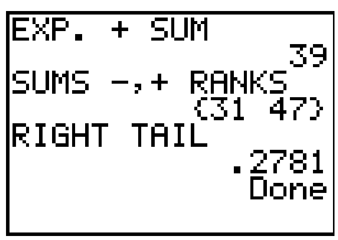
We see that the sum of the ranks of the positive differences is much higher than that of the negative differences. If the medians for each rate were the same, then there would be only a 0.0502 probability of the sum of positive ranks being as high as 10. They were expected to be 5 with the four subjects for which there is a difference. The relatively low *p*-value provides some evidence to reject H_0 and conclude that the median final heart rate is higher for the low rate test.

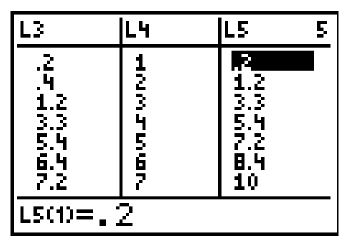
Example 15.4 Radon Detector Accuracy. Below are the readings from 12 home radon detectors exposed to 105 pCi/l of radon. We want to know if these detectors are accurate. Apply the Wilcoxon signed rank test to determine if the median reading from all such home radon detectors differs significantly from 105.

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

Solution. We will test the null hypothesis H_0: median = 105 versus H_A: median ≠ 105. First, we enter the given data into list L1 and then enter 105 into list L2 twelve times. If H_0 were true, then we would expect the sum of ranked positive differences of L2 – L1 to be (12*13) /4 = 39. But H_A implies that this sum of ranked positive differences will be either much higher than 39 or much lower than 39. To test the hypotheses, we execute the **SIGNRANK** program after entering these data into the lists.







Lists L3 and L4 will show that there were 12 nonzero differences in L2 – L1, and list L5 will show that eight of these were positive differences, meaning that there were eight times in which the home radon detector measured below 105.

The right-tail value is given as 0.2781; thus, the *p*-value for the two-sided alternative is 2*0.2781 = 0.5562. If the median home radon measurement were 105, then there would be a 0.5562 probability of the sum of positive ranks being as far away (in either direction) from the expected sum of 39 as the resulting sum of 47 is. Thus, we do not have significant evidence to reject H_0 and can conclude that, "on average," these detectors are accurate.

15.3 The Kruskal-Wallis Test

Our next program, **KRUSKAL**, is for the Kruskal-Wallis test, which simultaneously compares the distribution of more than two populations. We test the null hypothesis H_0: all populations have the same distribution versus the alternative H_A: measurements are systematically higher in some populations. To apply the test, we draw independent SRSs of sizes $n_1, n_2, \ldots, n_I$ from I populations. There are N observations in all. We rank all N observations and let R_i be the sum of the ranks for the ith sample. The Kruskal-Wallis statistic is

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{I} \frac{R_i^2}{n_i} - 3(N+1)$$

When the sample sizes are large and all I populations have the same continuous distribution, then H has an approximate chi-square distribution with $I-1$ degrees of freedom. When H is large, creating a small right-tail p-value, then we can reject the hypothesis that all populations have the same distribution.

The KRUSKAL Program

```
PROGRAM:KRUSKAL
:ClrList L3,L4,L5
:dim([B])↵L1
:sum(seq([B](1,J),J,1,L1(2))) ↵L
:1↵K
:For(J,1,L1(2))
:For(I,1,[B](1,J))
:[A](I,J) ↵L3(K)
:1+K↵K
:End:End
:SortA(L3)
;L3↵L6:L3(1)-1↵L3(L+1)
:1↵B:1↵I
:Lbl 1
:While I<(L)
:If L3(I)<L3(I+1)
:Then
:If I=L
:Then
:L↵L4(I)
:End
:1↵K
:Lbl 5
:While K≤L1(2)
:ClrList L2
:B↵L4(I)
:1+I↵I:1+B↵B
:Goto 1
:Else
:1↵J:B↵S
:Lbl 2
:While L3(I)=L3(I+J)
:S+B+J↵S:1+J↵J
:Goto 2
:End
:S/J↵T
:For(K,0,J-1)
:T↵L4(I+K)
:End
:I+J↵I:B+J↵B
:Goto 1
:End:End
:Then
:S+L4(I+J) ↵S:1+I↵I
:Else
:1+J↵J
:End
:Goto 3
:End
:S↵L5(K):1+K↵K
```

```
:For(I,1,[B](1,K))
:[A](I,K) ↵L2(I)
:SortA(L2)
:End
:1↵I:0↵S:0↵J
:Lbl 3
:While I≤[B](1,K)
:Lbl 4
:If L2(I)=L3(I+J)
```

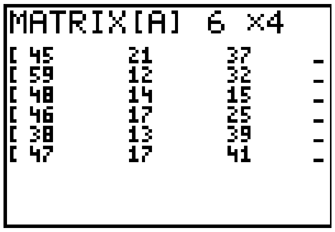
```
:Goto 5
:End
:L6↵L3:ClrList L2,L6
:12/L/(L+1)*sum(seq(L5(I)^2/
  [B](1,I),I,1,L1(2)))-3(L+1) ↵W
:1-↵χ2cdf(0,W,L/(2)-1) ↵P
:Disp "TEST STAT",W
:Disp "P-VALUE",P
```

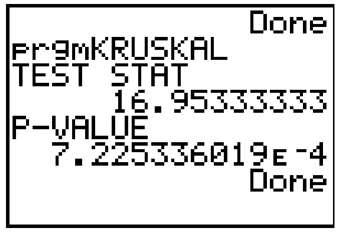
Before executing the **KRUSKAL** program, use `MATRX EDIT` to enter the data as columns into matrix [`A`] and to enter the sample sizes as a row into matrix [`B`**]**.

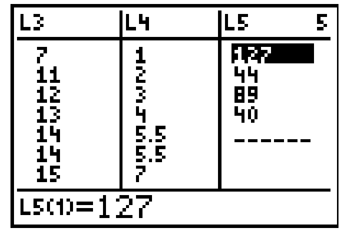
Example 15.5 Are Insects Colorblind? An experiment was conducted to determine if insects were equally attracted by different colors. Sticky boards were placed in a field of oats, and the number of cereal leaf beetles trapped was counted. Use the Kruskal-Wallis test to see if there are significant differences in the numbers of insects trapped by the different board colors.

Board color	**Insects trapped**					
Lemon yellow	45	59	48	46	38	47
White	21	12	14	17	13	17
Green	37	32	15	25	39	41
Blue	16	11	20	21	14	7

Solution. To execute the **KRUSKAL** program, we must use `MATRX EDIT` to enter the data into matrix `[A]` and enter the sample sizes into matrix `[B]`. First, enter the data into the columns of the 6×4 matrix `[A]` as you would normally enter data into lists. Next, enter the sample sizes into a 1×4 matrix `[B]`. Then execute the **KRUSKAL** program.



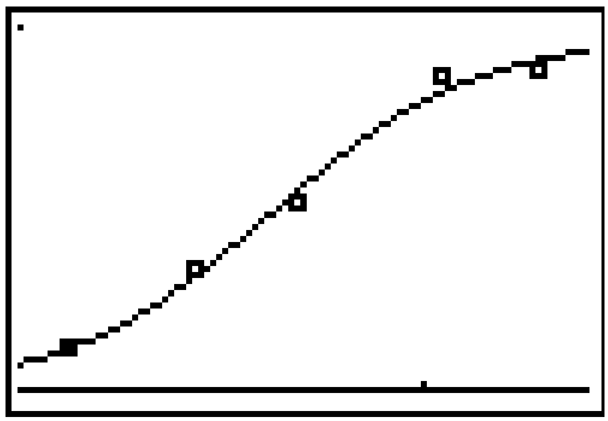




After the program completes, view the entries in lists L3, L4, and L5. List L3 contains the merged, sorted measurements, and list L4 contains their (averaged) ranks. List L5 contains the sum of ranks from each type of color. The low *p*-value of 0.00072 gives good evidence to reject the hypothesis that all colors yield the same distribution of insects trapped.

CHAPTER

16



Logistic Regression

Introduction

In this chapter, we give a brief discussion of two types of logistic regression fits. The first type is a linear fit for the logarithm of the odds ratio of two population proportions. The second type is the general logistic fit for several population proportions.

We also discuss the logistic regression model that is commonly used in modeling population sizes.

16.1 The Logistic Regression Model

First, we provide a supplementary program that computes appropriate mathematical odds for a given probability p of an event A. If $p \le 0.50$, then the odds *against* A are given as the ratio $(1-p):p$. If $p > 0.50$, then the odds *in favor of* A are given as the ratio $p:(1-p)$. The probability can be entered either as a decimal or as a fraction. If p is entered as a decimal, then the odds are computed after rounding p to four decimal places; thus some accuracy may be lost.

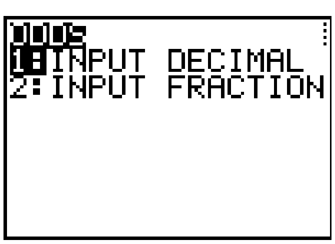
The ODDS Program

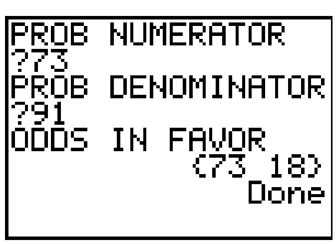
```
PROGRAM:ODDS
:Menu("ODDS","INPUT DECIMAL",
 1,"INPUT FRACTION",2)
:Lbl 1
:Disp "PROBABILITY"
:Input P
:round(P,4)↵P
:10000*P↵A
:10000*(1-P) ↵B
:A/gcd(A,B) ↵N
:B/gcd(A,B) ↵D
:Goto 3
:End
:Lbl 2
:Disp "PROB NUMERATOR"
:Input A
:Disp "PROB DENOMINATOR"
:Input B
:A/B↵P
:A/gcd(A,B-A) ↵N
:(B-A)/gcd(A,B-A) ↵D
:Lbl 3
:If P>.50
:Then
:Disp "ODDS IN FAVOR"
:Disp {N,D}
:Else
:Disp "ODDS AGAINST"
:Disp {D,N}
:Stop
```

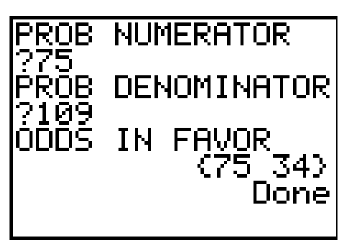
Example 16.1 Employee Stock Options. In a study of 91 high-tech companies and 109 non-high-tech companies, 73 of the high-tech companies and 75 of the non-high-tech companies offered incentive stock options to key employees.

(a) What proportion of high-tech companies offer stock options to their key employees? What are the odds?
(b) What proportion of non-high-tech companies offer stock options to their key employees? What are the odds?
(c) Find the odds ratio using the odds for the high-tech companies in the numerator.

Solution. (a) The proportion of high-tech companies that offer stock options is simply 73/91 = 0.8022. To compute the odds, we can use the **ODDS** program. After bringing up the program, enter the numerator value of **73** followed by the denominator value of **91**. The odds in favor are displayed as 73 to 18.

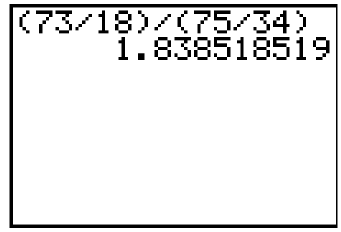






(b) For the non-high-tech companies, the proportion is 75/109 = 0.688. Because 34 of these companies do not offer stock options, the odds are 75 : 34 that such a company does offer stock options to their key employees. This result can be verified with the **ODDS** program.

(c) The odds-in-favor ratio can be computed by simple division of the odds (73/18)/(75/34). Thus, the odds in favor of a high-tech company offering stock options are about 1.8 times more than the odds for a non-high-tech company.



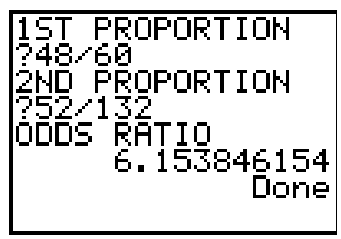
The logistic regression model is always based on the odds in *favor* of an event. It provides another method of studying the odds in favor ratio between two populations. As a lead-in, we now provide a program that specifically computes the odds-in-favor ratio. This program merely automates the division done just above.

The ODDS2 Program

PROGRAM:ODDS2	:Input R
:Disp "1ST PROPORTION"	:Disp "ODDS RATIO"
:Input P	:Disp P/(1-P)/(R/(1-R))
:Disp "2ND PROPORTION"	:Stop

Example 16.2 Gender Bias. In a study on gender bias in textbooks, 48 out of 60 female references were "girl." Also, 52 out of 132 male references were "boy." These two types of references were denoted as juvenile references. Compute the odds ratio for comparing the female juvenile references to the male juvenile references.

Solution. We simply enter the data into the **ODDS2** program that computes the ratio of odds. We see that the odds in favor of a juvenile female response are more than six times the odds in favor of a juvenile male response.



Model for Logistic Regression

The logistic regression model is given by the equation

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

where ln is the natural (base *e*) logarithm and *x* is either 1 or 0 to designate the explanatory variable. We now provide a program that computes and displays the regression coefficients for the fit ln(ODDS) = $b_0 + b_1 x$ as well as the odds ratio e^{b_1}.

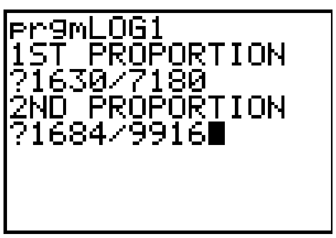
The LOG1 Program

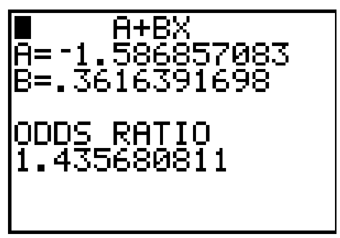
```
PROGRAM:LOG1
:Disp "1ST PROPORTION"
:Input P
:Disp "2ND PROPORTION"
:Input R
:ln(R/(1-R))↵A
:ln(P/(1-P))-A↵B
:ClrHome
:Output(1,6,"A+BX")
:Output(2,1,"A=")
:Output(2,3,A)
:Output(3,1,"B=")
:Output(3,3,B)
:Output(5,1,"ODDS RATIO")
:Output(6,1,e^(B))
:Stop
```

Example 16.3 Binge Drinking on Campus. The table below gives data on the numbers of men and women who responded "Yes" to being frequent binge drinkers in a survey of college students. Find the coefficients for the logistic regression model and the odds ratio of men to women.

Population	*X*	*n*
Men	1630	7180
Women	1684	9916

Solution. We simply enter the data into the **LOG1** program to obtain the equation $\log(ODDS) = -1.59 + .36x$. The odds ratio is also displayed.





16.2 Inference for Logistic Regression

To compute confidence intervals for the slope β_1 and the odds ratio of the logistic regression model, we can use the **ODDSINT** program that follows. To execute the program, separately enter the numerators and denominators of the two sample proportions and the desired level of confidence.

The ODDSINT Program

```
Program:ODDSINT
:Disp "1ST NO. OF YES"
:Input M
:Disp "1ST SAMPLE SIZE"
:Input P
:Disp "2ND NO. OF YES"
:Input N
:Disp "2ND SAMPLE SIZE"
:Input Q
:Disp "CONF. LEVEL"
:Input L
:√(1/M+1/(P-M)+1/N+1/(Q-N))↵S
:invNorm((L+1)/2,0,1) ↵R
:ln(M/P/(1-M/P))-ln(N/Q/(1-N/Q)) ↵B
:ClrHome
:Disp "STD. ERROR"
:Disp S
:Disp "SLOPE INTERVAL"
:Disp {round(B-R*S,4),round(B+R*S,4)}
:Disp "ODDS RATIO INT."
:Disp {round(e^(B-R*S),4),
 round(e^(B+R*S), 4)}
:Stop
```

To test the hypothesis that an odds ratio equals 1, we equivalently can test whether the logistic regression model coefficient b_1 equals 0. To do so, we use the *p*-value given by

$$P(\chi 2(1) \geq (b_1 / SE(b_1)^2))$$

where $SE(b_1)$ is the standard error of the coefficient b_1. The **ODDSTEST** program that follows computes this *p*-value upon entering the values of the two proportions under consideration and the value of the standard error $SE(b_1)$.

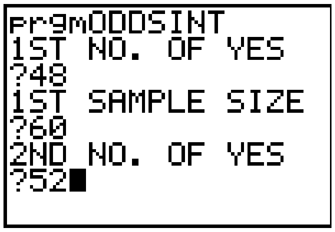
The ODDSTEST Program

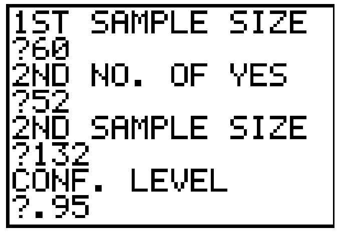
```
Program:ODDSTEST
:Disp "1ST PROPORTION"
:Input P
:Disp "2ND PROPORTION"
:Input R
:Disp "ST.ERROR OF B1"
:Input S
:ln(R/(1-R))↵A
:ln(P/(1-P))-A↵B
:(B/S)^2↵Z
:1-χ2cdf(0,Z,1) ↵P
:ClrHome
:Disp "TEST STAT",Z
:Disp "P-VALUE",P
```

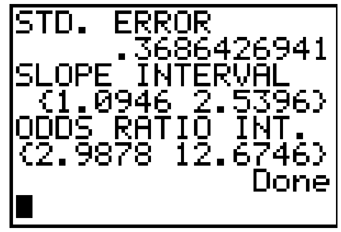
Example 16.4 More Gender Bias. In the study on gender bias in textbooks from Example 16.2, 48 out of 60 female references were "girl" and 52 out of 132 male references were "boy."

(a) Give a 95% confidence interval for the slope.
(b) Calculate the $\chi 2$ statistic for testing the null hypothesis that the slope is zero and give the approximate *p*-value.

Solution. (a) We execute the **ODDSINT** program and input the proportion information.

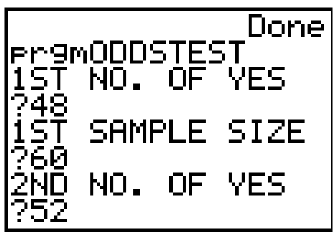


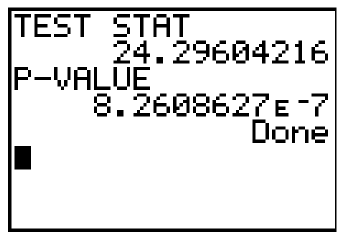




We obtain a 95% confidence interval for the slope of (1.0946, 2.5396) which is equivalent to an odds ration interval of (2.9878, 12.6746). Because 0 is not in the slope interval and 1 is not in the odds ratio interval, we have some evidence to reject the null hypothesis that the slope is zero; the odds of juvenile references are not equal between the two genders.

(b) We execute the **ODDSTEST** program and input the information from the samples.





The $\chi 2$ statistic of 24.296 yields a very low *p*-value of about 8.26×10^{-7} that gives significant evidence to reject the null hypothesis that the slope equals zero (or that the odds ratio is 1).

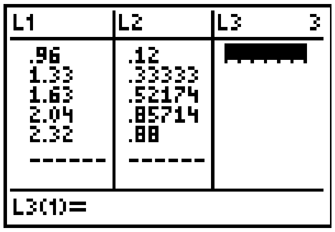
The Logistic Curve

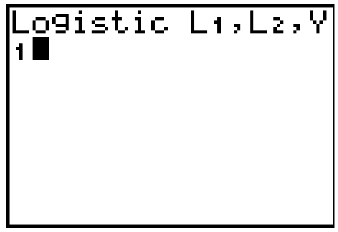
A general logistic curve is given by the function $p = \dfrac{c}{1+ae^{-bx}}$. Such a fit can be obtained with the **Logistic** command from the `STAT CALC` menu. Following is an example to illustrate this fit.

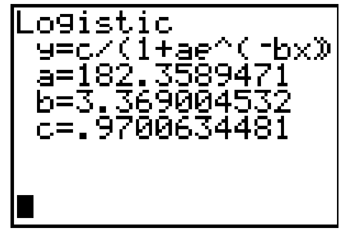
Example 16.5 Insecticide Effectiveness. An experiment was designed to examine how well an insecticide kills a certain type of insect. Find the logistic regression curve for the proportion of insects killed as a function of the insecticide concentration.

Concentration	Number of insects	Insects killed
0.96	50	6
1.33	48	16
1.63	46	24
2.04	49	42
2.32	50	44

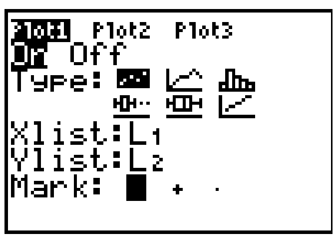
Solution. First, enter the data into the `STAT EDIT` screen with the concentrations in list `L1` and the proportions of insects killed in list `L2`. Next, enter the command `Logistic L1,L2,Y1` to compute the regression fit and to store the equation in function `Y1`. If you have forgotten, to get to `Y1`, press □, ~ to `Y-VARS`, ÷ for `Function`, ÷ for `Y1`.

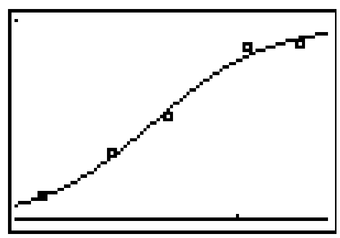






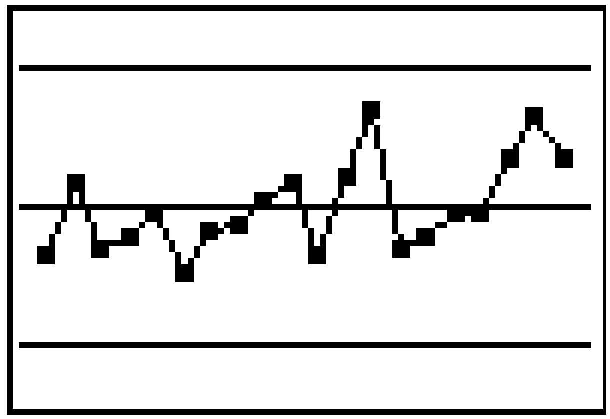
We obtain a logistic regression fit of $y = \dfrac{.97}{1+182.359e^{-3.369x}}$. If desired, we can make a scatterplot of the data along with the logistics regression curve. This clearly shows the increase in effectiveness of the insecticide until, finally, increasing the concentration does little to increase the effectiveness.





CHAPTER

17



Statistics for Quality: Control and Capability

Introduction

In this chapter, we provide several programs for computing control limits, graphing control charts, and for computing the capability indices of a process.

17.1 Statistical Process Control

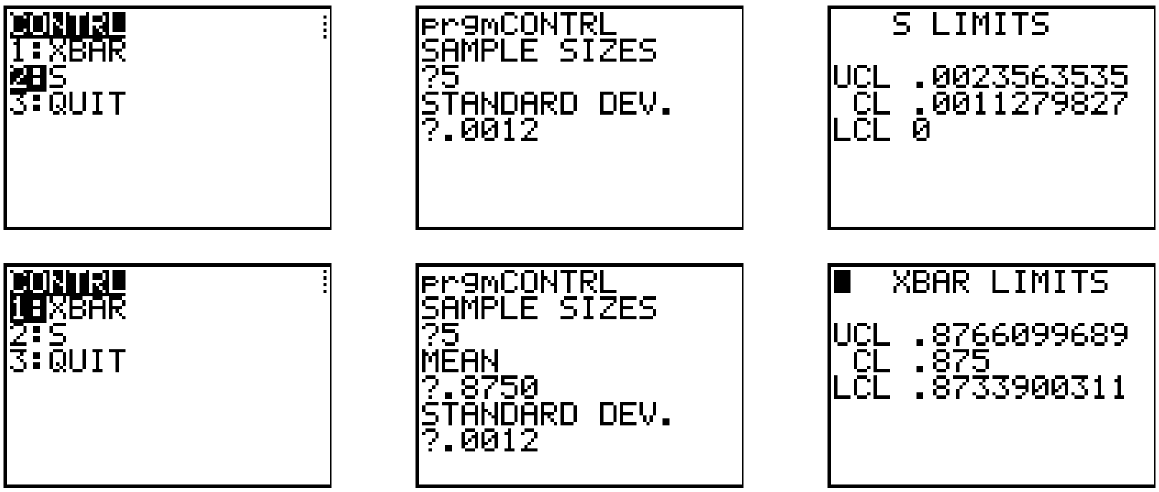
In this section, we provide a program that computes the upper and lower control limits and graphs the control charts for $\bar{x}$ and *s*.

The CONTRL Program

```
PROGRAM:CONTRL
:Menu("CONTRL","XBAR",1,"S",2,"
 QUIT",3)
:Lbl 1
:Disp "SAMPLE SIZES"
:Input N
:Disp "MEAN"
:Input M
:Disp "STANDARD DEV."
:Input S
:"M"→Y1
:"M+3*S/√(N)" →Y2
:"M-3*S/√ (N)" →Y3
:If dim(L1)>0
:Then
:seq(I,I,1,dim(L1))→L3
:0→Xmin
:dim(L1)+1→Xmax
:Xmax→Xscl
:min(min(L1),M-4*S/√ (N)) →Ymin
:max(max(L/),M+4*S/√ (N)) →Ymax
:Ymax-Ymin→Yscl
:Plot1(xyLine,L3,L1,·)
:End
:ClrHome
:Output(1,4,"XBAR LIMITS")
:Goto 4
:Lbl 2
:Disp "SAMPLE SIZES"
:Input N
:Disp "STANDARD DEV."
:Input S
:If int(N/2)=N/2
:Then
:((N/2-1)!)^2*2^(N-2)/(N-2)!*
 √ (2/Π/(N-1)) →C
:Else
: √ (2Π/(N-1))*(N-1)!/((N-1)/2)!
 /((N-3)/2)!/2^(N-1) →C
:End
:"C*S"→Y1
:"C*S+3S√ (1-C^2)"→Y2
:"max(C*S-3S√ (1-C^2),0)" →Y3
:If dim(L2)>0
:Then
:seq(I,I,1,dim(L2))→L3
:0→Xmin
:dim(L2)+1→Xmax
:Xmax→Xscl
:min(min(L2),C*S-4S√ (1-C^2))
 →Ymin
:max(max(L2),C*S+4S√ (1-C^2))
 →Ymax
:Ymax-Ymin→Yscl
:Plot1(xyLine,L3,L2, ·)
:End
:ClrHome
:Output(1,4,"S LIMITS")
:Lbl 4
:PlotsOff
:PlotsOn 1
:AxesOff
:Output(3,1,"UCL")
:Output(3,5,Y2)
:Output(4,1," CL")
:Output(4,5,Y1)
:Output(5,1,"LCL")
:Output(5,5,Y3)
:Lbl 3
:Stop
```

Example 17.1 Milling Hydraulic Systems. The width of a slot cut by a milling machine is important for the proper functioning of a hydraulic system for large tractors. The manufacturer checks the control of the milling process by measuring a sample of five consecutive items during each hour's production. The target width of a slot cut by the milling machine is μ = 0.8750 in. with a target standard deviation of 0.0012 in. What are the centerline and control limits for an *s* chart? For an $\overline{x}$ chart?

Solution. Bring up the **CONTRL** program and enter either **1** or **2** for the desired variable's control limits. Then enter the sample size of **5** and the target parameters. The centerline and control limits are then displayed.



The control limits for the standard deviation are 0 (the standard deviation cannot be negative) and 0.00236. The control limits for the mean are 0.8734 and 0.8766.

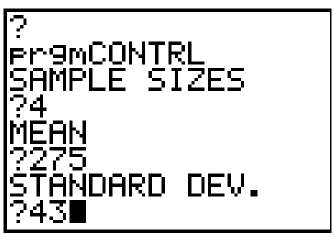
If we have the values of $\overline{x}$ and *s* from various samples, then we also can use the program to display the control charts. To do so, always enter the values of $\overline{x}$ into list `L1` and enter the values of *s* into list `L2`. After executing the **CONTRL** program, press ♦ to see the control chart.

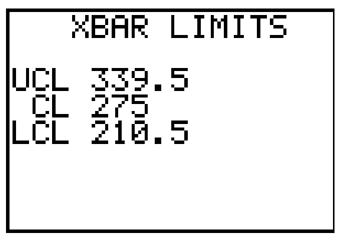
Example 17.2 Computer Monitors. A manufacturer of computer monitors must control the tension on the mesh of fine vertical wires that lies behind the surface of the viewing screen. Too much tension will tear the mesh, and too little will allow wrinkles. The following mesh tension data gives the sample mean and sample deviation from 20 different samples of size 4.

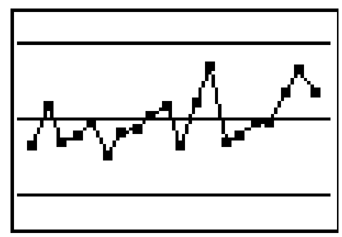
Sample mean	Standard deviation	Sample mean	Standard deviation
253.4	21.8	253.2	16.3
285.4	33.0	287.9	79.7
255.3	45.7	319.5	27.1
260.8	34.4	256.8	21.0
272.7	42.5	261.8	33.0
245.2	42.8	271.5	32.7
265.2	17.0	272.9	25.6
265.6	15.0	297.6	36.5
278.5	44.9	315.7	40.7
285.4	42.5	296.9	38.8

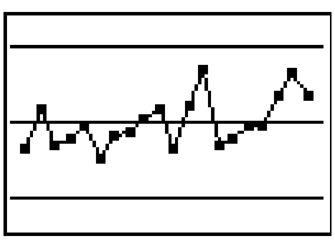
The target mean tension is μ = 275 mV with a target standard deviation of 43 mV. Find the centerline and control limits for $\bar{x}$ and for s. Graph the control charts for each.

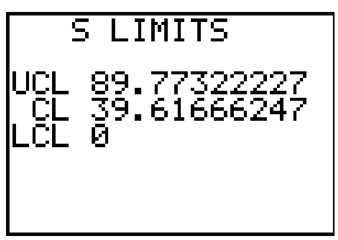
Solution. Enter the sample means into list L1 and the standard deviations into list L2. Then execute the **CONTRL** program for the desired variable to obtain the control limits, and press σ to see the control chart. If desired, press ρ and scroll right to see the individual points.

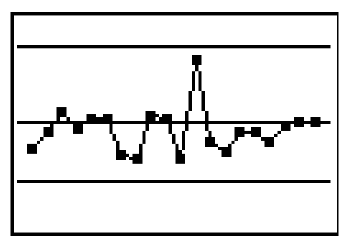












Since both charts are within the control limits, this process is in control; we see only common cause variation around the desired values of both the mean and standard deviation.

17.2 Using Control Charts

We now provide a variation of the **CONTRL** program that will compute the upper and lower control limits and graph the control charts for $\bar{x}$ or s based on past data.

In the program that follows, we note that data sets of *equal* sizes must be entered into lists `L1` and `L2` in order to obtain the control limits for $\bar{x}$. However, only the values of the sample deviations need to be entered into list `L2` in order to compute the control limits for s.

The CONTRL2 Program

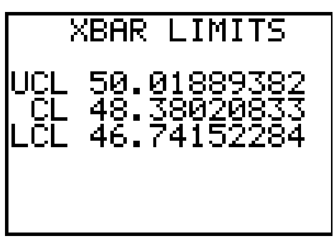
```
Program:CONTRL2
:Disp "SAMPLE SIZES"
:Input N
:If int(N/2)=N/2
:Then
:((N/2-1)!)^2*2^(N-2)/(N-2)!
 *√(2/Π/(N-1))↵C
:Else
: √ (2Π/(N-1))*(N-1)!/((N-1)/2)!
 /((N-3)/2)!/2^(N-1) ↵C
:End
:Disp "1=XBAR, 2=S"
:Input W
:If W=1
:Then
:2-Var Stats L1,L2
:⌡↵M
:□/C↵S
:"M"↵Y1
:"M+3*S/√ (N)" ↵Y2
:"M-3*S/√ (N)" ↵Y3
:seq(I,I,1,dim(L1))↵L3
:0↵Xmin
:dim(L1)+1↵Xmax
:Xmax↵Xscl
:min(min(L1),M-4*S/√ (N)) ↵Ymin
:max(max(L2),M+4*S/√ (N)) ↵Ymax
:Ymax-Ymin↵Yscl
:Plot1(xyLine,L3,L1,⎝)
:ClrHome
:Output(1,4,"XBAR LIMITS")
:Goto 4
:Else
:1-Var Stats L2
: ⌡/C↵S
:"C*S"↵Y1
:"C*S+3S√ (1-C^2)"↵Y2
:"max(C*S-3S√ (1-C^2),0)" ↵Y3
:seq(I,I,1,dim(L2))↵L3
:0↵Xmin
:dim(L2)+1↵Xmax
:Xmax↵Xscl
:min(min(L2),C*S-4S√ (1-C^2))
 ↵Ymin
:max(max(L2),C*S+4S√ (1-C^2))
 ↵Ymax
:Ymax-Ymin↵Yscl
:Plot1(xyLine,L3,L2, ⎝)
:ClrHome
:Output(1,4,"S LIMITS")
:End
:Lbl 4
:PlotsOff
:PlotsOn 1
:AxesOff
:Output(3,1,"UCL")
:Output(3,5,Y2)
:Output(4,1," CL")
:Output(4,5,Y`)
:Output(5,1,"LCL")
:Output(5,5,Y3)
:Stop
```

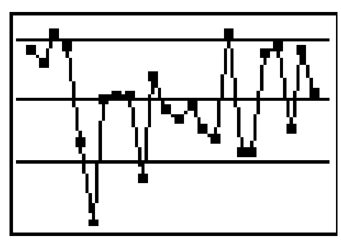
Example 17.3 Measuring Viscosity. The viscosity of a material is its resistance to flow under stress. Viscosity is a critical characteristic of rubber and rubberlike compounds called elastomers, which are used in many consumer products. A specialty chemical company is beginning production of an elastomer that is supposed to have viscosity 45 ± 5 "Mooney Units." The following data gives the mean and standard deviation of elastomer viscosity from samples of size 4 from the first 24 shifts as production begins.

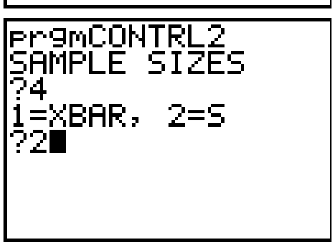
Sample	$\bar{x}$	s	**Sample**	$\bar{x}$	s
1	49.750	2.684	13	47.875	1.118
2	49.375	0.895	14	48.250	0.895
3	50.250	0.895	15	47.625	0.671
4	49.875	1.118	16	47.375	0.671
5	47.250	0.671	17	50.250	1.566
6	45.000	2.684	18	47.000	0.895
7	48.375	0.671	19	47.000	0.447
8	48.500	0.447	20	49.625	1.118
9	48.500	0.447	21	49.875	0.447
10	46.250	1.566	22	47.625	1.118
11	49.000	0.895	23	49.750	0.671
12	48.125	0.671	24	48.625	0.895

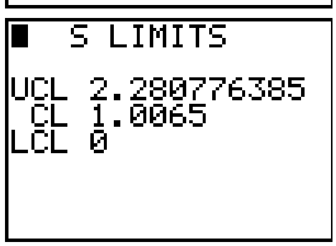
(a) Find the centerline and control limits for $\bar{x}$ and for s based on this past data. Graph the control charts for each.
(b) Remove the two values of s that are out of control and reevaluate the control limits for s based on the remaining data.
(c) Remove the corresponding two values of $\bar{x}$ from the samples that were removed in (b) and reevaluate the control limits for $\bar{x}$ based on the remaining data.

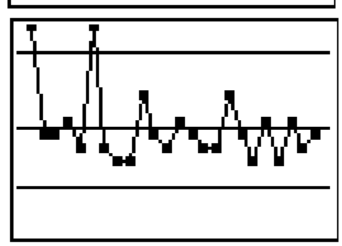
Solution. (a) First, enter the sample means into list `L1` and the sample deviations into list `L2`, then bring up the **CONTRL2** program. Enter 4 for the sample size, then enter **1** when prompted to calculate the control limits for $\bar{x}$ based on this past data. Press σ to see the control chart. Then reexecute the program, but enter **2** when prompted to calculate the control limits for s.





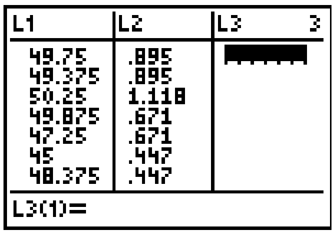


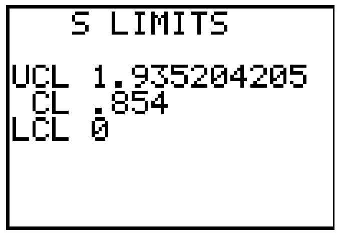


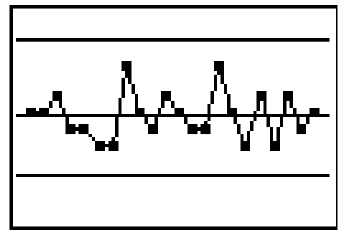


We notice that there were four samples out of control on the $\overline{x}$ chart: the third, sixth, tenth, and seventeenth. There were two samples (the first and sixth) out of control on the *s* chart.

(b) Because the upper control limit for *s* is 2.28, the values of *s* from the first and sixth samples are out of control. We now delete these two values (both 2.684) from list L2 and reexecute the **CONTRL2** program.

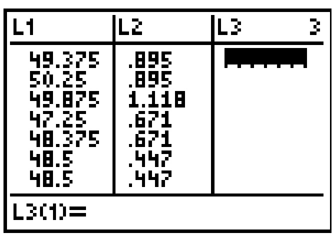


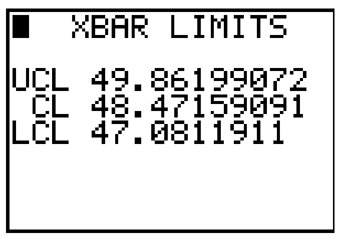


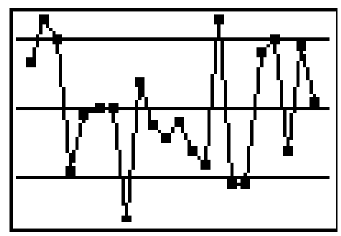


In terms of the standard deviation, this plot is much more desirable.

(c) We now delete the values of $\overline{x}$ from the first and sixth samples from list L1, then reexecute option 1 of the **CONTRL2** program to find the new control limits for $\overline{x}$.







We still see six points outside the control limits, which indicates that this process has not yet reached a controlled state where only common sources of variation are in play.

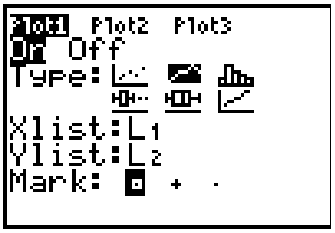
Example 17.4 Joe's Weight. Joe's weight has been stable for several years. An injury keeps Joe away from the gym for several months. The following data give Joe's weight, measured once each week, for the first 16 weeks back at the gym after his injury.

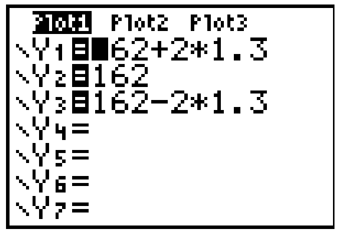
Week	1	2	3	4	5	6	7	8
Weight	168.7	167.6	165.8	167.5	165.3	163.4	163.0	165.5

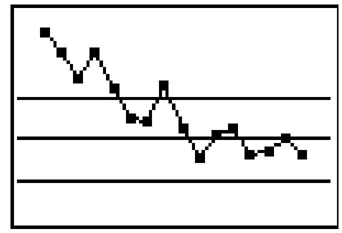
Week	9	10	11	12	13	14	15	16
Weight	162.6	160.8	162.3	162.7	160.9	161.3	162.1	161.0

Joe has a target of μ = 162 pounds for his weight. The short-term variation is estimated to be about σ = 1.3 pounds. Make a control chart for his measurements using control limits $\mu \pm 2\sigma$.

Solution. Simply make a time plot of the measurements together with graphs of the lines $y = \mu$ and $y = \mu \pm 2\sigma$. To do so, enter the weeks into list L1 and the weights into list L2, then adjust the STAT PLOT settings for a scatterplot of L1 versus L2. Next, enter the functions Y1 = 162+2*1.3, Y2 = 162, and Y3 = 162-2*1.3 into the o screen, and press ❑✹ to display the graph; adjust the WINDOW settings if needed so that the X range includes all weeks and the Y range includes all weights as well as the upper and lower control limits.







After his injury, Joe's weight was out of control; but it is getting into control.

17.3 Process Capability Indexes

In this section, we provide a program to compute the capability indexes. The program can be used when the parameters are given or when data is given in a list.

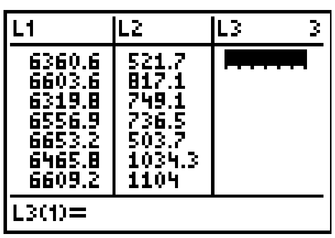
The CAPIND Program

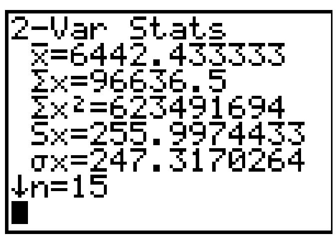
```
PROGRAM:CAPIND
:Disp "LSL"
:Input L
:Disp "USL"
:Input U
:Disp "1=STATS, 2=LIST"
:Input W
:If W=1
:Then
:Disp "MEAN"
:Input M
:Disp "STANDARD DEV."
:Input S
:Goto 1
:Else
:2-Var Stats L1,L2
:⌡↵M
:□↵S
:End
:Lbl 1
:(U-L)/(6S)↵C
:If M≤L or M≥U
:Then
:0↵D
:Else
:min(M-L,U-M)/(3S) ↵D
:End
:ClrHome
:Output(1,4,"CAP.INDEXES")
:Output(3,2,"Cp")
:Output(3,5,C)
:Output(5,1,"Cpk")
:Output(5,5,D)
:Stop
```

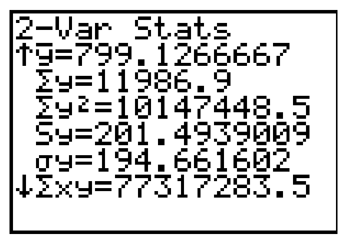
Example 17.5 Hospital Losses. Below are data on a hospital's losses for 120 DRG 209 (major joint replacement) patients collected as 15 monthly samples of eight patients each. The hospital has determined that suitable specification limits for its loss in treating one such patient are LSL = \$4000 and USL = \$8000.
(a) Estimate the percent of losses that meet the specifications.
(b) Estimate C_p and C_{pk}.

Sample	$\bar{x}$	s	**Sample**	$\bar{x}$	s
1	6360.6	521.7	9	6479.0	704.7
2	6603.6	817.1	10	6175.1	690.5
3	6319.8	749.1	11	6132.4	1128.6
4	6556.9	736.5	12	6237.9	596.6
5	6653.2	503.7	13	6828.0	879.8
6	6465.8	1034.3	14	5925.5	667.8
7	6609.2	1104.0	15	6838.9	819.5
8	6450.6	1033.0			

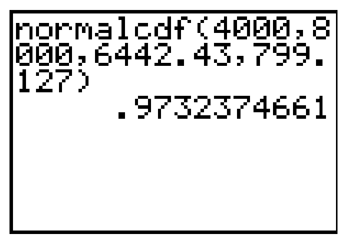
Solution. First, enter the sample means into list `L1` and the sample deviations into list `L2`. Then enter the command `2-Var Stats L1,L2` to compute $\bar{\bar{x}}$ and s. We see that $\bar{\bar{x}}$ = 6442.43 and $\bar{s}$ = 799.127 (from the } output).



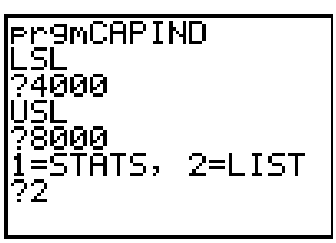


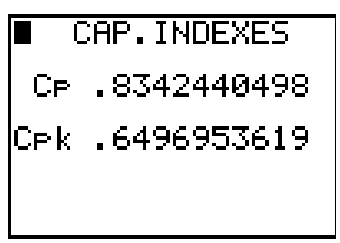


(a) Next, use the `normalcdf(` command from the `DISTR` menu to compute $P(4000 \le Z \le 8000)$ for $X \sim N(6442.43, 799.127)$. We find that about 97.32% of losses meet the specifications.



(b) Last, bring up the **CAPIND** program. Enter the LSL of 4000 and the USL of 8000, then enter **2** when prompted for **LIST** to output the capability index approximations.

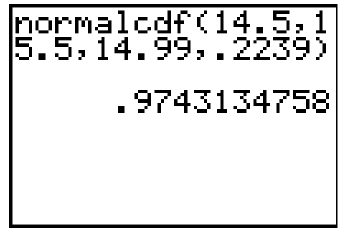




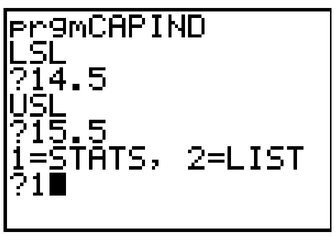
Example 17.6 Measuring Clips. The dimension of the opening of a clip has specifications 15 ± 0.5 millimeters. The production of the clip is monitored by $\bar{x}$ and s charts based on samples of five consecutive clips each hour. The 200 individual measurements from the past week's 40 samples have $\bar{x}$ = 14.99 mm and s = 0.2239 mm.

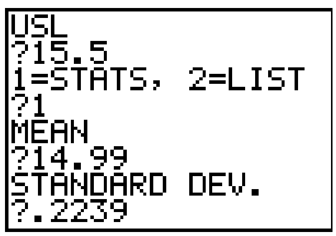
(a) What percent of clip openings will meet specifications if the process remains in its current state? (b) Estimate the capability index C_{pk}.

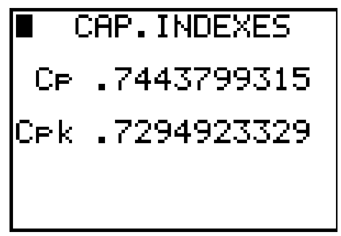
Solution. (a) We must compute $P(15-.5 \le X \le 15+.5)$ for $X \sim N$(14.99, .2239). Using the `normalcdf(` command, we find that about 97.43% of clip openings will meet specifications if production remains in its current state.



(b) To estimate C_{pk}, we shall use the **CAPIND** program. Bring up the program and enter the LSL of 14.5 and the USL of 15.5, then enter **1** when prompted for **STATS**. Enter the mean and standard deviation to receive the output. We find that C_{pk} = 0.7295.







17.4 Control Charts for Sample Proportions

We conclude this chapter with a program to compute the control limits for sample proportions. The program can be used with summary statistics or when a list of data is given.

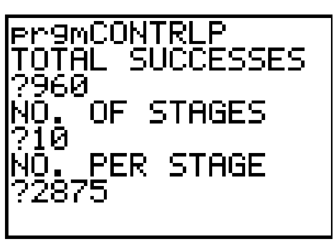
If data are in lists, the number of "successes" should be in `L1`, and the number of "trials" should be in `L2`.

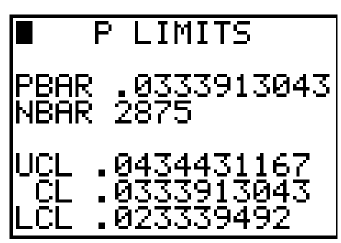
The CONTRLP Program

```
PROGRAM:CONTRLP
:Menu("CONTRLP","STATS",1,"LIST",2,"
 QUIT",3)
:Lbl 1
:Disp "TOTAL SUCCESSES"
:Input T
:Disp "NO. OF STAGES"
:Input M
:Disp "NO. PER STAGE"
:Input N
:T/(M*N)↵P
:Goto 4
:Lbl 2
:1-Var Stats L2
:⌡↵N
:sum(seq(L/(I),I,1,dim(L1)))/sum(
 seq(L2(I),I,1,dim(L2)))↵P
:L1/L2↵L3
:seq(I,I,1,dim(L1))↵L4
:dim(L1)+1↵Xmax
:Xmax↵Xscl
:min(Y3,min(L3))-.01↵Ymin
:max(Y2,max(L3))+.01↵Ymax
:1↵Yscl
:Plot1(xyLine,L4,L3,⍀)
:PlotsOff
:PlotsOn 1
:AxesOff
:Lbl 4
:"P"↵Y1
:"min(P+3*√ (P(1-P)/N),1)" ↵Y2
:"max(P-3*√ (P(1-P)/N),0)" ↵Y3
:0↵Xmin
:ClrHome
:Output(1,5,"P LIMITS")
:Output(3,1,"PBAR")
:Output(3,6,P)
:Output(4,1,"NBAR")
:Output(4,6,N)
:Output(6,1,"UCL")
:Output(6,5,Y2)
:Output(7,1," CL")
:Output(7,5,Y1)
:Output(8,1,"LCL")
:Output(8,5,Y3)
:Lbl 3
:Stop
```

Example 17.7 Unpaid Invoices. The controller's office of a corporation is concerned that invoices that remain unpaid after 30 days are damaging relations with vendors. To assess the magnitude of the problem, a manager searched payment records for invoices that have arrived in the last 10 months. In that period, an average of 2875 invoices per month have been received, with only 960 remaining unpaid after 30 days. Find $\overline{p}$. Give the centerline and control limits for a *p* chart.

Solution. Bring up the **CONTRLP** program and press **1** for **STATS**. Enter the values of **960** for total "successes," **10** for the number of "stages," and **2875** for the number per stage. We find that $\overline{p}$ = 0.0334 with an LCL of 0.02334 and a UCL of 0.04344.



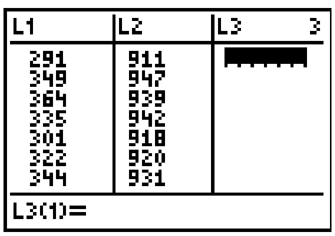


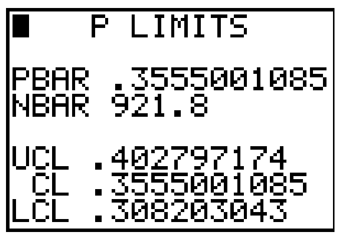
Example 17.8 School Absenteeism. Here are data on the total number of absentees among eighth-graders with three or more unexcused absences at an urban school district. Because the total number of students varies each month, these totals are also given for each month.

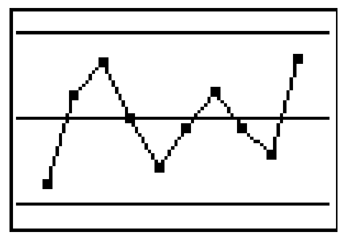
Month	**Sep.**	**Oct.**	**Nov.**	**Dec.**	**Jan.**	**Feb.**	**Mar.**	**Apr.**	**May**	**Jun.**
Students	911	947	939	942	918	920	931	925	902	883
Absent	291	349	364	335	301	322	344	324	303	344

(a) Find $\overline{p}$ and $\overline{n}$. (b) Make a *p* chart using control limits based on $\overline{n}$ students each month.

Solution. (a) First, enter the number of absentees ("successes") for each month into list L1 and the number of students into list L2. Next, execute option 2 of the **CONTRLP** program. We find that $\overline{p}$ = 0.3555 and $\overline{n}$ = 921.8. (b) After executing the program, the individual monthly proportions are stored in list L3. Press σ to see the *p* chart that has an LCL of 0.3082 and a UCL of 0.4028.

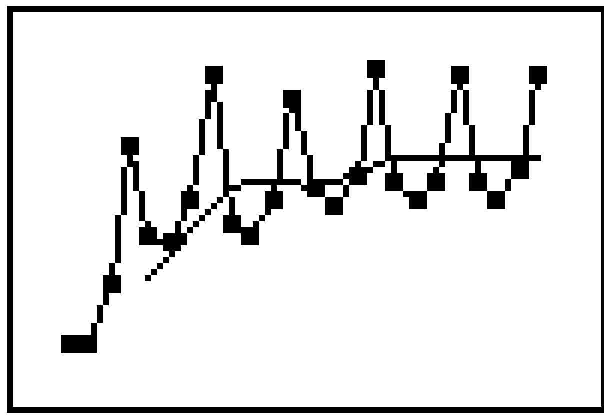






CHAPTER

18



Time Series Forecasting

Introduction

A time series is a sequence of observations on a single variable at equally spaced intervals.

In this chapter, we examine some basic ideas in time series forecasting and modeling. The most basic idea is to use a model to describe a trend (if any), then to add additional terms to describe seasons (or cycles within years). We also examine models in which a new observation is related to a prior ones (autoregressive models) as well as moving average models.

18.1 Trends and Seasons

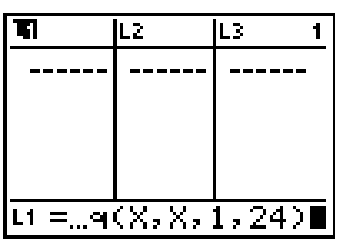
The most basic component of a time series is assessing whether or not there is a trend (systematic rise or fall) and whether there is some aspect that repeats regularly (a cycle, or seasonal component).

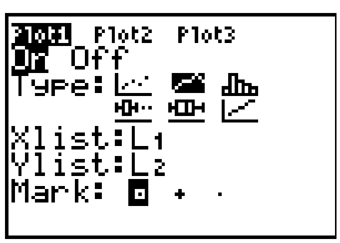
Example 18.1 JCPenney Sales. The table below contains retail sales for JCPenney in millions of dollars beginning with the first quarter of 1996 and ending with the fourth quarter of 2001.

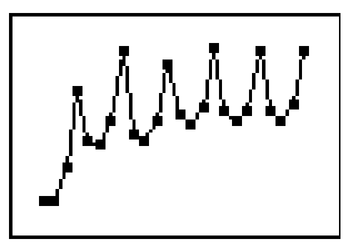
Year-quarter	Sales	Year-quarter	Sales
1996-1^{st}	4452	1999-1^{st}	7339
1996-2^{nd}	4507	1999-2^{nd}	7104
1996-3^{rd}	5537	1999-3^{rd}	7639
1996-4^{th}	8157	1999-4^{th}	9661
1997-1^{st}	6481	2000-1^{st}	7528
1997-2^{nd}	6420	2000-2^{nd}	7207
1997-3^{rd}	7208	2000-3^{rd}	7538
1997-4^{th}	9509	2000-4^{th}	9573
1998-1^{st}	6755	2001-1^{st}	7522
1998-2^{nd}	6483	2001-2^{nd}	7211
1998-3^{rd}	7129	2001-3^{rd}	7729
1998-4^{th}	9072	2001-4^{th}	9542

(a) Make a time plot of the data. Be sure to connect the points in your plot to highlight patterns.
(b) Is there an obvious trend in JCPenney quarterly sales? If so, is the trend positive or negative?
(c) Is there an obvious repeating pattern in this data? If so, clearly describe the repeating pattern.

Solution. (a) First, enter the data into two variables: one (in `L1`) for the quarter number (from 1 to 24; this is most easily done using the `List`, `Ops`, `seq(` command as shown below) and another for the sales amounts (in `L2`). Define the plot as a connected scatter plot, then press θ→ to display it.





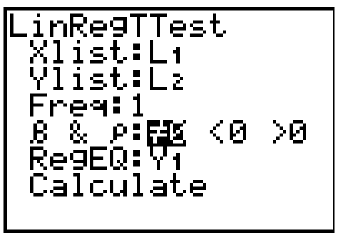


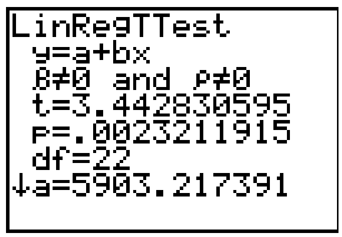
(b) There does appear to be some trend in this plot; however if we eliminate the first couple of years it is not as obvious.

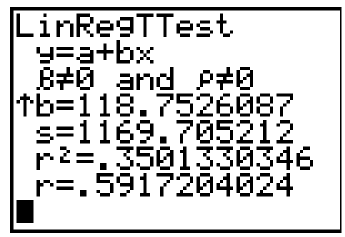
(c) There are definitely repeating patterns in this set of data. The fourth quarter (including Christmas) always has the highest sales; the second quarter always has the lowest sales. Third quarter is generally somewhat higher than the first quarter (back-to-school sales?).

Example 18.2 JCPenney Trends. Find any trend in the JCPenney data using $x = 1$ to correspond with the first quarter of 1996, $x = 2$ to the second quarter of 1996, and so forth. Interpret the slope and intercept in the model.

Solution. Since we already have the data entered, we need to compute the regression line where the dependent variable is Sales (in `L2`) and the independent variable is the quarter number (in `L1`). We can use either `Stat`, `Calc`, `Linreg(a+bx)` or `Stat`, `Tests`, `LinRegTTest` to do this.

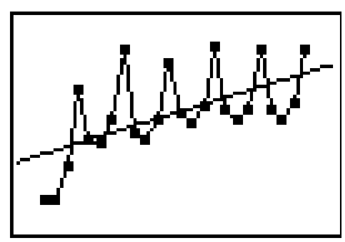






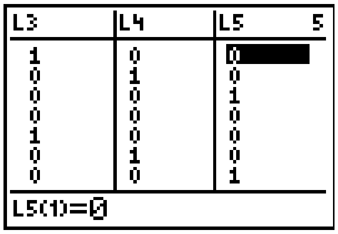
We find the equation *Sales* = 5903.217 + 118.753**Quarter_num*. With a *t*-statistic of 3.443 and *p*-value of 0.002, this trend is significantly nonzero. The slope indicates that JCPenney sales increase (on average) 118.753 million dollars per quarter. The intercept says estimated sales for quarter 0 (that would be fourth quarter 1995) are 5903.217 million dollars.

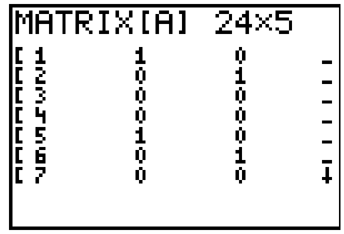
To see the trend line on the time plot, press σ. While significant, the trend does not do a good job of describing the major features of this series; also, most of the trend may be due to the two low values for the first two quarters.



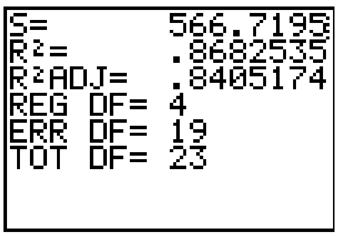
Example 18.3 JCPenney Cycles. Since sales seem to have an annual cycle, add indicator variables for the quarters to the trend-only model fit in Example 18.2. Compare the estimated intercept of this model with the intercept found in Example 18.2. Given the patterns of seasonal variation, which appears to be the better estimate?

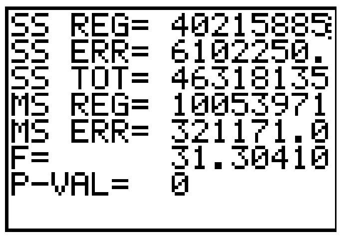
Solution. This becomes a multiple linear regression model. From our work in Chapter 11, we know that we need data in a matrix in order to use program **MULREG** to compute the regression. The actual sales data (the dependent variable) must be in the last column of matrix `[A]`. We will want four predictor variables (the quarter number, and indicator variables for quarters 1, 2, and 3 of each year). If we added a fourth indicator variable we would have linearly dependent columns (because of the intercept) and the model could not be solved. The easiest way to create matrix [A] is to enter the three indicator variable lists, and then use List ▸Matrix from the List, Ops menu as shown below.

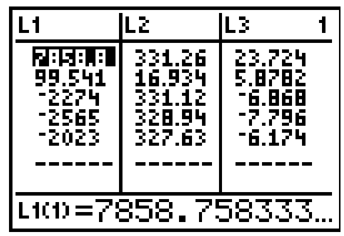




Now, execute program **MULREG** to fit the model (some of the output is omitted).







Our model is now (use the arrow keys to get more decimal places than display on the EDIT screen)

$$Sales = 7858.758 + 99.541 Quarter\_num - 2274.21 x1 - 2564.585 x2 - 2022.792 x3.$$

Since the indicator variables only come into play for the selected quarters (as a 1 or 0, which means these values affect the intercept), this really means we have the following set of equations:

$$4thQtrSales = 7858.738 + 99.541 Quarter\_num$$
$$1stQtrSales = 5584.448 + 99.541 Quarter\_num$$
$$2ndQtrSales = 5294.173 + 99.541 Quarter\_num$$
$$3rdQtrSales = 5835.966 + 99.541 Quarter\_num$$

Our trend is now that sales increase (on average) 99.541 million dollars per quarter. Since the intercept is still month 0 (fourth quarter 1995), this model forecasts 7858.738 million for that quarter. This is much more reasonable given that fourth quarter sales are always the highest for the year.

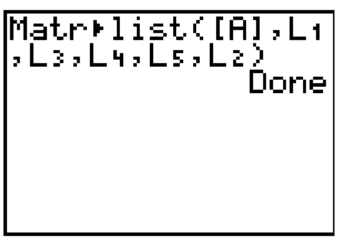
18.2 Time Series Models

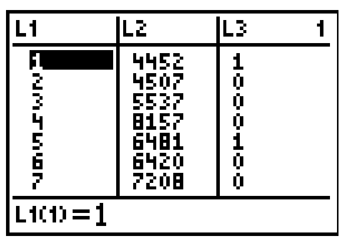
Time series models use past values to predict future values of the series. There are several ways to model these, and we have already looked at regression models. There are several other methods whose study can consume entire statistics courses on their own. We will look at two time series models – the autoregressive model and the moving average model.

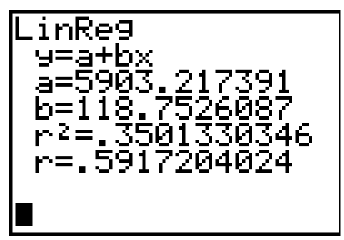
First-order autoregressive models use the equation $y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t$ as a model. This implies that the value one time period prior is the most useful in forecasting the next value. One way to look for an autoregressive relationship is to plot lagged residuals; that is, plot $(\varepsilon_1, \varepsilon_2), (\varepsilon_2, \varepsilon_3), ..., (\varepsilon_{n-1}, \varepsilon_n)$.

Example 18.4 JCPenney Trend-Only Residuals. Return to the linear trend only model of Example 18.2 and create a lagged residual plot to examine the data for autocorrelation. Calculate the correlation in the lagged residuals.

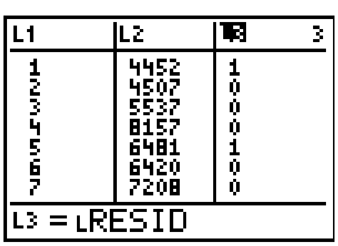
Solution. If you've been working along with this manual, the JCPenney sales data are now in matrix [A] and the original data have been overwritten with the coefficients of the multiple regression model. To reclaim the data (and not have to reenter it), use the command Matrix|List from the List, Ops menu. In order for us to have the lists we want easily accessible, the actual sales data (in the last column) are specified to end up in L2. Recompute the regression; the calculator automatically saves the residuals in a list named RESID.

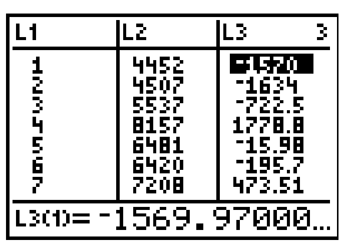


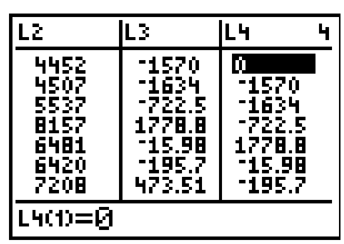




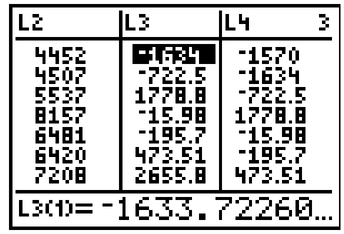
To use this list, move it to L3. Highlight the list name, and locate list RESID on the List, Names screen. Press ÷ to copy the list into L3. To lag the residuals, copy L3 into L4 (use L4=L3 with the L4 name highlighted on the EDIT screen), then ψ{ to insert a new observation (a 0) as the first entry in L4.



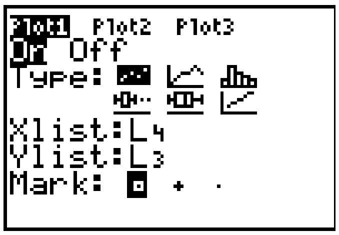


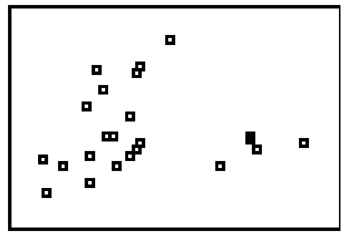


We have one more list manipulation to do before we can create the plot. Adding the new observation to lag the residuals makes the lists unequal in length; also that inserted 0 is not real data. Since calculators must have lists of equal length, delete the first observations in both `L3` and `L4`, and delete the last observation in `L4`.



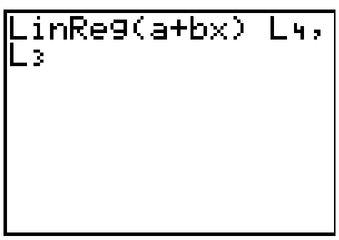
To plot the residuals, define a scatter plot using the lagged residuals (in `L4`) on the x- axis and the original residuals (in `L3`) on the y-axis.

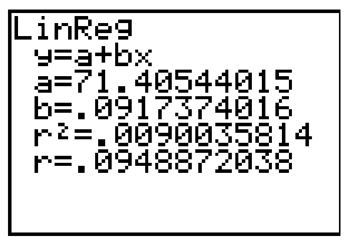




There appear to be two sets of points here: one with a significant amount of correlation in the left part of this plot, and an interesting group with large positive residuals. These are due to fourth quarter sales.

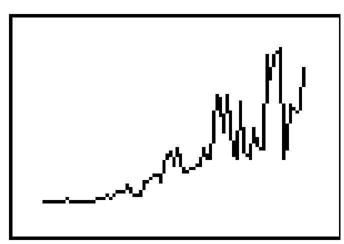
To compute the correlation, simply compute a regression between these two sets of residuals.





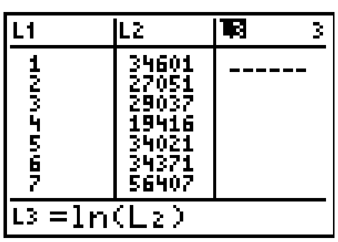
The correlation in these residuals is not significant because the given value of r is 0.095; the linear relationship explains only $r^2 = 0.009$ of the variation in the original residuals. If you use `LinRegTTest`, you will find the p-value for this relationship is 0.667 which is large. This result is most likely due to the clump of large positive residuals; remember, even a single point can make a correlation seem non-significant.

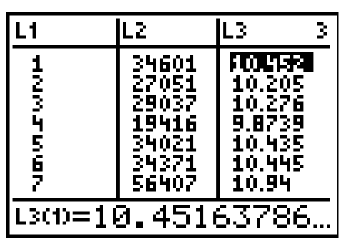
Example 18.5 Autoregressive DVD Sales. The popularity of the DVD format has exploded in the relatively short period of time since its introduction in March 1997. The Consumer Electronics Association tracks monthly sales of DVD players. In the table on the next page are data on DVD sales from April 1997 through June 2002 (63 months). A time plot of the data is shown.

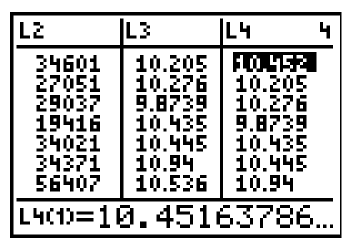


Date	Units sold	Date	Units sold	Date	Units sold
4-97	34601	1-99	125536	10-00	1236658
5-97	27051	2-99	109399	11-00	866507
6-97	29037	3-99	123466	12-00	1303091
7-97	19416	4-99	269107	1-01	572031
8-97	34021	5-99	279756	2-01	555856
9-97	34371	6-99	326668	3-01	1207489
10-97	56407	7-99	325151	4-01	631353
11-97	37657	8-99	260225	5-01	523225
12-97	42575	9-99	501501	6-01	920839
1-98	34027	10-99	603048	7-01	693013
2-98	34236	11-99	449242	8-01	673926
3-98	38336	12-99	646290	9-01	1768821
4-98	42889	1-00	370031	10-01	1516211
5-98	47805	2-00	401035	11-01	1781048
6-98	79044	3-00	412559	12-01	1862772
7-98	84709	4-00	409192	1-02	542698
8-98	81170	5-00	453435	2-02	736118
9-98	113558	6-00	654687	3-02	1162568
10-98	163074	7-00	537453	4-02	1090767
11-98	136908	8-00	557617	5-02	1171984
12-98	233505	9-00	1296280	6-02	1617098

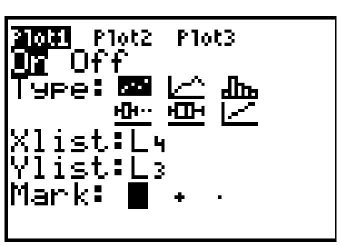
There is very little increase in sales at the beginning; then sales skyrocket and also exhibit considerable variation. As we will see, a scatter plot of lagged ln(DVD sales) versus ln(DVD Sales) is very linear. To first compute ln(DVD sales), with sales in L2, move the cursor to highlight the L3 list name. Enter the command to read L3=ln(L2) as shown below. Press ÷ to perform the computation. As we did to lag the residuals in Example 18.4, copy L3 to L4 and insert a 0 at the beginning, then delete the first observation from both L3 and L4, and the last observation from L4. When finished, L3 and L4 should have equal lengths; the first few rows are shown below.

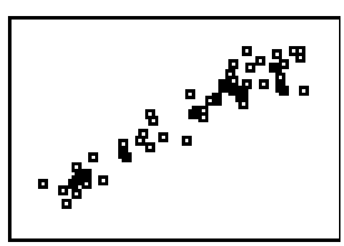




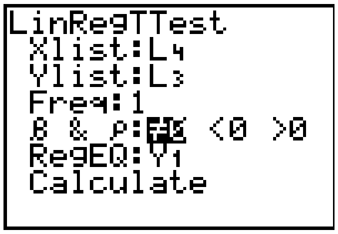


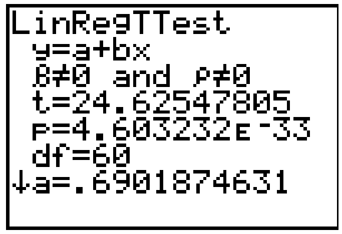
Now, define a scatterplot with L4 on the x axis and L3 on the y axis.

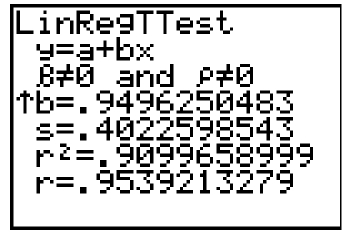




Since this plot is so linear, an autoregressive model appears reasonable. We will fit a linear model to the log sales data using the lagged data as the predictor.







We notice the *t*-statistic for this regression is 24.625 with a *p*-value of essentially 0. This gives an equation of $\ln(DVDsales)_t = .690 + .950\ln(DVDsales)_{t-1}$.

Suppose we want to predict sales for July 2002 based on this model. June 2002 sales were 1617098 units. Putting this into the equation, we have

$$\begin{aligned}\ln(July) &= .690 + .950\ln(1617098)\\ &= .690+.950*14.2961\\ &= 14.271\\ \text{July} &= e^{14.271} = 1577410\end{aligned}$$

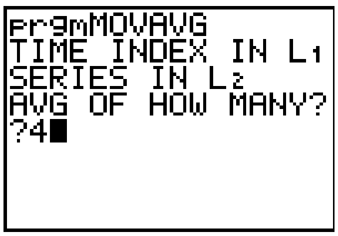
Moving average models forecast y_t as the average of the preceding *k* observations, in other words, $y_t = \dfrac{y_{t-1} + y_{t-2} + ... + y_{t-k}}{k}$. These models smooth out some of the irregular noise in a typical time series. Below, we provide a short program to find the moving averages. The time index must be in `L1`, and the series in `L2`. When complete, `L3` has the indices for the averaged values and `L4` has the averaged values. The program purposefully puts 0s in the first *N* entries in these lists so that one can page through the list and see the moving average estimate alongside the original value; to graph the moving average series along with the original, delete the 0 entries before graphing.

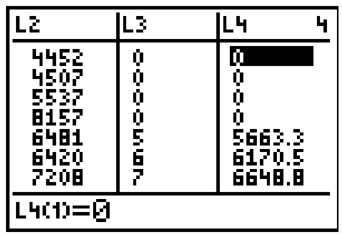
Program MOVAVG

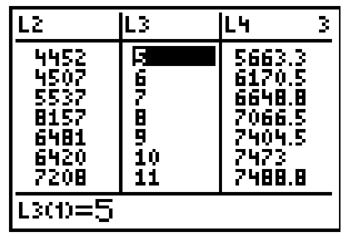
```
Disp "TIME INDEX IN L1"
Disp "SERIES IN L2"
Disp "AVG OF HOW MANY?"
Input N
Clrlist L3, L4
For(I,1,N)
0↵L3(I)
0↵L4(I)
End
dim(L2)↵M
For(K,N+1,M)
K-N↵J
sum(seq(L2(L),L,J,J+N-1)/N↵L4(K)
L1(K) ↵L3(K)
End
Stop
```

Example 18.6 JCPenney Moving Averages. JCPenney sales have an annual periodicity. In every year, the fourth quarter sales are highest, and second quarter sales are lowest. Find the moving average predictor using a span of 4. Graph the moving average model on the original time series.

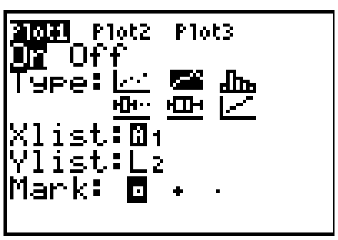
Solution. With the data entered, execute program **MOVAVG**. We want moving averages of length 4. When the program displays the Done message, go to the Statistics Editor. Here, we can see that the first four entries in L3 and L4 are 0s. The next entries are the first two moving averages; 5663.3 = (4453+4507+5537+8157)/4. (More exact values can be found by moving the cursor to the desired entry.) To prepare to graph these series, delete the first four entries in L3 and L4.

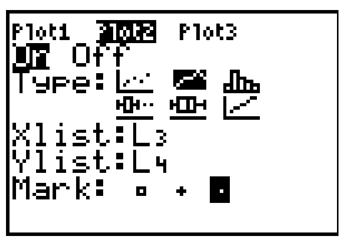


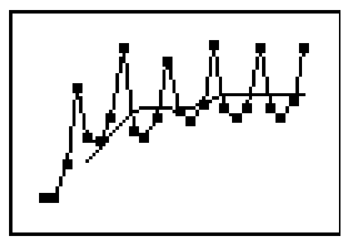




To graph both series at once, define two time plots. Then press ❑✹ to display them. It makes no difference what symbols are plotted for each data point, but we recommend using different ones to help keep the two series differentiated.







Notice that this graph shows the moving average "flattening out" through 2000 and 2001. The country was in a mild recession then, so one would not have expected increasing sales.

Index of Programs

Programs can be downloaded at
http://math.georgiasouthern.edu/~phumphre/TIpgms/ or from your text's portal.

ANOVA1 (page 134) displays the overall sample mean, the pooled sample deviation, the mean square error for groups MSG, the mean square for error MSE, the R^2 coefficient, the F-statistic, and the p-value of the ANOVA test for equality of means when data is entered as summary statistics. Before executing the program, enter the sample sizes into list `L1`, the sample means into list `L2`, and the sample deviations into list `L3`.

ANOVA2W (page 139) displays the p-values for two-way analysis of variance. For one observation per cell, enter the data into matrix **[A]** before executing the program. For C observations per cell, enter the means into matrix **[A]** and the standard deviations into matrix **[B]**. The program also stores the marginal means for the rows and columns into lists `L2` and `L4`. The overall mean is stored as the first value in list `L5`. The remainder of `L5` is the values SSA, SSB, and SSE.

BAYES (page 60) computes the total probability $P(C)$ and conditional probabilities associated with Bayes's rule. Before executing the program, enter values for $P(A_i)$ into list `L1` and the conditional probabilities $P(C \mid A_i)$ into list `L2`. The program displays $P(C)$, stores $P(C \cap A_i)$ in list `L3`, stores $P(A_i \mid C)$ in list `L4`, stores $P(A_i \mid C')$ in list `L5`, and stores $P(C \mid A'_i)$ in list `L6`.

BOOT (page 145) performs resampling on a random sample in list `L1`. If a bootstrap confidence interval for the statistic is desired, enter **1** for `CONF. INTERVAL?` when prompted; otherwise, enter **0**. The program takes resamples from the entered random sample and enters their means into list `L2`. The mean of all resamples, the bootstrap standard error, and the confidence interval (if specified) are displayed.

BOOTCORR (page 153) performs the bootstrap procedure on the correlation coefficient or the regression slope for paired sample data that has been entered into lists `L1` and `L2`. When prompted, enter **1** if you want to bootstrap the correlation coefficient, or enter **2** if you want to bootstrap the regression slope. The resampled statistics are stored in list `L3`. The statistic of the original sample data is displayed along with the bootstrap standard error and the confidence interval.

BOOTPAIR (page 150) computes a bootstrap t confidence interval for the difference in means based on random samples that have been entered into lists `L1` and `L2`. The resampled differences in mean are stored in list `L3`. The difference of the original sample averages is displayed along with the bootstrap standard error and the confidence interval.

BOOTTEST (page 156) performs a permutation test for the difference in means. Before executing, enter data from the first population into list `L1` and enter data from the second population into list `L2`. When prompted, enter **1**, **2**, or **3** to designate the desired alternative. The resampled differences in permuted mean are ordered and then stored in list `L3`. The program displays the difference between the original sample means and the p-value.

BOOTTRIM (page 149) computes a bootstrap t confidence interval for a trimmed mean on a random sample that has been entered into list `L1`. When prompted, enter the desired number of resamples, the decimal amount to be trimmed at each end, and the desired confidence level. The program takes resamples from the entered random sample and enters their trimmed means into list L2. The trimmed mean of the original sample, the trimmed bootstrap standard error, and the confidence interval are displayed.

BTPRTEST (page 158) performs a permutation test for either the difference in paired means or for the correlation. Before executing, enter the data set into lists L1 and L2. When prompted, enter **1** or **2** to designate the desired test, then enter **1**, **2**, or **3** to designate the desired alternative. The resampled permuted pair differences in mean (or correlations) are ordered and stored in list L3. The statistic from the original paired sample is displayed along with the *p*-value.

CAPIND (page 186) computes capability indexes based on given parameters or on data that has been entered into lists L1 and L2.

CONTRAST (page 136) computes the *p*-value for a significance test and a confidence interval for mean population contrasts. Before executing the program, enter the sample sizes into list L1, the sample means into list L2, the sample standard deviations into list L3, and the contrast equation coefficients into list L4. When prompted during the program, enter either **1** or **2** for a one-sided or two-sided alternative.

CONTRL (page 180) computes the upper and lower control limits and graphs the control charts for $\overline{x}$ and *s*.

CONTRL2 (page 183) computes the upper and lower control limits and graphs the control charts for $\overline{x}$ or *s* based on past data. Data sets of equal sizes must be entered into lists L1 and L2 in order to obtain the control limits for $\overline{x}$. Only the values of the sample deviations need to be entered into list L2 in order to compute the control limits for *s*.

CONTRLP (page 189) computes the control limits for sample proportions given either summary statistics or data entered into lists L1 and L2.

FITTEST (page 108) performs a goodness of fit test for a specified discrete distribution. Before executing, enter the specified proportions into list L1 and enter the observed cell counts into list L2. The expected cell counts are computed and stored in list L3; the contributions to the chi-square test statistic are stored in list L4. The program displays the test statistic and the *p*-value.

KRUSKAL (page 169) performs the Kruskal-Wallace test. Before executing, enter the data into the columns of matrix **[A]** and the sample sizes into a row matrix **[B]**. The program displays the test statistic and *p*-value. List L3 will contain the merged, sorted measurements, L4 will contain their (averaged) ranks, and L5 will contain the sum of ranks from each population.

LOG1 (page 175) computes and displays the coefficients of the linear regression model for the log of odds ratio $\ln(p/(1-p)) = \beta_0 + \beta_1 x$. Also displays the odds ratio.

MOVAVG (page 198) calculates the moving average of *N* values for a time series in L2 with time indices in L1. The time indices corresponding to the moving averages are stored in L3 and the averages in L4.

MULREG (page 123) computes the regression coefficients and an ANOVA table for the multiple linear regression model $\mu_Y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$. The squared correlation coefficient, *F*-statistic, *p*-value, and the standard deviation are also displayed. Before executing the program, enter sample data as columns into matrix [A] with the last column used for the dependent variable.

ODDS (page 173) computes the appropriate mathematical odds for a given probability p of an event *A*. If $p \le 0.50$, then the odds against *A* are given as the ratio $(1-p) : p$. If $p > 0.50$, then the odds in favor of *A* are given as the ratio $p : (1-p)$.

ODDS2 (page 174) computes the odds-in-favor ratio between two proportions.

ODDSINT (page 176) computes a confidence interval for the slope β_1 of the logistic regression model and the odds ratio.

ODDSTEST (page 176) computes the test statistic and *p*-value for the hypothesis test that an odds ratio equals 1.

PSAMPSZE (page 95) computes the required sample size that would give a maximum desired margin of error *m* for a confidence interval about a proportion.

RANDOM (page 46) randomly chooses a subset of specified size from the set {1, 2, . . ., *n*} and stores the values in list L1.

RANKSUM (page 163) performs the Wilcoxon rank sum test on data from two populations entered into lists L1 and L2. The program displays the expected sum of ranks from list L1, the sums of the ranks from each list, and the smallest tail value created by the test statistic which is the sum of the ranks from L1. List L3 then contains the merged, sorted measurements, and L4 contains their (averaged) ranks.

REGANV (page 118) computes the ANOVA table for linear regression and displays the associated *F*-statistic and *p*-value. Before executing the program, data must be entered into lists L1 and L2. The ANOVA table is stored into lists L4, L5, and L6.

REGINF (page 121) computes confidence intervals for the slope and for the intercept of the linear regression model $y = \beta_0 + \beta_1 x$, as well as confidence intervals for the mean value of *y* at a given *x* and prediction intervals *y* given a new value of *x*.

SIGNRANK (page 166) performs the Wilcoxon signed rank test on data sets of size *n* from two populations. Before executing, enter the data into lists L1 and L2. The program sorts the absolute value of the differences L2 – L1 into list L3, but disregards any zero differences. The (averaged) rank of each nonzero difference is stored in list L4. The sums of the ranks of the positive and negative differences are displayed. The program also displays the smallest tail value created by the test statistic, which is the sum of the ranks of the positive differences.

TSCORE (page 82) finds the critical value t^* of a *t* distribution upon specifying the degrees of freedom and confidence level.

Chapter 1 Problem Statements

1.3 The rating service Arbitron places U.S. radio stations into more than 50 categories that describe the kind of programs they broadcast. Which formats attract the largest audiences? Here are Arbitron's measurements of the share of the listening audience (aged 12 and over) for the most popular formats

Format	**Audience share**
Country	12.6%
News/Talk/Information	10.4%
Adult Contemporary	7.1%
Pop Contemporary Hit	5.5%
Classic Rock	4.7%
Rhythmic Contemporary Hit	4.2%
Urban Contemporary	4.1%
Urban Contemporary	3.4%
Oldies	3.3%
Hot Adult Contemporary	3.2%
Mexican Regional	3.1%

(a) What is the sum of the audience shares for these formats? What percent of the radio audience listens to stations with other formats?
(b) Make a bar graph to display these data. Be sure to include an ``Other format'' category.
(c) Would it be correct to display these data in a pie chart? Why?

1.5 Births are not, as you might think, evenly distributed across the days of the week. Here are the average numbers of babies born on each day of the week in 2005:

Day	**Births**
Sunday	7,374
Monday	11,704
Tuesday	13,169
Wednesday	13,038
Thursday	13,013
Friday	12,664
Saturday	8,459

Present these data in a well-labeled bar graph. Would it also be correct to make a pie chart? Suggest some possible reasons why there are fewer births on weekends.

1.11 Table 1.3 shows the annual spending per person on health care in the world's richer countries. Make a stemplot of the data after rounding to the nearest $100 (so that stems are thousands of dollars and leaves are hundreds of dollars). Split the stems, placing leaves 0 to 4 on the first stem and leaves 5 to 9 on the second stem of the same value. Describe the shape, center, and spread of the distribution. Which country is the high outlier?

Table 1.3 Annual spending per person on health care (in U.S. dollars)

Country	Dollars	Country	Dollars	Country	Dollars
Argentina	1067	Hungary	1269	Poland	745
Australia	2874	Iceland	3110	Portugal	1791
Austria	2306	Ireland	2496	Saudi Arabia	578
Belgium	2828	Israel	1911	Singapore	1156
Canada	2989	Italy	2266	Slovakia	777
Croatia	838	Japan	2244	Slovenia	1669
Czech Republic	1302	Korea	1074	South Africa	669
Denmark	2762	Kuwait	567	Spain	1853
Estonia	682	Lithuania	754	Sweden	2704
Finland	2108	Netherlands	2987	Switzerland	3776
France	2902	New Zealand	1893	United Kingdom	2389
Germany	3001	Norway	3809	United States	5711
Greece	1997	Oman	419		

1.25 The most popular colors for cars and light trucks change over time. Silver passed green in 2000 to become the most popular color worldwide, then gave way to shades of white in 2007. Here is the distribution of colors for vehicles sold in North America in 2007:

Color	Popularity
White	19%
Silver	18%
Black	16%
Red	13%
Gray	12%
Blue	12%
Beige, brown	5%
Other	

Fill in the percent of vehicles that are in other colors. Make a graph to display the distribution of color popularity.

1.27 Among persons aged 15 to 24 years in the United States, the leading causes of death and the number of deaths in 2005 were: accidents, 15,567; homicide, 5359; suicide, 4139; cancer, 1717; heart disease, 1067; congenital defects, 483.

(a) Make a bar graph to display these data.
(b) To make a pie chart, you need one additional piece of information. What is it?

1.29 Email spam is the curse of the Internet. Here is a compilation of the most common types of spam:

Type of Spam	Percent
Adult	19
Financial	20
Health	7
Internet	7
Leisure	6
Products	25
Scams	9

Make two bar graphs of these percents, one with bars ordered as in the table (alphabetically) and the other with bars in order from tallest to shortest. Comparisons are easier if you order the bars by height.

1.35 Table 1.5 gives the number of active medical doctors per 100,000 people in each state.

Medical doctors per 100,000 people, by state					
State	**Doctors**	**State**	**Doctors**	**State**	**Doctors**
Alabama	213	Louisiana	264	Ohio	261
Alaska	222	Maine	267	Oklahoma	171
Arizona	208	Maryland	411	Oregon	263
Arkansas	203	Massachusetts	450	Pennsylvania	294
California	259	Michigan	240	Rhode Island	351
Colorado	258	Minnesota	281	South Carolina	230
Connecticut	363	Mississippi	181	South Dakota	219
Delaware	248	Missouri	239	Tennessee	261
Florida	245	Montana	221	Texas	212
Georgia	220	Nebraska	239	Utah	209
Hawaii	310	Nevada	186	Vermont	362
Idaho	169	New Hampshire	260	Virginia	270
Illinois	272	New Jersey	306	Washington	265
Indiana	213	New Mexico	240	West Virginia	229
Iowa	187	New York	389	Wisconsin	254
Kansas	220	North Carolina	253	Wyoming	188
Kentucky	230	North Dakota	242	Dist. of Columbia	798

(a) Why is the number of doctors per 100,000 people a better measure of the availability of health care than a simple count of the number of doctors in a state?
(b) Make a histogram that displays the distribution of doctors per 100,000 people. Write a brief description of the distribution. Are there any outliers? If so, can you explain them?

1.37 "Recruitment," the addition of new members to a fish population, is an important measure of the health of ocean ecosystems. The table gives data on the recruitment of rock sole in the Bering Sea from 1973 to 2000. Make a stemplot to display the distribution of yearly rock sole
recruitment. (Round to the nearest hundred and split the stems.) Describe the shape, center, and spread of the distribution and any striking deviations that you see.

Year	Recruitment (millions)	Year	Recruitment (millions)	Year	Recruitment (millions)	Year	Recruitment (millions)
1973	173	1980	1411	1987	4700	1994	505
1974	234	1981	1431	1988	1702	1995	304
1975	616	1982	1250	1989	1119	1996	425
1976	344	1983	2246	1990	2407	1997	214
1977	515	1984	1793	1991	1049	1998	385
1978	576	1985	1793	1992	505	1999	445
1979	727	1986	2809	1993	998	2000	767

1.39 Make a time plot of the rock sole recruitment data in Exercise 1.37. What does the time plot show that your stemplot in Exercise 1.37 did not show? When you have time series data, a time plot is often needed to understand what is happening.

1.43 The impression that a time plot gives depends on the scales you use on the two axes. If you stretch the vertical axis and compress the time axis, change appears to be more rapid. Compressing the vertical axis and stretching the time axis make change appear slower. Make two more time plots of the college tuition data in Exercise 1.12 (page 24), one that makes tuition appear to increase very rapidly and one that shows only a gentle increase. The moral of this exercise is: pay close attention to the scales when you look at a time plot.

1.45 Here are data on the number of people bitten by alligators in Florida over a 36-year period:

Year	Number	Year	Number	Year	Number	Year	Number
1972	4	1981	10	1990	17	1999	16
1973	3	1982	7	1991	20	2000	23
1974	4	1983	9	1992	15	2001	25
1975	5	1984	9	1993	19	2002	17
1976	2	1985	7	1994	20	2003	12
1977	14	1986	23	1995	22	2004	13
1978	7	1987	13	1996	13	2005	15
1979	2	1988	18	1997	8	2006	18
1980	5	1989	13	1998	9	2007	18

(a) Make a histogram of the counts of people bitten by alligators. The distribution has an irregular shape. What is the midpoint of the yearly counts of people bitten?

(b) Make a time plot. There is great variation from year to year, but also an increasing trend. How many of the 22 years from 1986 to 2007 had more people bitten by alligators than your midpoint from (a)? The trend reflects Florida's growing population, which brings more people close to alligators.

Chapter 2 Problem Statements

2.1 Example 1.9 (page 20) gives the breaking strength in pounds of 20 pieces of Douglas fir. Find the mean breaking strength. How many of the pieces of wood have strengths less than the mean? What feature of the stemplot (Figure 1.11, page 21) explains the fact that the mean is smaller than most of the observations?

2.3 Find the mean of the travel times to work for the 20 New York workers in Example 2.3. Compare the mean and median for these data. What general fact does your comparison illustrate?

2.5 Table 1.4 (page 33) gives the ratio of two essential fatty acids in 30 food oils. Find the mean and the median for these data. Make a histogram of the data. What feature of the distribution explains why the mean is more than 10 times as large as the median?

2.11 The mean $\bar{x}$ and standard deviation s measure center and spread but are not a complete description of a distribution. Data sets with different shapes can have the same mean and standard deviation. To demonstrate this fact, use your calculator to find $\bar{x}$ and s for these two small data sets. Then make a stemplot of each and comment on the shape of each distribution.

Data A	9.14	8.14	8.74	8.77	9.26	8.10	6.13	3.10	9.13	7.26	4.74
Data B	6.58	5.76	7.71	8.84	8.47	7.04	5.25	5.56	7.91	6.89	12.50

2.13 "Conservationists have despaired over destruction of tropical rain forest by logging, clearing, and burning." These words begin a report on a statistical study of the effects of logging in Borneo. Charles Cannon of Duke University and his coworkers compared forest plots that had never been logged (Group 1) with similar plots nearby that had been logged 1 year earlier (Group 2) and 8 years earlier (Group 3). All plots were 0.1 hectare in area. Here are the counts of trees for plots in each group:

Group 1	27	22	29	21	19	33	16	20	24	27	28	19
Group 2	12	12	15	9	20	18	17	14	14	1	27	19
Group 3	18	4	22	15	18	19	22	12	12			

To what extent has logging affected the count of trees? Follow the four-step process in reporting your work.

2.29 An alternative presentation of the flower length data in Table 2.1 reports the five-number summary and uses boxplots to display the distributions. Do this. Do the boxplots fail to reveal any important information visible in the stemplots in Figure 2.5?

2.31 Here is the distribution of the weight at birth for all babies born in the United States in 2005:

Weight (grams)	Count	Weight (grams)	Count
Less than 500	6,599	3,000 to 3,499	1,596,944
500 to 999	23,864	3,500 to 3,999	1,114,887
1,000 to 1,499	31,325	4,000 to 4,499	289,098
1,500 to 1,999	66,453	4,500 to 4,999	42,119
2,000 to 2,499	210,324	5,000 to 5,499	4,715
2,500 to 2,999	748,042		

(a) For comparison with other years and with other countries, we prefer a histogram of the *percents* in each weight class rather than the counts. Explain why.
(b) How many babies were there? Make a histogram of the distribution, using percents on the vertical scale.
(c) What are the positions of the median and quartiles in the ordered list of all birth weights? In which weight classes do the median and quartiles fall?

2.35 Here are the survival times in days of 72 guinea pigs after they were injected with infectious bacteria in a medical\ experiment. Survival times, whether of machines under stress or cancer patients after treatment, usually have distributions that are skewed to the right.

43	45	53	56	56	57	58	66	67	73	74	79
80	80	81	81	81	82	83	83	84	88	89	91
91	92	92	97	99	99	100	100	101	102	102	102
103	104	107	108	109	113	114	118	1121	123	126	128
137	138	139	144	145	147	156	162	174	178	179	184
191	198	211	214	243	249	329	380	403	511	522	598

(a) Graph the distribution and describe its main features. Does it show the expected right skew?
(b) Which numerical summary would you choose for these data? Calculate your chosen summary. How does it reflect the skewness of the distribution?

2.37 Table 1.1 (page 12) gives the percent of foreign-born residents in each of the states. For the nation as a whole, 12.5% of residents are foreign-born. Find the mean of the 51 entries in Table 1.1. It is *not* 12.5%. Explain carefully why this happens. (*Hint*: The states with the largest populations are California, Texas, New York, and Florida. Look at their entries in Table 1.1.)

2.43 In 2007, the Boston Red Sox won the World Series for the second time in 4 years. Table 2.2 gives the salaries of the 25 players on the Red Sox World Series roster. Provide the team owner with a full description of the distribution of salaries and a brief summary of its most important features.

Salaries for the 2007 Boston Red Sox World Series team					
Player	**Salary**	**Player**	**Salary**	**Player**	**Salary**
Josh Beckett	$6,666,667	Jon Lester	$384,000	Jonathan Papelbon	$425,000
Alex Cora	$2,000,000	Javier López	$402,000	Dustin Pedroia	$380,000
Coco Crisp	$3,833,333	Mike Lowell	$9,000,000	Manny Ramirez	$17,016,381
Manny Delcarmen	$380,000	Julio Lugo	$8,250,000	Curt Schilling	$13,000,000
J. D. Drew	$14,400,000	D Matsuzaka	$6,333,333	Kyle Snyder	$535,000
Jacoby Ellsbury	$380,000	Doug Mirabelli	$750,000	Mike Timlin	$2,800,000
Eric Gagné	$6,000,000	Hideki Okajimi	$1,225,000	Jason Varitek	$11,000,000
Eric Hinske	$5,725,000	David Ortiz	$13,250,000	Kevin Youkilis	$424,000
Bobby Kielty	$2,100,000				

2.45 Businesses know that customers often respond to background music. Do they also respond to odors? Nicolas Guéguen and his colleagues studied this question in a small pizza restaurant in France on Saturday evenings in May. On one evening, a relaxing lavender odor was spread through the restaurant; on another evening, a stimulating lemon odor; a third evening served as a control, with no odor. Table 2.3 shows the amounts (in euros) that customers spent on each of these evenings. Compare the three distributions. What effect did the two odors have on customer spending?

Amount spent (euros) by customers in a restaurant when exposed to odors									
No odor									
15.9	18.5	15.9	18.5	18.5	21.9	15.9	15.9	15.9	15.9
15.9	18.5	18.5	18.5	20.5	18.5	18.5	15.9	15.9	15.9
18.5	18.5	15.9	18.5	15.9	18.5	15.9	25.5	12.9	15.9
Lemon Odor									
18.5	15.9	18.5	18.5	18.5	15.9	18.5	15.9	18.5	18.5
15.9	18.5	21.5	15.9	21.9	15.9	18.5	18.5	18.5	18.5
25.9	15.9	15.9	15.9	18.5	18.5	18.5	18.5		
Lavender Odor									
21.9	18.5	22.3	21.9	18.5	24.9	18.5	22.5	21.5	21.9
21.5	18.5	25.5	18.5	18.5	21.9	18.5	18.5	24.9	21.9
25.9	21.9	18.5	18.5	22.8	18.5	21.9	20.7	21.9	22.5

2.47 Farmers know that driving heavy equipment on wet soil compresses the soil and hinders the growth of crops. Table 2.5 gives data on the "penetrability" of the same soil at three levels of compression. Penetrability is a measure of the resistance plant roots meet when they grow through the soil. Low penetrability means high resistance. How does increasing compression affect penetrability?

Table 2.5 Penetrability of soil at three compression levels					
Compressed		**Intermediate**		**Loose**	
2.86	3.08	3.14	3.54	3.99	4.11
2.68	2.82	3.38	3.36	4.20	4.30
2.92	2.78	3.10	3.18	3.94	3.96
2.82	2.98	3.40	3.12	4.16	4.03
2.76	3.00	3.38	3.86	4.29	4.89
2.81	2.78	3.14	2.92	4.19	4.12
2.78	2.96	3.18	3.46	4.13	4.00
3.08	2.90	3.26	3.44	4.41	4.34
2.94	3.18	2.96	3.62	3.98	4.27
2.86	3.16	3.02	4.26	4.41	4.91

2.51 Which members of the Boston Red Sox (Table 2.2) have salaries that are suspected outliers by the 1.5 × IQR rule?

Chapter 3 Problem Statements

3.9 The heights of women aged 20 to 29 are approximately Normal with mean 64 inches and standard deviation 2.7 inches. Men the same age have mean height 69.3 inches with standard deviation 2.8 inches. What are the z-scores for a woman 6 feet tall and a man 6 feet tall? Say in simple language what information the z-scores give that the actual heights do not.

3.11 The summer monsoon rains in India follow approximately a Normal distribution with mean 852 millimeters (mm) of rainfall and standard deviation 82 mm.

(a) In the drought year 1987, 697 mm of rain fell. In what percent of all years will India have 697 mm or less of monsoon rain?
(b) "Normal rainfall" means within 20% of the long-term average, or between 683 mm and 1022 mm. In what percent of all years is the rainfall normal?

3.13 Use Table A to find the value z of a standard Normal variable that satisfies each of the following conditions. (Use the value of z from Table A that comes closest to satisfying the condition.) In each case, sketch a standard Normal curve with your value of z marked on the axis.

(a) The point z with 20% of the observations falling below it.
(b) The point z with 40% of the observations falling above it.

3.29

(a) Find the number z such that the proportion of observations that are less than z in a standard Normal distribution is 0.8.
(b) Find the number z such that 35% of all observations from a standard Normal distribution are greater than z.

3.31 Emissions of sulfur dioxide by industry set off chemical changes in the atmosphere that result in "acid rain." The acidity of liquids is measured by pH on a scale of 0 to 14. Distilled water has pH 7.0, and lower pH values indicate acidity. Normal rain is somewhat acidic, so acid rain is sometimes defined as rainfall with a pH below 5.0. The pH of rain at one location varies among rainy days according to a Normal distribution with mean 5.4 and standard deviation 0.54. What proportion of rainy days have rainfall with pH below 5.0?

3.33 Automated manufacturing operations are quite precise but still vary, often with distributions that are close to Normal. The width in inches of slots cut by a milling machine follows approximately the $N(0.8750, 0.0012)$ distribution. The specifications allow slot widths between 0.8720 and 0.8780 inch. What proportion of slots meet these specifications?

3.35 The 2008 Chevrolet Malibu with a four-cylinder engine has combined gas mileage 25 mpg. What percent of all vehicles have worse gas mileage than the Malibu?

3.37 The quartiles of any distribution are the values with cumulative proportions 0.25 and 0.75. They span the middle half of the distribution. What are the quartiles of the distribution of gas mileage?

3.39 Reports on a student's ACT or SAT usually give the percentile as well as the actual score. The percentile is just the cumulative proportion stated as a percent: the percent of all scores that were lower than this one. In 2007, composite ACT scores were close to Normal with mean 21.2 and standard deviation 5.0. Jacob scored 16. What was his percentile?

3.41 The heights of women aged 20 to 29 follow approximately the $N(64, 2.7)$ distribution. Men the same age have heights distributed as $N(69.3, 2.8)$. What percent of young women are taller than the mean height of young men?

3.43 Changing the mean and standard deviation of a Normal distribution by a moderate amount can greatly change the percent of observations in the tails. Suppose that a college is looking for applicants with SAT math scores 750 and above.

(a) In 2007, the scores of men on the math SAT followed the $N(533, 116)$ distribution. What percent of men scored 750 or better?
(b) Women's SAT math scores that year had the $N(499, 110)$ distribution. What percent of women scored 750 or better? You see that the percent of men above 750 is almost three times the percent of women with such high scores. Why this is true is controversial. (On the other hand, women score higher than men on the new SAT writing test, though by a smaller amount.)

3.47 Scores on the ACT test for the 2007 high school graduating class had mean 21.2 and standard deviation 5.0. In all, 1,300,599 students in this class took the test. Of these, 149,164 had scores higher than 27 and another 50,310 had scores exactly 27. ACT scores are always whole numbers. The exactly Normal $N(21.2, 5.0)$ distribution can include any value, not just whole numbers. What is more, there is *no* area exactly above 27 under the smooth Normal curve. So ACT scores can be only approximately Normal. To illustrate this fact, find

(a) the percent of 2007 ACT scores greater than 27.
(b) the percent of 2007 ACT scores greater than or equal to 27.
(c) the percent of observations from the $N(21.2,5.0)$ distribution that are greater than 27. (The percent greater than or equal to 27 is the same, because there is no area exactly over 27.)

3.49 Here are the lengths in millimeters of the thorax for 49 male fruit flies:

0.64	0.64	0.64	0.68	0.68	0.68	0.72	0.72	0.72	0.72
0.74	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.78
0.80	0.80	0.80	0.80	0.80	0.82	0.82	0.84	0.84	0.84
0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.88	0.88	0.88
0.88	0.88	0.88	0.88	0.88	0.92	0.92	0.92	0.94	

(a) Make a histogram of the distribution. Although the result depends a bit on your choice of classes, the distribution appears roughly symmetric with no outliers.
(b) Find the mean, median, standard deviation, and quartiles for these data. Comparing the mean and the median and comparing the distances of the two quartiles from the median suggest that the distribution is quite symmetric. Why?
(c) If the distribution were exactly Normal with the mean and standard deviation you found in (b), what proportion of observations would lie between the two quartiles you found in (b)? What proportion of the actual observations lie between the quartiles (include observations equal to either quartile value). Despite the discrepancy, this distribution is ``close enough to Normal" for statistical work in later chapters.

3.51 Table 2.5 (page 65) gives data on the penetrability of soil at each of three levels of compression. We might expect the penetrability of specimens of the same soil at the same level of compression to follow a Normal distribution. Make stemplots of the data for loose and for intermediate compression. Does either sample seem roughly Normal? Does either appear distinctly non-Normal? If so, what kind of departure from Normality does your stemplot show?

3.53 How many standard deviations above and below the mean do the quartiles of any Normal distribution lie? (Use the standard Normal distribution to answer this question.)

Chapter 4 Problem Statements

4.5 Airlines have increasingly outsourced the maintenance of their planes to other companies. Critics say that the maintenance may be less carefully done, so that outsourcing creates a safety hazard. As evidence, they point to government data on percent of major maintenance outsourced and percent of flight delays blamed on the airline (often due to maintenance problems):

Airline	Outsource percent	Delay percent	Airline	Outsource percent	Delay percent
AirTran	66	14	Frontier	65	31
Alaska	92	42	Hawaiian	80	70
American	46	26	JetBlue	68	18
America West	76	39	Northwest	76	43
ATA	18	19	Southwest	68	20
Continental	69	20	United	63	27
Delta	48	26	US Airways	77	24

Make a scatterplot that shows how delays depend on outsourcing.

4.7 Does your plot for Exercise 4.5 show a positive association between maintenance outsourcing and delays caused by the airline? One airline is a high outlier in delay percent. Which airline is this? Aside from the outlier, does the plot show a roughly linear form? Is the relationship very strong?

4.9 The study of dieting described in Exercise 4.4 collected data on the lean body mass (in kilograms) and metabolic rate (in calories) for both female and male subjects:

Sex	F	F	F	F	F	F	F	F	F	F
Mass	36.1	54.6	48.5	42.0	50.6	42.0	40.3	33.1	42.4	34.5
Rate	995	1425	1396	1418	1502	1256	1189	913	1124	1052
Sex	F	F	M	M	M	M	M	M	M	
Mass	51.1	41.2	51.9	46.9	62.0	62.9	47.4	48.7	51.9	
Rate	1347	1204	1867	1439	1792	1666	1362	1614	1460	

(a) Make a scatterplot of metabolic rate versus lean body mass for all 19 subjects. Use separate symbols to distinguish women and men.
(b) Does the same overall pattern hold for both women and men? What is the most important difference between women and men?

4.13 The gas mileage of an automobile first increases and then decreases as the speed increases. Suppose that this relationship is very regular, as shown by the following data on speed (miles per hour) and mileage (miles per gallon):

Speed	20	30	40	50	60
Mileage	24	28	30	28	24

Make a scatterplot of mileage versus speed. Show that the correlation between speed and mileage is $r = 0$. Explain why the correlation is 0 even though there is a strong relationship between speed and mileage.

4.27 Coffee is a leading export from several developing countries. When coffee prices are high, farmers often clear forest to plant more coffee trees. Here are five years of data on prices paid to coffee growers in Indonesia and the percent of forest area lost in a national park that lies in a coffee-producing region:

Price (cents per pound)	29	40	54	55	72
Forest Loss (percent)	0.49	1.59	1.69	1.82	3.10

(a) Make a scatterplot. Which is the explanatory variable? What kind of pattern does your plot show?
(b) Find the correlation r between coffee price and forest loss. Do your scatterplot and correlation support the idea that higher coffee prices increase the loss of forest?
(c) The price of coffee in international trade is given in dollars and cents. If the prices in the data were translated into the equivalent prices in euros, would the correlation between coffee price and percent of forest loss change? Explain your answer.

4.29 Most people dislike losses more than they like gains. In money terms, people are about as sensitive to a loss of $10 as to a gain of $20. To discover what parts of the brain are active in decisions about gain and loss, psychologists presented subjects with a series of gambles with different odds and different amounts of winnings and losses. From a subject's choices, they constructed a measure of “behavioral loss aversion.” Higher scores show greater sensitivity to losses. Observing brain activity while subjects made their decisions pointed to specific brain regions. Here are data for 16 subjects on behavioral loss aversion and “neural loss aversion,” a measure of activity in one region of the brain:

Neural	–50.0	–39.1	–25.9	–26.7	–28.6	–19.8	–17.6	5.5
Behavioral	0.08	0.81	0.01	0.12	0.68	0.11	0.36	0.34
Neural	2.6	20.7	12.1	15.5	28.8	41.7	55.3	155.2
Behavioral	0.53	0.68	0.99	1.04	0.66	0.86	1.29	1.94

(a) Make a scatterplot that shows how behavior responds to brain activity.
(b) Describe the overall pattern of the data. There is one clear outlier.

(c) Find the correlation *r* between neural and behavioral loss aversion both with and without the outlier. Does the outlier have a strong influence on the value of *r*? By looking at your plot, explain why adding the outlier to the other data points causes *r* to increase.

4.31 Japanese researchers measured the growth of icicles in a cold chamber under various conditions of temperature, wind, and water flow. Table 4.2 contains data produced under two sets of conditions. In both cases, there was no wind and the temperature was set at –11°C. Water flowed over the icicle at a higher rate (29.6 milligrams per second) in Run 8905 and at a slower rate (11.9 mg/s) in Run 8903.

Run 8903				**Run 8905**			
Time (min)	**Length (cm)**	**Time (min)**	**Length (cm)**	**Time (min)**	**Length (cm)**	**Time (min)**	**Length (cm)**
10	0.6	130	18.1	10	0.3	130	10.4
20	1.8	140	19.9	20	0.6	140	11.0
30	2.9	150	21.0	30	1.0	150	11.9
40	4.0	160	23.4	40	1.3	160	12.7
50	5.0	170	24.7	50	3.2	170	13.9
60	6.1	180	27.8	60	4.0	180	14.6
70	7.9			70	5.3	190	15.8
80	10.1			80	6.0	200	16.2
90	10.9			90	6.9	210	17.9
100	12.7			100	7.8	220	18.8
110	14.4			110	8.3	230	19.9
120	16.6			120	9.6	240	21.1

(a) Make a scatterplot of the length of the icicle in centimeters versus time in minutes, using separate symbols for the two runs.
(b) What does your plot show about the pattern of growth of icicles? What does it show about the effect of changing the rate of water flow on icicle growth?

4.33 To detect the presence of harmful insects in farm fields, we can put up boards covered with a sticky material and examine the insects trapped on the boards. Which colors attract insects best? Experimenters placed six boards of each of four colors at random locations in a field of oats and measured the number of cereal leaf beetles trapped. Here are the data:

Board color	**Beetles trapped**					
Blue	16	11	20	21	14	7
Green	37	32	20	29	37	32
White	21	12	14	17	13	20
Yellow	45	59	48	46	38	47

(a) Make a plot of beetles trapped against color (space the four colors equally on the horizontal axis). Which color appears best at attracting beetles?
(b) Does it make sense to speak of a positive or negative association between board color and beetles trapped? Why? Is correlation *r* a helpful description of the relationship? Why?

4.43 We have data from a house in the Midwest that uses natural gas for heating. Will installing solar panels reduce the amount of gas consumed? Gas consumption is higher in cold weather, so the relationship between outside temperature and gas consumption is important. Here are data for 16 consecutive months:

	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June
Degree-days per day	24	51	43	33	26	13	4	0
Gas used per day	6.3	10.9	8.9	7.5	5.3	4.0	1.7	1.2
	July	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.
Degree-days per day	0	1	6	12	30	32	52	30
Gas used per day	1.2	1.2	2.1	3.1	6.4	7.2	11.0	6.9

Outside temperature is recorded in degree-days, a common measure of demand for heating. A day's degree-days are the number of degrees its average temperature falls below 65°. Gas used is recorded in hundreds of cubic feet. Here are data for 23 more months after installing solar panels:

Degree-days	19	3	3	0	0	0	8	11	27	46	38	34
Gas used	3.2	2.0	1.6	1.0	0.7	0.7	1.6	3.1	5.1	7.7	7.0	6.1
Degree-days	16	9	2	1	0	2	3	18	32	34	40	
Gas used	3.0	2.1	1.3	1.0	1.0	1.0	1.2	3.4	6.1	6.5	7.5	

What do the before-and-after data show about the effect of solar panels? (Start by plotting both sets of data on the same plot, using two different plotting symbols.)

4.49 We often describe our emotional reaction to social rejection as "pain." Does social rejection cause activity in areas of the brain that are known to be activated by physical pain? If it does, we really do experience social and physical pain in similar ways. Psychologists first included and then deliberately excluded individuals from a social activity while they measured changes in brain activity. After each activity, the subjects filled out questionnaires that assessed how excluded they felt. Here are data for 13 subjects.

The explanatory variable is "social distress" measured by each subject's questionnaire score after exclusion relative to the score after inclusion. (So values greater than 1 show the degree of distress caused by exclusion.) The response variable is change in activity in a region of the brain that is activated by physical pain. Discuss what the data show.

Subject	Social distress	Brain activity	Subject	Social distress	Brain activity
1	1.26	–0.055	8	2.18	0.025
2	1.85	–0.040	9	2.58	0.027
3	1.10	–0.026	10	2.75	0.033
4	2.50	–0.017	11	2.75	0.064
5	2.17	–0.017	12	3.33	0.077
6	2.67	0.017	13	3.65	0.124
7	2.01	0.021			

Chapter 5 Problem Statements

5.3 An outbreak of the deadly Ebola virus in 2002 and 2003 killed 91 of the 95 gorillas in 7 home ranges in the Congo. To study the spread of the virus, measure "distance" by the number of home ranges separating a group of gorillas from the first group infected. Here are data on distance and number of days until deaths began in each later group:

Distance	1	3	4	4	4	5
Days	4	21	33	41	43	46

As you saw in Exercise 4.10 (page 106), there is a linear relationship between distance x and days y.

(a) Use your calculator to find the mean and standard deviation of both x and y and the correlation r between x and y. Use these basic measures to find the equation of the least-squares line for predicting y from x.
(b) Enter the data into your software or calculator and use the regression function to find the least-squares line. The result should agree with your work in (a) up to roundoff error.

5.11 Exercise 4.5 (page 99) gives data for 14 airlines on the percent of major maintenance outsourced and the percent of flight delays blamed on the airline.

(a) Make a scatterplot with outsourcing percent as x and delay percent as y. Hawaiian Airlines is a high outlier in the y direction. Because several other airlines have similar values of x, the influence of this outlier is unclear without actual calculation.
(b) Find the correlation r with and without Hawaiian Airlines. How influential is the outlier for correlation?
(c) Find the least-squares line for predicting y from x with and without Hawaiian Airlines. Draw both lines on your scatterplot. Use both lines to predict the percent of delays blamed on an airline
that has outsourced 76% of its major maintenance. How influential is the outlier for the least-squares line?

5.35 Keeping water supplies clean requires regular measurement of levels of pollutants. The measurements are indirect—a typical analysis involves forming a dye by a chemical reaction with the dissolved pollutant, then passing light through the solution and measuring its "absorbence." To calibrate such measurements, the laboratory measures known standard solutions and uses regression to relate absorbence and pollutant concentration. This is usually done every day. Here is one series of data on the absorbence for different levels of nitrates. Nitrates are measured in milligrams per liter of water.

Nitrates	50	50	100	200	400	800	1200	1600	2000	2000
Absorbence	7.0	7.5	12.8	24.0	47.0	93.0	138.0	183.0	230.0	226.0

(a) Chemical theory says that these data should lie on a straight line. If the correlation is not at least 0.997, something went wrong and the calibration procedure is repeated. Plot the data and find the correlation. Must the calibration be done again?
(b) The calibration process sets nitrate level and measures absorbence. The linear relationship that results is used to estimate the nitrate level in water from a measurement of absorbence. What is the equation of the line used to estimate nitrate level? What is the estimated nitrate level in a water specimen with absorbence 40?
(c) Do you expect estimates of nitrate level from absorbence to be quite accurate? Why?

5.37 Exercise 4.29 (page 116) describes an experiment that showed a linear relationship between how sensitive people are to monetary losses ("behavioral loss aversion"") and activity in one part of their brains ("neural loss aversion").

(a) Make a scatterplot with neural loss aversion as x and behavioral loss aversion as y. One point is a high outlier in both the x and y directions.
(b) Find the least-squares line for predicting y from x, *leaving out the outlier*, and add the line to your plot.
(c) The outlier lies very close to your regression line. Looking at the plot, you now expect that adding the outlier will increase the correlation but will have little effect on the least-squares line. Explain why.
(d) Find the correlation and the equation of the least-squares line with and without the outlier. Your results verify the expectations from (c).

5.39 People with diabetes must manage their blood sugar levels carefully. They measure their fasting plasma glucose (FPG) several times a day with a glucose meter. Another measurement, made at regular medical checkups, is called HbA. This is roughly the percent of red blood cells that have a glucose molecule attached. It measures average exposure to glucose over a period of several months. Table 5.2 gives data on both HbA and FPG for 18 diabetics five months after they had completed a diabetes education class.

Table 5.2 Two measures of glucose level in diabetics

Subject	HbA (%)	FPG (mg/ml)	Subject	HbA (%)	FPG (mg/ml)	Subject	HbA (%)	FPG (mg/ml)
1	6.1	141	7	7.5	96	13	10.6	103
2	6.3	158	8	7.7	87	14	10.7	172
3	6.4	112	9	7.9	148	15	10.7	359
4	6.8	153	10	8.7	172	16	11.2	145
5	7.0	143	11	9.4	200	17	13.7	147
6	7.1	95	12	10.5	271	18	19.3	255

(a) Make a scatterplot with HbA as the explanatory variable. There is a positive linear relationship, but it is surprisingly weak.
(b) Subject 15 is an outlier in the *y* direction. Subject 18 is an outlier in the *x* direction. Find the correlation for all 18 subjects, for all except Subject 15, and for all except Subject 18. Are either or both of these subjects influential for the correlation? Explain in simple language why *r* changes in opposite directions when we remove each of these points.

5.41 Add three regression lines for predicting FPG from HbA to your scatterplot from Exercise 5.39: for all 18 subjects, for all except Subject 15, and for all except Subject 18. Is either Subject 15 or Subject 18 strongly influential for the least-squares line? Explain in simple language what features of the scatterplot explain the degree of influence.

5.51 Do beavers benefit beetles? Researchers laid out 23 circular plots, each 4 meters in diameter, in an area where beavers were cutting down cottonwood trees. In each plot, they counted the number of stumps from trees cut by beavers and the number of clusters of beetle larvae. Ecologists think that the new sprouts from stumps are more tender than other cottonwood growth, so that beetles prefer them. If so, more stumps should produce more beetle larvae. Here are the data:

Stumps	2	2	1	3	3	4	3	1	2	5	1	3
Beetle Larvae	10	30	12	24	36	40	43	11	27	56	18	40
Stumps	2	1	2	2	1	1	4	1	2	1	4	
Beetle Larvae	25	8	21	14	16	6	54	9	13	14	50	

Analyze these data to see if they support the “beavers benefit beetles” idea.

5.55 Exercise 4.43 (page 121) gives monthly data on outside temperature (in degree-days per day) and natural gas consumed for a house in the Midwest both before and after installing solar panels. A cold winter month in this location may average 45 degree-days per day (temperature 20°). Use before-and-after regression lines to estimate the savings in gas consumption due to solar panels.

Chapter 6 Problem Statements

6.1 Recycling is supposed to save resources. Some people think recycled products are lower in quality than other products, a fact that makes recycling less practical. Here are data on attitudes toward coffee filters made of recycled paper among people who had bought these filters and people who had not:

	Think the quality of the recycled product is		
	Higher	**The same**	**Lower**
Buyers	20	7	9
Nonbuyers	29	25	43

(a) How many people does this table describe? How many of these were buyers of coffee filters made of recycled paper?
(b) Give the marginal distribution of opinion about the quality of recycled filters. What percent of consumers think the quality of the recycled product is the same or higher than the quality of other filters?

6.3 Exercise 6.1 gives data on the opinions of people who have and have not bought coffee filters made from recycled paper. To see the relationship between opinion and experience with the product, find the conditional distributions of opinion (the response variable) for buyers and nonbuyers. What do you conclude?

6.19 Will giving cocaine addicts an antidepressant drug help them break their addiction? An experiment assigned 24 chronic cocaine users to take the antidepressant drug desipramine, another 24 to take lithium, and another 24 to take a placebo. (Lithium is a standard drug to treat cocaine addiction. A placebo is a dummy pill, used so that the effect of being in the study but not taking any drug can be seen.) After three years, 14 of the 24 subjects in the desipramine group had remained free of cocaine, along with 6 of the 24 in the lithium group and 4 of the 24 in the placebo group.

(a) Make up a two-way table of “Treatment received” by whether or not the subject remained free of cocaine.
(b) Compare the effectiveness of the three treatments in preventing use of cocaine by former addicts. Use percents and draw a bar graph. What do you conclude?

6.27 The University of Chicago's General Social Survey asked a representative sample of adults this question: “Which of the following statements best describes how your daily work is organized? 1: I am free to decide how my daily work is organized. 2: I can decide how my daily work is organized, within certain limits. 3: I am not free to decide how my daily work is organized." Here is a two-way table of the responses for three levels of education:

Response	Highest Degree Completed		
	Less than high school	**High school**	**Bachelor's**
1	31	161	81
2	49	269	85
3	47	112	14

How does freedom to organize you work depend on level of education?

6.29 "Colleges and universities across the country are grappling with the case of the mysteriously vanishing male." So said an article in the *Washington Post*. Here are data on the numbers of degrees earned in 2009–2010, as projected by the National Center for Education Statistics. The table entries are counts of degrees in thousands.

	Female	Male
Associate's	447	268
Bachelor's	945	651
Master's	397	251
Professional	49	44
Doctor's	26	25

Briefly contrast the participation of men and women in earning degrees.

Chapter 7 Problem Statements

7.3 The Pew Research Center asked a random sample of adults whether they had favorable or unfavorable opinions of a number of major companies. Answers to such questions depend a lot on recent news. Here are the percents with favorable opinions for several of the companies:

Company	Percent Favorable
Apple	71
Ben and Jerry's	59
Coors	53
Exxon/Mobil	44
Google	73
Haliburton	25
McDonald's	71
Microsoft	78
Starbucks	64
Wal-Mart	68

Make a graph that displays these data.

7.5 Here are the weights (in milligrams) of 58 diamonds from a nodule carried up to the earth's surface in surrounding rock. This represents a single population of diamonds formed in a single event deep in the earth.

13.8	3.7	33.8	11.8	27.0	18.9	19.3	20.8	25.4	23.1	7.8	10.9
9.0	9.0	14.4	6.5	7.3	5.6	18.5	1.1	11.2	7.0	7.6	9.0
9.5	7.7	7.6	3.2	6.5	5.4	7.2	7.8	3.5	5.4	5.1	5.3
3.8	2.1	2.1	4.7	3.7	3.8	4.9	2.4	1.4	0.1	4.7	1.5
2.0	0.1	0.1	1.6	3.5	3.7	2.6	4.0	2.3	4.5		

Make a graph that shows the distribution of weights of diamonds. Describe the shape of the distribution and any outliers. Use numerical measures appropriate for the shape to describe the center and spread.

7.9 The Aleppo pine and the Torrey pine are widely planted as ornamental trees in Southern California. Here are the lengths (centimeters) of 15 Aleppo pine needles:

10.2 7.2 7.6 9.3 12.1 10.9 9.4 11.3 8.5 8.5 12.8 8.7 9.0 9.0 9.4

Here are the lengths of 18 needles from Torrey pines:

33.7	21.2	26.8	29.7	21.6	21.7	33.7	32.5	23.1
23.7	30.2	29.0	24.2	24.4	25.5	26.6	28.9	29.7

Use five-number summaries and boxplots to compare the two distributions. Given only the length of a needle, do you think you could say which pine species it comes from?

7.15 The lengths of needles from Aleppo pines follow approximately the Normal distribution with mean 9.6 centimeters (cm) and standard deviation 1.6 cm. According to the 68–95–99.7 rule, what range of lengths covers the center 95% of Aleppo pine needles? What percent of needles are less than 6.4 cm long?

7.17 Almost all medical schools in the United States require applicants to take the Medical College Admission Test (MCAT). The scores of applicants on the biological sciences part of the MCAT in 2007 were approximately Normal with mean 9.6 and standard deviation 2.2. For applicants who actually entered medical school, the mean score was 10.6 and the standard deviation was 1.7.

(a) What percent of all applicants had scores higher than 13?
(b) What percent of those who entered medical school had scores between 8 and 12?

7.19 From Rex Boggs in Australia comes an unusual data set: before showering in the morning, he weighed the bar of soap in his shower stall. The weight goes down as the soap is used. The data appear below (weights in grams). Notice that Mr.\ Boggs forgot to weigh the soap on some days.

Day	**Weight**	**Day**	**Weight**	**Day**	**Weight**
1	124	8	84	16	27
2	121	9	78	18	16
5	103	10	71	19	12
6	96	12	58	20	8
7	90	13	50	21	6

Plot the weight of the bar of soap against day. Is the overall pattern roughly linear? Based on your scatterplot, is the correlation between day and weight close to 1, positive but not close to 1, close to 0, negative but not close to –1, or close to –1? Explain your answer. Then find the correlation *r* to verify what you concluded from the graph.

7.23 Animals and people that take in more energy than they expend will get fatter. Here are data on 12 rhesus monkeys: 6 lean monkeys (4% to 9% body fat) and 6 obese monkeys (13% to 44% body fat). The data report the energy expended in 24 hours (kilojoules per minute) and the lean body mass (kilograms, leaving out fat) for each monkey.

Lean		**Obese**	
Mass	**Energy**	**Mass**	**Energy**
6.6	1.17	7.9	0.93
7.8	1.02	9.4	1.39
8.9	1.46	10.7	1.19
9.8	1.68	12.2	1.49
9.7	1.06	12.1	1.29
9.3	1.16	10.8	1.31

(a) What is the mean lean body mass of the lean monkeys? Of the obese monkeys? Because animals with higher lean mass usually expend more energy, we can't directly compare energy expended

(b) Instead, look at how energy expended is related to body mass. Make a scatterplot of energy versus mass, using different plot symbols for lean and obese monkeys. Then add to the plot two regression lines, one for lean monkeys and one for obese monkeys. What do these lines suggest about the monkeys?

7.25 That animal species produce more offspring when their supply of food goes up isn't surprising. That some animals appear able to anticipate unusual food abundance is more surprising. Red squirrels eat seeds from pine cones, a food source that occasionally has very large crops (called seed masting). Here are data on an index of the abundance of pine cones and average number of offspring per female over 16 years:

Cone Index	0.00	2.02	0.25	3.22	4.68	0.31	3.37	3.09
Offspring	1.49	1.10	1.29	2.71	4.07	1.29	3.36	2.41
Cone Index	2.44	4.81	1.88	0.31	1.61	1.88	0.91	1.04
Offspring	1.97	3.41	1.49	2.02	3.34	2.41	2.15	2.12

Describe the relationship with both a graph and numerical measures, then summarize in words. What is striking is that the offspring are conceived in the spring, *before* the cones mature in the fall to feed the new young squirrels through the winter.

7.27 The usual way to study the brain's response to sounds is to have subjects listen to "pure tones." The response to recognizable sounds may differ. To compare responses, researchers anesthetized macaque monkeys. They fed pure tones and also monkey calls directly to their brains by inserting electrodes. Response to the stimulus was measured by the firing rate (electrical spikes per second) of neurons in various areas of the brain. Table 7.1 contains the responses for 37 neurons.

Table 7.1 Neuron response to tones and monkey calls								
Neuron	**Tone**	**Call**	**Neuron**	**Tone**	**Call**	**Neuron**	**Tone**	**Call**
1	474	500	14	145	42	26	71	134
2	256	138	15	141	241	27	68	65
3	241	485	16	129	294	28	59	182
4	226	338	17	113	123	29	59	97
5	185	194	18	112	182	30	57	318
6	174	159	19	102	141	31	56	201
7	176	341	20	100	118	32	47	279
8	168	85	21	74	62	33	46	62
9	161	303	22	72	112	34	41	84
10	150	208	23	20	193	35	26	203
11	19	66	24	21	129	36	28	192
12	20	54	25	26	135	37	31	70
13	35	103						

(a) One important finding is that responses to monkey calls are generally stronger than responses to pure tones. For how many of the 37 neurons is this true?

(b) We might expect some neurons to have strong responses to any stimulus and others to have consistently weak responses. There would then be a strong relationship between tone response and call response. Make a scatterplot of monkey call response against pure tone response (explanatory variable). Find the correlation r between tone and call responses. How strong is the linear relationship?

7.37 We have 91 years of data on the date of ice breakup on the Tanana River. Describe the distribution of the breakup date with both a graph or graphs and appropriate numerical summaries. What is the median date (month and day) for ice breakup?

Table 7.3 Days from April 20 for the Tanana River tripod to fall											
Year	**Day**	**Year**	**Day**	**Year**	**Day**	**Year**	**Day**	**Year**	**Day**	**Year**	**Day**
1917	11	1933	19	1949	25	1965	18	1981	11	1997	11
1918	22	1934	11	1950	17	1966	19	1982	21	1998	1
1919	14	1935	26	1951	11	1967	15	1983	10	1999	10
1920	22	1936	11	1952	23	1968	19	1984	20	2000	12
1921	22	1937	23	1953	10	1969	9	1985	23	2001	19
1922	23	1938	17	1954	17	1970	15	1986	19	2002	18
1923	20	1939	10	1955	20	1971	19	1987	16	2003	10
1924	22	1940	1	1956	12	1972	21	1988	8	2004	5
1925	16	1941	14	1957	16	1973	15	1989	12	2005	9
1926	7	1942	11	1958	10	1974	17	1990	5	2006	13
1927	23	1943	9	1959	19	1975	21	1991	12	2007	8

1928	17	1944	15	1960	13	1976	13	1992	25	
1929	16	1945	27	1961	16	1977	17	1993	4	
1930	19	1946	16	1962	23	1978	11	1994	10	
1931	21	1947	14	1963	16	1979	11	1995	7	
1932	12	1948	24	1964	31	1980	10	1996	16	

7.43 Here is one way that nature regulates the size of animal populations: High population density attracts predators, who remove a higher proportion of the population than when the density of the prey is low. One study looked at kelp perch and their common predator, the kelp bass. The researcher set up four large circular pens on the sandy ocean bottom in southern California. He chose young perch at random from a large group and placed 10, 20, 40, and 60 perch in the four pens. Then he dropped the nets protecting the pens, allowing bass to swarm in, and counted the perch left after 2 hours. Here are data on the proportions of perch eaten in four repetitions of this setup:

Perch	**Proportion killed**			
10	0.0	0.1	0.3	0.3
20	0.2	0.3	0.3	0.6
40	0.075	0.3	0.6	0.725
60	0.517	0.55	0.7	0.817

Do the data support the principle that "more prey attract more predators, who drive down the number of prey"? Follow the four-step process (page 55) in your answer.

7.49 Table 7.1 (page 188) contains data on the response of 37 monkey neurons to pure tones and to monkey calls. You made a scatterplot of these data in Exercise 7.27.

(a) Find the least-squares line for predicting a neuron's call response from its pure tone response. Add the line to your scatterplot. Mark on your plot the point (call it A) with the largest residual (either positive or negative) and also the point (call it B) that is an outlier in the x direction.
(b) How influential are each of these points for the correlation r?
(c) How influential are each of these points for the regression line?

Chapter 8 Problem Statements

8.7 A firm wants to understand the attitudes of its minority managers toward its system for assessing management performance. Below is a list of all the firm's managers who are members of minority groups. Use Table B at line 139 to choose six to be interviewed in detail about the performance appraisal system.

1	Abdulhamid	8	Duncan	15	Huang	22	Puri
2	Agarwal	9	Fernandez	16	Kim	23	Richards
3	Baxter	10	Fleming	17	Lumumba	24	Rodriguez
4	Bonds	11	Gates	18	Mourning	25	Santiago
5	Brown	12	Gomez	19	Nguyen	26	Shen
6	Castro	13	Gupta	20	Peters	27	Vargas
7	Chavez	14	Hernandez	21	Peña	28	Wang

8.11 Cook County, Illinois, has the second-largest population of any county in the United States (after Los Angeles County, California). Cook County has 30 suburban townships and an additional 8 townships that make up the city of Chicago. The suburban townships are

Barrington	Elk Grove	Maine	Orland	Riverside
Berwyn	Evanston	New Trier	Palatine	Schaumburg
Bloom	Hanover	Niles	Palos	Stickney
Bremen	Lemont	Northfield	Proviso	Thornton
Calumet	Leyden	Norwood	Park Rich	Wheeling
Cicero	Lyons	Oak Park	River Forest	Worth

The Chicago townships are

Hyde Park	Lake	North Chicago	South Chicago
Jefferson	Lake View	Rogers Park	West Chicago

Because city and suburban areas may differ, the first stage of a multistage sample chooses a stratified sample of 6 suburban townships and 4 of the more heavily populated Chicago townships. Use Table B or software to choose this sample. (If you use Table B, assign labels in alphabetical order and start at line 101 for the suburbs and at line 110 for Chicago.)

8.27 You want to ask a sample of college students the question "How much do you trust information about health that you find on the Internet—a great deal, somewhat, not much, or not at all?" You try out this and other questions on a pilot group of 10 students chosen from your class. The class members are

Anderson	Deng	Glaus	Nguyen	Samuels
Arroyo	De Ramos	Helling	Palmiero	Shen
Batista	Drasin	Husain	Percival	Tse
Bell	Eckstein	Johnson	Prince	Velasco
Burke	Fernandez	Kim	Puri	Wallace
Cabrera	Fullmer	Molina	Richards	Washburn
Calloway	Gandhi	Morgan	Rider	Zabidi
Delluci	Garcia	Murphy	Rodriguez	Zhao

Choose an SRS of 10 students. If you use Table B, start at line 117.

8.29 To gather data on a 1200-acre pine forest in Louisiana, the U.S. Forest Service laid a grid of 1410 equally spaced circular plots over a map of the forest. A ground survey visited a sample of 10% of these plots.

(a) How would you label the plots?
(b) Choose the first 5 plots in an SRS of 141 plots. (If you use Table B, start at line 105.)

8.39 At a large block party there are 290 men and 110 women. You want to ask opinions about how to improve the next party. To be sure that women's opinions are adequately represented, you decide to choose a stratified random sample of 20 men and 20 women. Explain how you will assign labels to the names of the people at the party. Give the labels of the first 3 men and the first 3 women in your sample. If you use Table B, start at line 130.

Chapter 9 Problem Statements

9.9 The changing climate will probably bring more rain to California, but we don't know whether the additional rain will come during the winter wet season or extend into the long dry season in spring and summer. Kenwyn Suttle of the University of California at Berkeley and his coworkers carried out a randomized controlled experiment to study the effects of more rain in either season. They randomly assigned plots of open grassland to 3 treatments: added water equal to 20% of annual rainfall either during January to March (winter) or during April to June (spring), and no added water (control). Thirty-six circular plots of area 70 square meters were available (see the photo), of which 18 were used for this study. One response variable was total plant biomass, in grams per square meter, produced in a plot over a year.

(a) Outline the design of the experiment, following the model of Figure 9.4.
(b) Number all 36 plots and choose 6 at random for each of the 3 treatments. Be sure to explain how you did the random selection.

9.33 Elementary schools in rural India are usually small, with a single teacher. The teachers often fail to show up for work. Here is an idea for improving attendance: give the teacher a digital camera with a tamper-proof time and date stamp and ask a student to take a photo of the teacher and class at the beginning and end of the day. Offer the teacher better pay for good attendance, verified by the photos. Will this work? A randomized comparative experiment started with 120 rural schools in Rajasthan and assigned 60 to this treatment and 60 to a control group. Random checks for teacher attendance showed that 21% of teachers in the treatment group were absent, as opposed to 42% in the control group
(a) Outline the design of this experiment.
(b) Label the schools and choose the first 10 schools for the treatment group. If you use Table B, start at line 108.

9.35 Some people think that red wine protects moderate drinkers from heart disease better than other alcoholic beverages. This calls for a randomized comparative experiment. The subjects were healthy men aged 35 to 65. They were randomly assigned to drink red wine (9 subjects), drink white wine (9 subjects), drink white wine and also take polyphenols from red wine (6 subjects), take polyphenols alone (9 subjects), or drink vodka and lemonade (6 subjects). Outline the design of the experiment and randomly assign the 39 subjects to the 5 groups. If you use Table B, start at line 107.

9.37 Doctors identify “chronic tension-type headaches” as headaches that occur almost daily for at least six months. Can antidepressant medications or stress management training reduce the number and severity of these headaches? Are both together more effective than either alone?

(a) Use a diagram like Figure 9.2 to display the treatments in a design with two factors: medication yes or no and stress management yes or no. Then outline the design of a completely randomized experiment to compare these treatments.

(b) The headache sufferers named below have agreed to participate in the study. Randomly assign the subjects to the treatments. If you use the *Simple Random Sample* applet or other software, assign all the subjects. If you use Table A, start at line 130 and assign subjects to only the first treatment group.

Abbott	Decker	Herrera	Lucero	Richter
Abdalla	Devlin	Hersch	Masters	Riley
Alawi	Engel	Hurwitz	Morgan	Samuels
Broden	Fuentes	Irwin	Nelson	Smith
Chai	Garrett	Jiang	Nho	Suarez
Chuang	Gill	Kelley	Ortiz	Upasani
Cordoba	Glover	Kim	Ramdas	Wilson
Custer	Hammond	Landers	Reed	Xiang

9.41 Here's the opening of a Starbucks press release: “Starbucks Corp. on Monday said it would roll out a line of blended coffee drinks intended to tap into the growing popularity of reduced-calorie and reduced-fat menu choices for Americans.” You wonder if Starbucks customers like the new ``Mocha Frappuccino Light" as well as the regular Mocha Frappuccino coffee.

(a) Describe a matched pairs design to answer this question. Be sure to include proper blinding of your subjects.

(b) You have 20 regular Starbucks customers on hand. Use the *Simple Random Sample* applet or Table B at line 141 to do the randomization that your design requires.

Chapter 10 Problem Statements

10.47 A couple plans to have three children. There are 8 possible arrangements of girls and boys. For example, GGB means the first two children are girls and the third child is a boy. All 8 arrangements are (approximately) equally likely.

(a) Write down all 8 arrangements of the sexes of three children. What is the probability of any one of these arrangements?
(b) Let X be the number of girls the couple has. What is the probability that $X = 2$?
(c) Starting from your work in (a), find the distribution of X. That is, what values can X take, and what are the probabilities for each value?

10.51 A sample survey contacted an SRS of 663 registered voters in Oregon shortly after an election and asked respondents whether they had voted. Voter records show that 56\% of registered voters had actually voted. We will see later that in this situation the proportion of the sample who voted (call this proportion V) has approximately the Normal distribution with mean $\mu = 0.56$ and standard deviation $\sigma = 0.019$.

(a) If the respondents answer truthfully, what is $P(0.52 \leq V \leq 0.60)$? This is the probability that the sample proportion V estimates the population proportion 0.56 within plus or minus 0.04.
(b) In fact, 72% of the respondents said they had voted ($V = 0.72$). If respondents answer truthfully, what is $P(V \geq 0.72)$? This probability is so small that it is good evidence that some people who did not vote claimed that they did vote.

Chapter 11 Problem Statements

11.7 Let's illustrate the idea of a sampling distribution in the case of a very small sample from a very small population. The population is the scores of 10 students on an exam:

Student	0	1	2	3	4	5	6	7	8	9
Score	82	62	80	58	72	73	65	66	74	62

The parameter of interest is the mean score μ in this population. The sample is an SRS of size $n = 4$ drawn from the population. Because the students are labeled 0 to 9, a single random digit from Table B chooses one student for the sample.

(a) Find the mean of the 10 scores in the population. This is the population mean μ

(b) Use the first digits in row 116 of Table B to draw an SRS of size 4 from this population. What are the four scores in your sample? What is their mean $\bar{x}$? This statistic is an estimate of μ.

(c) Repeat this process 9 more times, using the first digits in rows 117 to 125 of Table B. Make a histogram of the 10 values of $\bar{x}$. You are constructing the sampling distribution of $\bar{x}$. Is the center of your histogram close to μ?

11.9 Suppose that in fact the blood cholesterol level of all men aged 20 to 34 follows the Normal distribution with mean $\mu = 188$ milligrams per deciliter (mg/dl) and standard deviation $\sigma = 41$ mg/dl.

(a) Choose an SRS of 100 men from this population. What is the sampling distribution of $\bar{x}$? What is the probability that $\bar{x}$ takes a value between 185 and 191 mg/dl? This is the probability that $\bar{x}$ estimates μ within ± 3 mg/dl.

(b) Choose an SRS of 1000 men from this population. Now what is the probability that $\bar{x}$ falls within ± 3 mg/dl of μ? The larger sample is much more likely to give an accurate estimate of μ.

11.13 An insurance company knows that in the entire population of millions of homeowners, the mean annual loss from fire is $\mu = 250$ and the standard deviation of the loss is $\sigma = 1000$. The distribution of losses is strongly right-skewed: most policies have $0 loss, but a few have large losses. If the company sells 10,000 policies, can it safely base its rates on the assumption that its average loss will be no greater than $275? Follow the four-step process as illustrated in Example 11.8.

11.27 Shelia's doctor is concerned that she may suffer from gestational diabetes (high blood glucose levels during pregnancy). There is variation both in the actual glucose level and in the blood test that measures the level. A patient is classified as having gestational diabetes if the glucose level is above 140 milligrams per deciliter (mg/dl) one hour after having a sugary drink. Shelia's measured glucose level one hour after the sugary drink varies according to the Normal distribution with $\mu = 125$ mg/dl and $\sigma = 10$ mg/dl.

(a) If a single glucose measurement is made, what is the probability that Shelia is diagnosed as having gestational diabetes?
(b) If measurements are made on 4 separate days and the mean result is compared with the criterion 140 mg/dl, what is the probability that Shelia is diagnosed as having gestational diabetes?

11.29 Shelia's measured glucose level one hour after a sugary drink varies according to the Normal distribution with $\mu = 125$ mg/dl and $\sigma = 10$ mg/dl. What is the level L such that there is probability only 0.05 that the mean glucose level of 4 test results falls above L? (*Hint*: This requires a backward Normal calculation. See page 83 in Chapter 3 if you need to review.)

11.31 The number of accidents per week at a hazardous intersection varies with mean 2.2 and standard deviation 1.4. This distribution takes only whole-number values, so it is certainly not Normal.

(a) Let $\bar{x}$ be the mean number of accidents per week at the intersection during a year (52 weeks). What is the approximate distribution of $\bar{x}$ according to the central limit theorem?
(b) What is the approximate probability that $\bar{x}$ is less than 2?
(c) What is the approximate probability that there are fewer than 100 accidents at the intersection in a year? (*Hint*: Restate this event in terms of $\bar{x}$.)

11.33 Andrew plans to retire in 40 years. He plans to invest part of his retirement funds in stocks, so he seeks out information on past returns. He learns that over the entire 20th century, the real (that is, adjusted for inflation) annual returns on U.S. common stocks had mean 8.7% and standard deviation 20.2%. The distribution of annual returns on common stocks is roughly symmetric, so the mean return over even a moderate number of years is close to Normal. What is the probability (assuming that the past pattern of variation continues) that the mean annual return on common stocks over the next 40 years will exceed 10%? What is the probability that the mean return will be less than 5%? Follow the four-step process as illustrated in Example 11.8.

11.39 Unlike Joe (see the previous exercise) the operators of the numbers racket can rely on the law of large numbers. It is said that the New York City mobster Casper Holstein took as many as 25,000 bets per day in the Prohibition era. That's 150,000 bets in a week if he takes Sunday off. Casper's mean winnings per bet are $0.40 (he pays out 60 cents of each dollar bet to people like Joe and keeps the other 40 cents.) His standard deviation for single bets is about $18.96, the same as Joe's.

(a) What are the mean and standard deviation of Casper's average winnings $\overline{x}$ on his 150,000 bets?
(b) According to the central limit theorem, what is the approximate probability that Casper's average winnings per bet are between $0.30 and $0.50? After only a week, Casper can be pretty confident that his winnings will be quite close to $0.40 per bet.

Chapter 12 Problem Statements

12.9 Suppose that 10% of adults belong to health clubs, and 40% of these health club members go to the club at least twice a week. What percent of all adults go to a health club at least twice a week? Write the information given in terms of probabilities and use the general multiplication rule.

12.15 Continue your work from Exercise 12.13. What is the conditional probability that exactly 1 of the people will be allergic to peanuts or tree nuts, given that at least 1 of the 5 people suffers from one of these allergies?

12.27 New York State's “Quick Draw” lottery moves right along. Players choose between one and ten numbers from the range 1 to 80; 20 winning numbers are displayed on a screen every four minutes. If you choose just one number, your probability of winning is 20/80, or 0.25. Lester plays one number 8 times as he sits in a bar. What is the probability that all 8 bets lose?

12.29 Slot machines are now video games, with outcomes determined by random number generators. In the old days, slot machines were like this: you pull the lever to spin three wheels; each wheel has 20 symbols, all equally likely to show when the wheel stops spinning; the three wheels are independent of each other. Suppose that the middle wheel has 9 cherries among its 20 symbols, and the left and right wheels have 1 cherry each.

(a) You win the jackpot if all three wheels show cherries. What is the probability of winning the jackpot?
(b) There are three ways that the three wheels can show two cherries and one symbol other than a cherry. Find the probability of each of these ways.
(c) What is the probability that the wheels stop with exactly two cherries showing among them?

12.47 You are tossing a pair of balanced dice in a board game. Tosses are independent. You land in a danger zone that requires you to roll doubles (both faces show the same number of spots) before you are allowed to play again. How long will you wait to play again?

(a) What is the probability of rolling doubles on a single toss of the dice? (If you need review, the possible outcomes appear in Figure 10.2 (page 267). All 36 outcomes are equally likely.)
(b) What is the probability that you do not roll doubles on the first toss, but you do on the second toss?

(c) What is the probability that the first two tosses are not doubles and the third toss is doubles? This is the probability that the first doubles occurs on the third toss.

(d) Now you see the pattern. What is the probability that the first doubles occurs on the fourth toss? On the fifth toss? Give the general result: what is the probability that the first doubles occurs on the *k*th toss?

(*Comment*: The distribution of the number of trials to the first success is called a *geometric distribution*. In this problem you have found geometric distribution probabilities when the probability of a success on each trial is 1/6. The same idea works for any probability of success.)

Chapter 13 Problem Statements

13.5 Typing errors in a text are either nonword errors (as when “the” is typed as “teh”) or word errors that result in a real but incorrect word. Spell-checking software will catch nonword errors but not word errors. Human proofreaders catch 70% of word errors. You ask a fellow student to proofread an essay in which you have deliberately made 10 word errors.

(a) If the student matches the usual 70% rate, what is the distribution of the number of errors caught? What is the distribution of the number of errors missed?
(b) Missing 3 or more out of 10 errors seems a poor performance. What is the probability that a proofreader who catches 70% of word errors misses exactly 3 out of 10? If you use software, also find the probability of missing 3 or more out of 10.

13.11 A small liberal arts college would like to have an entering class of 415 students next year. Past experience shows that about 27% of the students admitted will decide to attend. The college therefore plans to admit 1535 students. Suppose that students make their decisions independently and that the probability is 0.27 that a randomly chosen student will accept the offer of admission.

(a) What are the mean and standard deviation of the number of students who accept the admissions offer from this college?
(b) Use the Normal approximation: what is the approximate probability that the college gets more students than they want?
(c) Use software to compute the exact probability that the college gets more students than they want. How good is the approximation in part (b)?

13.25 A believer in the random walk theory of stock markets thinks that an index of stock prices has probability 0.65 of increasing in any year. Moreover, the change in the index in any given year is not influenced by whether it rose or fell in earlier years. Let X be the number of years among the next 5 years in which the index rises.

(a) X has a binomial distribution. What are n and p?
(b) What are the possible values that X can take?
(c) Find the probability of each value of X. Draw a probability histogram for the distribution of X. (See Figure 13.2 for an example of a probability histogram.)
(d) What are the mean and standard deviation of this distribution? Mark the location of the mean on your histogram.

13.27 Many women take oral contraceptives to prevent pregnancy. Under ideal conditions, 1% of women taking the pill become pregnant within one year. In typical use, however, 5% become pregnant. Choose at random 20 women taking the pill. How many become pregnant in the next year?

(a) Explain why this is a binomial setting.
(b) What is the probability that at least one of the women becomes pregnant under ideal conditions? What is the probability in typical use?

13.29 A study of the effectiveness of oral contraceptives interviews a random sample of 500 women who are taking the pill.

(a) Based on the information about typical use in Exercise 13.27, what is the probability that at least 25 of these women become pregnant in the next year? (Check that the Normal approximation is permissible and use it to find this probability. If your software allows, find the exact binomial probability and compare the two results.)
(b) We can't use the Normal approximation to the binomial distributions to find this probability under ideal conditions as described in Exercise 13.27. Why not?

13.31 According to genetic theory, the blossom color in the second generation of a certain cross of sweet peas should be red or white in a 3:1 ratio. That is, each plant has probability 3/4 of having red blossoms, and the blossom colors of separate plants are independent.

(a) What is the probability that exactly 6 out of 8 of these plants have red blossoms?
(b) What is the mean number of red-blossomed plants when 80 plants of this type are grown from seeds?
(c) What is the probability of obtaining at least 60 red-blossomed plants when 80 plants are grown from seeds? Use the Normal approximation. If your software allows, find the exact binomial probability and compare the two results.

13.33 The Census Bureau says that 21% of Americans aged 18 to 24 do not have a high school diploma. A vocational school wants to attract young people who may enroll in order to achieve high school equivalency. The school mails an advertising flyer to 25,000 persons between the ages of 18 and 24.

(a) If the mailing list can be considered a random sample of the population, what is the mean number of high school dropouts who will receive the flyer?
(b) What is the approximate probability that at least 5000 dropouts will receive the flyer?

13.35 Here is a simple probability model for multiple-choice tests. Suppose that each student has probability p of correctly answering a question chosen at random from a universe of possible questions. (A strong student has a higher p than a weak student.) Answers to different questions are independent.

(a) Jodi is a good student for whom $p = 0.75$. Use the Normal approximation to find the probability that Jodi scores between 70% and 80% on a 100-question test.
(b) If the test contains 250 questions, what is the probability that Jodi will score between 70% and 80%? You see that Jodi's score on the longer test is more likely to be close to her "true score."

13.39 In 2007, Bob Jones University ended its fall semester a week early because of a whooping cough outbreak; 158 students were isolated and another 1200 given antibiotics as a precaution. Authorities react strongly to whooping cough outbreaks because the disease is so contagious. Because the effect of childhood vaccination often wears off by late adolescence, treat the Bob Jones students as if they were unvaccinated. It appears that about 1400 students were exposed. What is the probability that at least 75% of these students develop infections if not treated? (Fortunately, whooping cough is much less serious after infancy.)

13.41 We would like to find the probability that exactly 2 of the 20 exposed children in the previous exercise develop whooping cough.

(a) One way to get 2 infections is to get 1 among the 17 vaccinated children and 1 among the 3 unvaccinated children. Find the probability of exactly 1 infection among the 17 vaccinated children. Find the probability of exactly 1 infection among the 3 unvaccinated children. These events are independent: what is the probability of exactly 1 infection in each group?
(b) Write down all the ways in which 2 infections can be divided between the two groups of children. Follow the pattern of part (a) to find the probability of each of these possibilities. Add all of your results (including the result of part (a)) to obtain the probability of exactly 2 infections among the 20 children.

Chapter 14 Problem Statements

14.3 The critical value z^* for confidence level 97.5% is not in Table C. Use software or Table A of standard Normal probabilities to find z^*. Include in your answer a sketch like Figure 14.3 with C = 0.975 and your critical value z^* marked on the axis.

14.5 Here are the IQ test scores of 31 seventh-grade girls in a Midwest school district:

114	100	104	89	102	91	114	114	103	105	
108	130	120	132	111	128	118	119	86	72	
111	103	74	112	107	103	98	96	112	112	93

(a) These 31 girls are an SRS of all seventh-grade girls in the school district. Suppose that the standard deviation of IQ scores in this population is known to be $\sigma = 15$. We expect the distribution of IQ scores to be close to Normal. Make a stemplot of the distribution of these 31 scores (split the stems) to verify that there are no major departures from Normality. You have now checked the "simple conditions" to the extent possible.
(b) Estimate the mean IQ score for all seventh-grade girls in the school district, using a 99% confidence interval. Follow the four-step process as illustrated in Example 14.3.

14.13 The *P*-value for the first cola in Example 14.7 is the probability (taking the null hypothesis $\mu = 0$ to be true) that $\bar{x}$ takes a value at least as large as 0.3.

(a) What is the sampling distribution of $\bar{x}$ when $\mu = 0$? This distribution appears in Figure 14.6.
(b) Do a Normal probability calculation to find the *P*-value. Your result should agree with Example 14.7 up to roundoff error.

14.17 Exercise 14.7 describes 6 measurements of the electrical conductivity of a liquid. You stated the null and alternative hypotheses in Exercise 14.9.

(a) One set of measurements has mean conductivity $\bar{x}$ = 4.98. Enter this $\bar{x}$, along with the other required information, into the *P-Value of a Test of Significance* applet. What is the *P*-value? Is this outcome statistically significant at the $\alpha = 0.05$ level? At the $\alpha = 0.01$ level?
(b) Another set of measurements has $\bar{x}$ = 4.7. Use the applet to find the *P*-value for this outcome. Is it statistically significant at the $\alpha = 0.05$ level? At the $\alpha = 0.01$ level?
(c) Explain briefly why these *P*-values tell us that one outcome is strong evidence against the null hypothesis and that the other outcome is not.

14.19 Here are 6 measurements of the electrical conductivity of a liquid:

5.32 4.88 5.10 4.73 5.15 4.75

The liquid is supposed to have conductivity 5. Do the measurements give good evidence that the true conductivity is not 5?

The 6 measurements are an SRS from the population of all results we would get if we kept measuring conductivity forever. This population has a Normal distribution with mean equal to the true conductivity of the liquid and standard deviation 0.2. Use this information to carry out a test, following the four-step process as illustrated in Example 14.9.

14.21 A test of $H_0 : \mu = 1$ against $H_a : \mu > 1$ has test statistic z = 1.776. Is this test significant at the 5% level ($\alpha = 0.05$)? Is it significant at the 1% level ($\alpha = 0.01$)?

14.23 A random number generator is supposed to produce random numbers that are uniformly distributed on the interval from 0 to 1. If this is true, the numbers generated come from a population with $\mu = 0.5$ and $\sigma = 0.2887$. A command to generate 100 random numbers gives outcomes with mean $\bar{x}$ = 0.4365. Assume that the population σ remains fixed. We want to test

$$H_0 : \mu = 0.5$$
$$H_a : \mu \neq 0.5$$

(a) Calculate the value of the z test statistic.
(b) Use Table C: is z significant at the 5% level ($\alpha = 0.05$)?
(c) Use Table C: is z significant at the 1% level ($\alpha = 0.01$)?
(d) Between which two Normal critical values z* in the bottom row of Table C does z lie? Between what two numbers does the *P*-value lie? Does the test give good evidence against the null hypothesis?

14.35 Young men in North America and Europe (but not in Asia) tend to think they need more muscle to be attractive. One study presented 200 young American men with 100 images of men with various levels of muscle. Researchers measure level of muscle in kilograms per square meter (kg/m^2) of fat-free body mass. Typical young men have about 20 kg/m^2. Each subject chose two images, one that represented his own level of body muscle and one that he thought represented "what women prefer." The mean gap between self-image and "what women prefer" was 2.35 kg/m^2.

Suppose that the ``muscle gap" in the population of all young men has a Normal distribution with standard deviation 2.5 kg/m^2. Give a 90% confidence interval for the mean amount of muscle young men think they should add to be attractive to women. (They are wrong: women actually prefer a level close to that of typical men.)

14.41 If young men thought that their own level of muscle was about what women prefer, the mean "muscle gap" in the study described in Exercise 14.35 would be 0. We suspect (before seeing the data) that young men think women prefer more muscle than they themselves have.

(a) State null and alternative hypotheses for testing this suspicion.
(b) What is the value of the test statistic z?
(c) You can tell just from the value of z that the evidence in favor of the alternative is very strong (that is, the P-value is very small). Explain why this is true.

14.51 Breast-feeding mothers secrete calcium into their milk. Some of the calcium may come from their bones, so mothers may lose bone mineral. Researchers measured the percent change in mineral content of the spines of 47 mothers during three months of breast-feeding. Here are the data:

–4.7	–2.5	–4.9	–2.7	–0.8	–5.3	–8.3	–2.1	–6.8	–4.3
2.2	–7.8	–3.1	–1.0	–6.5	–1.8	–5.2	–5.7	–7.0	–2.2
–6.5	–1.0	–3.0	–3.6	–5.2	–2.0	–2.1	–5.6	–4.4	–3.3
–4.0	–4.9	–4.7	–3.8	–5.9	–2.5	–0.3	–6.2	–6.8	1.7
0.3	–2.3	0.4	–5.3	0.2	–2.2	–5.1			

(a) The researchers are willing to consider these 47 women as an SRS from the population of all nursing mothers. Suppose that the percent change in this population has standard deviation $\sigma = 2.5\%$. Make a stemplot of the data to see that they appear to follow a Normal distribution quite closely. (Don't forget that you need both a 0 and a –0 stem because there are both positive and negative values.)
(b) Use a 99 % confidence interval to estimate the mean percent change in the population.

14.53 Exercise 14.51 gives the percent change in the mineral content of the spine for 47 mothers during three months of nursing a baby. As in that exercise, suppose that the percent change in the population of all nursing mothers has a Normal distribution with standard deviation $\sigma = 2.5\%$. Do these data give good evidence that on the average nursing mothers lose bone mineral?

14.55 Athletes performing in bright sunlight often smear black eye grease under their eyes to reduce glare. Does eye grease work? In one study, 16 student subjects took a test of sensitivity to contrast after 3 hours facing into bright sun, both with and without eye grease. This is a matched pairs design. Here are the differences in sensitivity, with eye grease minus without eye grease:

0.07	0.64	–0.12	–0.05	–0.18	0.14	–0.16	0.03
0.05	0.02	0.43	0.24	–0.11	0.28	0.05	0.29

We want to know whether eye grease increases sensitivity on the average.

(a) What are the null and alternative hypotheses? Say in words what mean μ your hypotheses concern.

(b) Suppose that the subjects are an SRS of all young people with normal vision, that contrast differences follow a Normal distribution in this population, and that the standard deviation of differences is $\sigma = 0.22$. Carry out a test of significance.

Chapter 15 Problem Statements

15.5 Example 14.1 (page 360) described NHANES survey data on the body mass index (BMI) of 654 young women. The mean BMI in the sample was $\bar{x} = 26.8$. We treated these data as an SRS from a Normally distributed population with standard deviation $\sigma = 7.5$.

(a) Suppose that we had an SRS of just 100 young women. What would be the margin of error for 95% confidence?
(b) Find the margins of error for 95% confidence based on SRSs of 400 young women and 1600 young women.
(c) Compare the three margins of error. How does increasing the sample size change the margin of error of a confidence interval when the confidence level and population standard deviation remain the same?

15.9 Give a 95% confidence interval for the mean pH μ for each sample size in the previous exercise. The intervals, unlike the *P*-values, give a clear picture of what mean pH values are plausible for each sample.

15.11 Example 14.1 (page 360) assumed that the body mass index (BMI) of all American young women follows a Normal distribution with standard deviation $\sigma = 7.5$. How large a sample would be needed to estimate the mean BMI μ in this population to within ±1 with 95% confidence?

15.41 Software can generate samples from (almost) exactly Normal distributions. Here is a random sample of size 5 from the Normal distribution with mean 10 and standard deviation 2:

6.47	7.51	10.10	13.63	9.91

These data match the conditions for a *z* test better than real data will: the population is very close to Normal and has known standard deviation $\sigma = 2$, and the population mean is $\mu = 10$. Test the hypotheses

$$H_0 : \mu = 8$$
$$H_a : \mu \neq 8$$

(a) What are the *z* statistic and its *P*-value? Is the test significant at the 5% level?
(b) We know that the null hypothesis does not hold, but the test failed to give strong evidence against H_0. Explain why this is not surprising.

15.49 The previous exercise shows how to calculate the power of a one-sided *z* test. Power calculations for two-sided tests follow the same outline. We will find the power of a test based on 6 measurements of the conductivity of a liquid, reported in Exercise 15.13. The hypotheses are

$$H_0 : \mu = 5$$
$$H_a : \mu \neq 5$$

The population of all measurements is Normal with standard deviation $\sigma = 0.2$, and the alternative we hope to be able to detect is $\mu = 5.1$. (If you used the *Power of a Test* applet for Exercise 15.15, the two Normal curves for $n = 6$ illustrate parts (a) and (b) below.)

(a) Write the *z* test statistic in terms of the sample mean $\bar{x}$. For what values of *z* does this two-sided test reject H_0 at the 5 % significance level?

(b) Restate your result from part (a): what values of $\bar{x}$ lead to rejection of H_0?

(c) Now suppose that $\mu = 5.1$. What is the probability of observing an $\bar{x}$ that leads to rejection of H_0? This is the power of the test.

Chapter 16 Problem Statements

16.3 A university's financial aid office wants to know how much it can expect students to earn from summer employment. This information will be used to set the level of financial aid. The population contains 3478 students who have completed at least one year of study but have not yet graduated. The university will send a questionnaire to an SRS of 100 of these students, drawn from an alphabetized list.

(a) Describe how you will label the students in order to select the sample.
(b) Use Table B, beginning at line 105, to select the first 5 students in the sample.
(c) What is the response variable in this study?

16.5 Elephants sometimes damage crops in Africa. It turns out that elephants dislike bees. They recognize beehives in areas where they are common and avoid them. Can this be used to keep elephants away from trees? A group in Kenya placed active beehives in some trees and empty beehives in others. Will elephant damage be less in trees with hives? Will even empty hives keep elephants away?

(a) Outline the design of an experiment to answer these questions using 72 acacia trees (be sure to include a control group).
(b) Use software or the *Simple Random Sample* to choose the trees for the active-hive group, or Table B at line 137 to choose the first 4 trees in that group.
(c) What is the response variable in this experiment?

16.13 The distribution of blood cholesterol level in the population of young men aged 20 to 34 years is close to Normal with standard deviation $\sigma = 41$ milligrams per deciliter (mg/dl). You measure the blood cholesterol of 14 cross-country runners. The mean level is $\bar{x}$ =172 mg/dl. Assuming that σ is the same as in the general population, give a 90% confidence interval for the mean level μ among cross-country runners.

16.15 How large a sample is needed to cut the margin of error in Exercise 16.13 in half? How large a sample is needed to cut the margin of error to ±5 mg/dl?

16.17 The level of pesticides found in the blubber of whales is a measure of pollution of the oceans by runoff from land and can also be used to identify different populations of whales. A sample of 8 male minke whales in the West Greenland area of the North Atlantic found the mean concentration of the insecticide dieldrin to be $\bar{x}$ =357 nanograms per gram of blubber (ng/g). Suppose that the concentration in all such whales varies Normally with standard deviation $\sigma = 50$ ng/g. Use a 95% confidence interval to estimate the mean level. Be sure to state your conclusion in plain language.

16.19 Use the information in Exercise 16.17 to give an 80% confidence interval and a 90% confidence interval for the mean concentration of dieldrin in the whale population. What general fact about confidence intervals do the margins of error of your three intervals illustrate?

16.21 IQ tests are scaled so that the mean score in a large population should be $\mu = 100$. We suspect that the very-low-birth-weight population has mean score less than 100. Does the study described in the previous exercise give good evidence that this is true? State hypotheses, carry out a test assuming that the ``simple conditions" (page 360) hold, and give your conclusion in plain language.

16.27 The time that people require to react to a stimulus usually has a right-skewed distribution, as lack of attention or tiredness causes some lengthy reaction times. Reaction times for children with attention deficit hyperactivity disorder (ADHD) are more skewed, as their condition causes more frequent lack of attention. In one study, children with ADHD were asked to press the spacebar on a computer keyboard when any letter other than X appeared on the screen. With 2 seconds between letters, the mean reaction time was 445 milliseconds (ms) and the standard deviation was 82 ms. Take these values to be the population μ and σ for ADHD children.

(a) What are the mean and standard deviation of the mean reaction time $\overline{x}$ for a randomly chosen group of 15 ADHD children? For a group of 150 such children?
(b) The distribution of reaction time is strongly skewed. Explain briefly why we hesitate to regard $\overline{x}$ as Normally distributed for 15 children but are willing to use a Normal distribution for the mean reaction time of 150 children.
(c) What is the approximate probability that the mean reaction time in a group of 150 ADHD children is greater than 450 ms?

16.45 Here are the daily average body temperatures (degrees Fahrenheit) for 20 healthy adults:

98.74	98.83	96.80	98.12	97.89	98.09	97.87	97.42	97.30	97.84
100.27	97.90	99.64	97.88	98.54	98.33	97.87	97.48	98.92	98.33

(a) Make stemplot of the data. The distribution is roughly symmetric and single-peaked. There is one mild outlier. We expect the distribution of the sample mean $\overline{x}$ to be close to Normal.
(b) Do these data give evidence that the mean body temperature for all healthy adults is not equal to the traditional 98.6 degrees? Follow the four-step process for significance tests (page 378). (Suppose that body temperature varies Normally with standard deviation 0.7 degree.)

16.47 Use the data in Exercise 16.45 to estimate mean body temperature with 90% confidence. Follow the four-step process for confidence intervals (page 366).

Chapter 17 Problem Statements

17.3 Use Table C or software to find

(a) the critical value for a one-sided test with level $\alpha = 0.05$ based on the *t*(5) distribution.
(b) the critical value for a 98% confidence interval based on the *t*(21) distribution.

17.7 The composition of the earth's atmosphere may have changed over time. To try to discover the nature of the atmosphere long ago, we can examine the gas in bubbles inside ancient amber. Amber is tree resin that has hardened and been trapped in rocks. The gas in bubbles within amber should be a sample of the atmosphere at the time the amber was formed. Measurements on specimens of amber from the late Cretaceous era (75 to 95 million years ago) give these percents of nitrogen:

63.4 65.0 64.4 63.3 54.8 64.5 60.8 49.1 51.0

Assume (this is not yet agreed on by experts) that these observations are an SRS from the late Cretaceous atmosphere. Use a 90% confidence interval to estimate the mean percent of nitrogen in ancient air. Follow the four-step process as illustrated in Example 17.2.

17.9 The one-sample *t* statistic from a sample of *n* = 25 observations for the two-sided test of

$$H_0 : \mu = 64$$
$$H_a : \mu \neq 64$$

has the value *t* = 1.12.

(a) What are the degrees of freedom for *t*?
(b) Locate the two critical values *t** from Table C that bracket *t*. What are the two-sided *P*-values for these two entries?
(c) Is the value *t* = 1.12 statistically significant at the 10% level? At the 5% level?

17.11 The usual way to study the brain's response to sounds is to have subjects listen to "pure tones." The response to recognizable sounds may differ. To compare responses, researchers anesthetized macaque monkeys. They fed pure tones and also monkey calls directly to their brains by inserting electrodes. Response to the stimulus was measured by the firing rate (electrical spikes per second) of neurons in various areas of the brain. Table 17.2 contains the responses for 37 neurons. Researchers suspected that the response to monkey calls would be stronger than the response to a pure tone. Do the data support this idea?

Table 17.2 Neuron response to tones and monkey calls								
Neuron	Tone	Call	Neuron	Tone	Call	Neuron	Tone	Call
1	474	500	14	145	42	26	71	134
2	256	138	15	141	241	27	68	65
3	241	485	16	129	294	28	59	182
4	226	338	17	113	123	29	59	97
5	185	194	18	112	182	30	57	318
6	174	159	19	102	141	31	56	201
7	176	341	20	100	118	32	47	279
8	168	85	21	74	62	33	46	62
9	161	303	22	72	112	34	41	84
10	150	208	23	20	193	35	26	203
11	19	66	24	21	129	36	28	192
12	20	54	25	26	135	37	31	70
13	35	103						

Complete the *Plan*, *Solve*, and *Conclude* steps of the four-step process, following the model of Example 17.4.

17.13 A group of earth scientists studied the small diamonds found in a nodule of rock carried up to the earth's surface in surrounding rock. This is an opportunity to examine a sample from a single population of diamonds formed in a single event deep in the earth. Table 17.3 (page 460) presents data on the nitrogen content (parts per million) and the abundance of carbon-13 in these diamonds.

Table 17.3 Nitrogen and carbon-13 in a sample of diamonds					
Diamond	Nitrogen (ppm)	Carbon-13 Ratio	Diamond	Nitrogen (ppm)	Carbon-13 Ratio
1	487	–2.78	13	273	–2.73
2	1430	–1.39	14	94	–3.57
3	60	–4.26	15	69	–3.83
4	244	–1.19	16	262	–2.04
5	196	–2.12	17	120	–2.82
6	274	–2.87	18	302	–0.84
7	41	–3.68	19	75	–3.57
8	54	–3.29	20	242	–2.42
9	473	–3.79	21	115	–3.89
10	30	–4.06	22	65	–3.87
11	98	–1.83	23	311	–1.58
12	41	–4.03	24	61	–3.97

(Carbon has several isotopes, forms with different numbers of neutrons in the nuclei of their atoms. Carbon-12 makes up almost 99% of natural carbon. The abundance of carbon-13 is measured by the ratio of carbon-13 to carbon-12, in parts per thousand more or less than a standard. The minus signs in the data mean that the ratio is smaller in these diamonds than in standard carbon.)

We would like to estimate the mean abundance of both nitrogen and carbon-13 in the population of diamonds represented by this sample. Examine the data for nitrogen. Can we use a *t* confidence interval for mean nitrogen? Explain your answer. Give a 95% confidence interval if you think the result can be trusted.

17.25 You read in the report of a psychology experiment: "Separate analyses for our two groups of 12 participants revealed no overall placebo effect for our student group (mean = 0.08, SD = 0.37, $t(11) = 0.49$) and a significant effect for our non-student group (mean = 0.35, SD = 0.37, $t(11) = 3.25$, $p < 0.01$)." The null hypothesis is that the mean effect is zero. What are the correct values of the two *t* statistics based on the means and standard deviations? Compare each correct *t*-value with the critical values in Table C. What can you say about the two-sided *P*-value in each case?

17.27 The Trial Urban District Assessment (TUDA) is a government-sponsored study of student achievement in large urban school districts. TUDA gives a reading test scored from 0 to 500. A score of 243 is a "basic" reading level and a score of 281 is "proficient." Scores for a random sample of 1470 eighth-graders in Atlanta had $\bar{x} = 240$ with standard error 1.1.

(a) We don't have the 1470 individual scores, but use of the *t* procedures is surely safe. Why?
(b) Give a 99% confidence interval for the mean score of all Atlanta eighth-graders. (Be careful: the report gives the standard error of $\bar{x}$, not the standard deviation *s*.)
(c) Urban children often perform below the basic level. Is there good evidence that the mean for all Atlanta eighth-graders is less than the basic level?

17.29 The placebo effect is particularly strong in patients with Parkinson's disease. To understand the workings of the placebo effect, scientists measure activity at a key point in the brain when patients receive a placebo that they think is an active drug and also when no treatment is given. The same six patients are measured both with and without the placebo, at different times.

(a) Explain why the proper procedure to compare the mean response to placebo with control (no treatment) is a matched pairs *t* test.
(b) The six differences (treatment minus control) had $\bar{x} = -0.326$ and $s = 0.181$. Is there significant evidence of a difference between treatment and control?

17.31 Blissymbols are pictographs (think of Egyptian hieroglyphics) sometimes used to help learning-disabled children. In a study of computer-assisted learning, 12 normal-ability schoolchildren were assigned at random to each of four computer learning programs. After they used the program, they attempted to recognize 24 Blissymbols. Here are the counts correct for one of the programs:

12 22 9 14 20 15 9 10 11 11 15 6

(a) Make a stemplot (split the stems). Are there outliers or strong skewness that would forbid use of the *t* procedures?
(b) Give a 90% confidence interval for the mean count correct among all children of this age who use the program.

17.33 Our bodies have a natural electrical field that is known to help wounds heal. Does changing the field strength slow healing? A series of experiments with newts investigated this question. In one experiment, the two hind limbs of 12 newts were assigned at random to either experimental or control groups. This is a matched pairs design. The electrical field in the experimental limbs was reduced to zero by applying a voltage. The control limbs were left alone. Here are the rates at which new cells closed a razor cut in each limb, in micrometers per hour:

Newt	1	2	3	4	5	6	7	8	9	10	11	12
Control limb	36	41	39	42	44	39	39	56	33	20	49	30
Experimental limb	28	31	27	33	33	38	45	25	28	33	47	23

(a) Make a stemplot of the differences between limbs of the same newt (control limb minus experimental limb). There is a high outlier.
(b) A good way to judge the effect of an outlier is to do your analysis twice, once with the outlier and a second time without it. Carry out two *t* tests to see if the mean healing rate is significantly lower in the experimental limbs, one including all 12 newts and another that omits the outlier. What are the test statistics and their *P*-values? Does the outlier have a strong influence on your conclusion?

17.35 Here's a new idea for treating advanced melanoma, the most serious kind of skin cancer. Genetically engineer white blood cells to better recognize and destroy cancer cells, then infuse these cells into patients. The subjects in a small initial study were 11 patients whose melanoma had not responded to existing treatments. One question was how rapidly the new cells would multiply after infusion, as measured by the doubling time in days. Here are the doubling times:

1.4 1.0 1.3 1.0 1.3 2.0 0.6 0.8 0.7 0.9 1.9

(a) Examine the data. Is it reasonable to use the *t* procedures?
(b) Give a 90% confidence interval for the mean doubling time. Are you willing to use this interval to make an inference about the mean doubling time in a population of similar patients?

17.37 The concentration of carbon dioxide (CO_2) in the atmosphere is increasing rapidly due to our use of fossil fuels. Because plants use CO_2 to fuel photosynthesis, more CO_2 may cause trees and other plants to grow faster. An elaborate apparatus allows researchers to pipe extra CO_2 to a 30-meter circle of forest. They selected two nearby circles in each of three parts of a pine forest and randomly chose one of each pair to receive extra CO_2. The response variable is the mean increase in base area for 30 to 40 trees in a circle during a growing season. We measure this in percent increase per year. The following are one year's data.

Pair	Control plot	Treatment plot
1	9.752	10.587
2	7.263	9.244
3	5.742	8.675

(a) State the null and alternative hypotheses. Explain clearly why the investigators used a one-sided alternative.
(b) Carry out a test and report your conclusion in simple language.
(c) The investigators used the test you just carried out. Any use of the *t* procedures with samples this size is risky. Why?

17.39 Velvetleaf is a particularly annoying weed in cornfields. It produces lots of seeds, and the seeds wait in the soil for years until conditions are right. How many seeds do velvetleaf plants produce? Here are counts from 28 plants that came up in a cornfield when no herbicide was used:

2450	2504	2114	1110	2137	8015	1623	1531	2008	1716
721	863	1136	2819	1911	2101	1051	218	1711	164
2228	363	5973	1050	1961	1809	130	880		

We would like to give a confidence interval for the mean number of seeds produced by velvetleaf plants. Alas, the *t* interval can't be safely used for these data. Why not?

17.43 Give a 90% confidence interval for the difference in healing rates (control minus experimental) in the previous exercise.

17.45 The design of controls and instruments affects how easily people can use them. Timothy Sturm investigated this effect in a course project, asking 25 right-handed students to turn a knob (with their right hands) that moved an indicator by screw action. There were two identical instruments, one with a right-hand thread (the knob turns clockwise) and the other with a left-hand thread (the knob turns counterclockwise). Table 17.5 gives the times in seconds each subject took to move the indicator a fixed distance.

Table 17.5 Performance times (seconds) using right-hand and left-hand threads

Subject	Right Thread	Left Thread	Subject	Right Thread	Left Thread
1	113	137	14	107	87
2	105	105	15	118	166
3	130	133	16	103	146
4	101	108	17	111	123
5	138	115	18	104	135
6	118	170	19	111	112
7	87	103	20	89	93
8	116	145	21	78	76
9	75	78	22	100	116
10	96	107	23	89	78
11	122	84	24	85	101
12	103	148	25	88	123
13	116	147			

(a) Each of the 25 students used both instruments. Explain briefly how you would use randomization in arranging the experiment.
(b) The project hoped to show that right-handed people find right-hand threads easier to use. Do an analysis that leads to a conclusion about this issue.

17.47 Give a 90% confidence interval for the mean time advantage of right-hand over left-hand threads in the setting of Exercise 17.45. Do you think that the time saved would be of practical importance if the task were performed many times—for example, by an assembly-line worker? To help answer this question, find the mean time for right-hand threads as a percent of the mean time for left-hand threads.

Chapter 18 Problem Statements

18.5 "Conservationists have despaired over destruction of tropical rain forest by logging, clearing, and burning." These words begin a report on a statistical study of the effects of logging in Borneo. Here are data on the number of tree species in 12 unlogged forest plots and 9 similar plots logged 8 years earlier:

Unlogged	22	18	22	20	15	21	13	13	19	13	19	15
Logged	17	4	18	14	18	15	15	10	12			

(a) The study report says, "Loggers were unaware that the effects of logging would be assessed." Why is this important? The study report also explains why the plots can be considered to be randomly assigned.
(b) Does logging significantly reduce the mean number of species in a plot after 8 years? Follow the four-step process as illustrated in Examples 18.2 and 18.3.

18.7 Use the data in Exercise 18.5 to give a 90% confidence interval for the difference in mean number of species between unlogged and logged plots.

18.9 Businesses know that customers often respond to background music. Do they also respond to odors? One study of this question took place in a small pizza restaurant in France on Saturday evenings in May. On one of these evenings, a relaxing lavender odor was spread through the restaurant. Table 18.2 gives the time (minutes) that two samples of 30 customers spent in the restaurant and the amount they spent (in euros). The two evenings were comparable in many ways (weather, customer count, and so on), so we are willing to regard the data as independent SRSs from spring Saturday evenings at this restaurant. The authors say, "Therefore at this stage it would be impossible to generalize the results to other restaurants."

(a) Does a lavender odor encourage customers to stay longer in the restaurant? Examine the time data and explain why they are suitable for two-sample *t* procedures. Use the two-sample *t* test to answer the question posed.
(b) Does a lavender odor encourage customers to spend more while in the restaurant? Examine the spending data. In what ways do these data deviate from Normality? With 30 observations, the *t* procedures are nonetheless reasonably accurate. Use the two-sample *t* test to answer the question posed.

Table 18.2 Time (minutes) and spending (Euros) by restaurant customers

No odor		Lavender	
Minutes	Euros	Minutes	Euros
103	15.9	92	21.9
68	18.5	126	18.5
79	15.9	114	22.3
106	18.5	106	21.9
72	18.5	89	18.5
121	21.9	137	24.9
92	15.9	93	18.5
84	15.9	76	22.5
72	15.9	98	21.5
92	15.9	108	21.9
85	15.9	124	21.5
69	18.5	105	18.5
73	18.5	129	25.5
87	18.5	103	18.5
109	20.5	107	18.5
115	18.5	109	21.9
91	18.5	94	18.5
84	15.9	105	18.5
76	15.9	102	24.9
96	15.9	108	21.9
107	18.5	95	25.9
98	18.5	121	21.9
92	15.9	109	18.5
107	18.5	104	18.5
93	15.9	116	22.8
118	18.5	88	18.5
87	15.9	109	21.9
101	25.5	97	20.7
75	12.9	101	21.9
86	15.9	106	22.5

18.25 Equip male and female students with a small device that secretly records sound for a random 30 seconds during each 12.5-minute period over two days. Count the words each subject speaks during each recording period, and from this, estimate how many words per day each subject speaks. The published report includes a table summarizing six such studies. Here are two of the six:

	Sample Size		Estimated Average Number (SD) Of Words Spoken per Day	
Study	Women	Men	Women	Men
1	56	56	16,177 (7520)	16,569 (9108)
2	27	20	16,496 (7914)	12,867 (8343)

Readers are supposed to understand that, for example, the 56 women in the first study had $\bar{x} = 16{,}177$ and s = 7520. It is commonly thought that women talk more than men. Does either of the two samples support this idea? For each study:

(a) State hypotheses in terms of the population means for men (μ_M) and women (μ_F).
(b) Find the two-sample t statistic.
(c) What degrees of freedom does Option 2 use to get a conservative P -value?
(d) Compare your value of t with the critical values in Table C. What can you say about the P-value of the test?
(e) What do you conclude from the results of these two studies?

18.27 In a study of the presence of whelks along the Pacific coast, investigators put down a frame that covers 0.25 square meter and counted the whelks on the sea bottom inside the frame. They did this at 7 locations in California and 6 locations in Oregon. The report says that whelk densities "were twice as high in Oregon as in California (mean ± SEM, 26.9±1.56 versus 11.9±2.68 whelks per 0.25 m^2, Oregon versus California, respectively; Student's t test, $P < 0.001$)."

(a) SEM stands for the standard error of the mean, $s/\sqrt{n}$. Fill in the values in this summary table:

Group	Location	n	$\bar{x}$	s
1	Oregon	?	?	?
2	California	?	?	?

(b) What degrees of freedom would you use in the conservative two-sample t procedures to compare Oregon and California?
(c) What is the two-sample t test statistic for comparing the mean densities of whelks in Oregon and California?
(d) Test the null hypothesis of no difference between the two population means against the two-sided alternative. Use your statistic from part (c) with degrees of freedom from part (b). Does your conclusion agree with the published report?

18.29 The post-lunch dip is the drop in mental alertness after a midday meal. Does an extract of the leaves of the ginkgo tree reduce the post-lunch dip? Assign healthy people aged 18 to 40 to take either ginkgo extract or a placebo pill. After lunch, ask them to read seven pages of random letters and place an X over every e. Count the number of misses per line read.

(a) What is a placebo and why was one group given a placebo?
(b) What is the double-blind method and why should it be used in this experiment?
(c) Here are summaries of performance after 13 weeks of either ginkgo extract or placebo:

Group	Group size	Mean	Std. dev.
Ginkgo	21	0.06383	0.01462
Placebo	18	0.05342	0.01549

Is there a significant difference between the two groups? What do these data show about the effect of ginkgo extract?

18.31 The SAFE (Social, Attitudinal, Familial, and Environmental Stress) scale measures the stress level of adults adjusting to a different culture. Scores range from 1 (not stressful) to 5 (extremely stressful). In a study of stress among immigrant mothers in a university community, mothers of children between 2 and 10 years of age whose families had come to the United States for professional or academic reasons took the SAFE questionnaire. Here are summaries for mothers from Asia and Europe:

Origin	Sample size	Mean	Std. dev
Asian	12	1.92	0.60
European	9	1.74	0.57

Is there evidence of a difference in mean stress levels between mothers from Asia and Europe?

18.33 What we really want to know is whether coached students improve more than uncoached students, and whether any advantage is large enough to be worth paying for. Use the information in the previous exercise to answer these questions:

(a) Is there good evidence that coached students gained more on the average than uncoached students?
(b) How much more do coached students gain on the average? Give a 99% confidence interval.
(c) Based on your work, what is your opinion: do you think coaching courses are worth paying for?

18.35 Here are the IQ test scores of 31 seventh-grade girls in a Midwest school district.

114	100	104	89	102	91	114	114	103	105
108	130	120	132	111	128	118	119	86	72

111	103	74	112	107	103	98	96	112	112	93

The IQ test scores of 47 seventh-grade boys in the same district are

111	107	100	107	115	111	97	112	104	106	113
109	113	128	128	118	113	124	127	136	106	123
124	126	116	127	119	97	102	110	120	103	115
93	123	79	119	110	110	107	105	105	110	77
90	114	106								

(a) Make stemplots or histograms of both sets of data. Because the distributions are reasonably symmetric with no extreme outliers, the *t* procedures will work well.
(b) Treat these data as SRSs from all seventh-grade students in the district. Is there good evidence that girls and boys differ in their mean IQ scores?

18.37 Use the data in Exercise 18.35 to give a 95% confidence interval for the difference between the mean IQ scores of all boys and all girls in the district.

18.39 Of course, the reason for durable press treatment is to reduce wrinkling. ``Wrinkle recovery angle" measures how well a fabric recovers from wrinkles. Higher is better. Here are data on the wrinkle recovery angle (in degrees) for the same fabric swatches discussed in the previous exercise:

Permafresh	136	135	132	137	134
Hylite	143	141	146	141	145

Is there a significant difference in wrinkle resistance?

(a) Do the sample means suggest that one process has better wrinkle resistance?
(b) Make stemplots for both samples. There are no obvious deviations from Normality.
(c) Test the hypothesis $H_0 : \mu_1 = \mu_2$ against the two-sided alternative. What do you conclude from part (a) and from the result of your test?

18.41 In Exercise 18.39, you found that the Hylite process results in significantly greater wrinkle resistance than the Permafresh process. How large is the difference in mean wrinkle recovery angle? Give a 90% confidence interval.

18.45 Kathleen Vohs of the University of Minnesota and her coworkers carried out several randomized comparative experiments on the effects of thinking about money. Here's part of one such experiment. Ask student subjects to unscramble 30 sets of five words to make a meaningful phrase from four of the five words. The control group unscrambled phrases like “cold it desk outside is” into “it is cold outside.” The treatment group unscrambled phrases that lead to thinking about money, turning "high a salary desk paying" into "a high-paying salary." Then each subject worked a hard puzzle, knowing that he or she could ask for help. Here are the times in seconds until subjects asked for help, for the treatment group,

609	444	242	199	174	55	251	466	443
531	135	241	476	482	362	69	160	

and for the control group,

118	272	413	291	140	104	55	189	126
400	92	64	88	142	141	373	156	

The researchers suspected that money is connected with self-sufficiency, so that the treatment group will ask for help less quickly on the average. Do the data support this idea?

18.47 A "subliminal" message is below our threshold of awareness but may nonetheless influence us. Can subliminal messages help students learn math? A group of students who had failed the mathematics part of the City University of New York Skills Assessment Test agreed to participate in a study to find out.

All received a daily subliminal message, flashed on a screen too rapidly to be consciously read. The treatment group of 10 students (chosen at random) was exposed to "Each day I am getting better in math." The control group of 8 students was exposed to a neutral message, "People are walking on the street." All students participated in a summer program designed to raise their math skills, and all took the assessment test again at the end of the program. Table 18.3 gives data on the subjects' scores before and after the program. Is there good evidence that the treatment brought about a greater improvement in math scores than the neutral message? How large is the mean difference in gains between treatment and control? (Use 90% confidence.)

Table 18.3 Mathematics skills scores before and after a subliminal message

Treatment group		**Control group**	
Before	**After**	**Before**	**After**
18	24	18	29
18	25	24	29
21	33	20	24
18	29	18	26
18	33	24	38
20	36	22	27
23	34	15	22
23	36	19	31
21	34		
17	27		

Chapter 19 Problem Statements

19.5 Canada has much stronger gun control laws than the United States, and Canadians support gun control more strongly than do Americans. A sample survey asked a random sample of 1505 adult Canadians, "Do you agree or disagree that all firearms should be registered?" Of the 1505 people in the sample, 1288 answered either "Agree strongly" or "agree somewhat"

(a) The survey dialed residential telephone numbers at random in all ten Canadian provinces (omitting the sparsely populated northern territories). Based on what you know about sample surveys, what is likely to be the biggest weakness in this survey?
(b) Nonetheless, act as if we have an SRS from adults in the Canadian provinces. Give a 95% confidence interval for the proportion who support registration of all firearms.

19.7 Sample surveys usually contact large samples, so we can use the large-sample confidence interval if the sample design is close to an SRS. Scientific studies often use small samples that require the plus four method. For example, the small round holes you often see in sea shells were drilled by other sea creatures, who ate the former owners of the shells. Whelks often drill into mussels, but this behavior appears to be more or less common in different locations. Investigators collected whelk eggs from the coast of Oregon, raised the whelks in the laboratory, and then put each whelk in a container with some delicious mussels. Only 9 of 98 whelks drilled into mussels.

(a) Why can't we use the large-sample confidence interval for the proportion p of Oregon whelks that will spontaneously drill mussels?
(b) The plus four method adds four observations, two successes and two failures. What are the sample size and the number of successes after you do this? What is the plus four estimate $\tilde{p}$ of p?
(c) Give the plus four 90% confidence interval for the proportion of Oregon whelks that will spontaneously drill mussels.

19.9 The plus four method is particularly useful when there are *no* successes or *no* failures in the data. The study of Spanish currency described in Example 19.5 found that in Seville, all 20 of a sample of 20 euro bills had cocaine traces.

(a) What is the sample proportion $\hat{p}$ of contaminated bills? What is the large-sample 95% confidence interval for p? It's not plausible that *every* bill in Seville has cocaine traces, as this interval says.
(b) Find the plus four estimate $\tilde{p}$ and the plus four 95% confidence interval for p. These results are more reasonable.

19.11 PTC is a substance that has a strong bitter taste for some people and is tasteless for others. The ability to taste PTC is inherited. About 75% of Italians can taste PTC, for example. You want to estimate the proportion of Americans with at least one Italian grandparent who can taste PTC. Starting with the 75% estimate for Italians, how large a sample must you collect in order to estimate the proportion of PTC tasters within ± 0.04 with 90% confidence?

19.13 We often judge other people by their faces. It appears that some people judge candidates for elected office by their faces. Psychologists showed head-and-shoulders photos of the two main candidates in 32 races for the U.S. Senate to many subjects (dropping subjects who recognized one of the candidates) to see which candidate was rated “more competen” based on nothing but the photos. On election day, the candidates whose faces looked more competent won 22 of the 32 contests. If faces don't influence voting, half of all races in the long run should be won by the candidate with the better face. Is there evidence that the candidate with the better face wins more than half the time? Follow the four-step process as illustrated in Example 19.7.

19.25 The Harris Poll asked a sample of smokers, “Do you believe that smoking will probably shorten your life, or not” Of the 1010 people in the sample, 848 said “Yes.”

(a) Harris called residential telephone numbers at random in an attempt to contact an SRS of smokers. Based on what you know about national sample surveys, what is likely to be the biggest weakness in the survey?
(b) We will nonetheless act as if the people interviewed are an SRS of smokers. Give a 95% confidence interval for the percent of smokers who agree that smoking will probably shorten their lives.

19.29 The Pew Research Center asked a random sample of 1128 adult women, “How satisfied are you with your life overall?” Of these women, 56 said either “Mostly dissatisfied'” or “Very dissatisfied.”

(a) Pew dialed residential telephone numbers at random in the continental United States in an attempt to contact a random sample of adults. Based on what you know about national sample surveys, what is likely to be the biggest weakness in the survey?
(b) Act as if the sample is an SRS. Give a large-sample 90% confidence interval for the proportion p of all adult women who are mostly or very dissatisfied with their lives.
(c) Give the plus four confidence interval for p. If you express the two confidence intervals in percents and round to the nearest tenth of a percent, how do they differ? (As always, the plus four method pulls results away from 0% or 100%, whichever is closer. Although the condition for the large-sample interval is met, we can place more trust in the plus four interval.)

19.31 Most soybeans grown in the United States are genetically modified to, for example, resist pests and so reduce use of pesticides. Because some nations do not accept genetically modified (GM) foods, grain-handling facilities routinely test soybean shipments for the presence of GM beans. In a study of the accuracy of these tests, researchers submitted shipments of soybeans containing 1% of GM beans to 23 randomly selected facilities. Eighteen detected the GM beans.

(a) Show that the conditions for the large-sample confidence interval are not met. Show that the conditions for the plus four interval are met.
(b) Use the plus four method to give a 90% confidence interval for the percent of all grain-handling facilities that will correctly detect 1% of GM beans in a shipment.

19.37 Some shrubs have the useful ability to resprout from their roots after their tops are destroyed. Fire is a particular threat to shrubs in dry climates, as it can injure the roots as well as destroy the aboveground material. One study of resprouting took place in a dry area of Mexico. The investigators clipped the tops of samples of several species of shrubs. In some cases, they also applied a propane torch to the stumps to simulate a fire. Of 12 specimens of the shrub *Krameria cytisoides*, 5 resprouted after fire. Estimate with 90% confidence the proportion of all shrubs of this species that will resprout after fire.

19.39 A sample survey funded by the National Science Foundation asked a random sample of American adults about biological evolution. One question asked subjects to answer ”True,” “False,” or “Not sure” to the statement “Human beings, as we know them today, developed from earlier species of animals.” Of the 1484 respondents, 594 said “True.” What can you say with 95% confidence about the percent of all American adults who think that humans developed from earlier species of animals?

19.41 Does the sample in Exercise 19.39 give good evidence to support the claim “Fewer than half of American adults think that humans developed from earlier species of animals”?

Chapter 20 Problem Statements

20.1 Younger people use online instant messaging (IM) more often than older people. A random sample of IM users found that 73 of the 158 people in the sample aged 18 to 27 said they used IM more often than email. In the 28 to 39 age group, 26 of 143 people used IM more often than email. Give a 95% confidence interval for the difference between the proportions of IM users in these age groups who use IM more often than email. Follow the four-step process as illustrated in Examples 20.1 and 20.2.

20.3 A government survey randomly selected 6889 female high school students and 7028 male high school students. Of these students, 1915 females and 3078 males met recommended levels of physical activity. (These levels are quite high: at least 60 minutes of activity that makes you breathe hard on at least 5 of the past 7 days.) Give a 99% confidence interval for the difference between the proportions of all female and male high school students who meet the recommended levels of activity.

20.5 We don't like to find broken crackers when we open the package. How can makers reduce breaking? One idea is to microwave the crackers for 30 seconds right after baking them. Breaks start as hairline cracks called "checking." Assign 65 newly baked crackers to the microwave and another 65 to a control group that is not microwaved. After one day, none of the microwave group and 16 of the control group show checking. Give the 95% plus four confidence interval for the amount by which microwaving reduces the proportion of checking. The plus four method is particularly helpful when, as here, a count of successes is zero. Follow the four-step process as illustrated in Example 20.3.

20.7 Most alpine skiers and snowboarders do not use helmets. Do helmets reduce the risk of head injuries? A study in Norway compared skiers and snowboarders who suffered head injuries with a control group who were not injured. Of 578 injured subjects, 96 had worn a helmet. Of the 2992 in the control group, 656 wore helmets. Is helmet use less common among skiers and snowboarders who have head injuries? Follow the four-step process as illustrated in Example 20.5. (Note that this is an observational study that compares injured and uninjured subjects. An experiment that assigned subjects to helmet and no-helmet groups would be more convincing.)

20.17 Many teens have posted profiles on sites such as MySpace. A sample survey asked random samples of teens with online profiles if they included false information in their profiles. Of 170 younger teens (ages 12 to 14), 117 said "Yes." Of 317 older teens (ages 15 to 17), 152 said "Yes."

(a) Do these samples satisfy the guidelines for the large-sample confidence interval?

(b) Give a 95% confidence interval for the difference between the proportions of younger and older teens who include false information in their online profiles.

20.19 Genetic influences on cancer can be studied by manipulating the genetic makeup of mice. One of the processes that turn genes on or off (so to speak) in particular locations is called "DNA methylation." Do low levels of this process help cause tumors? Compare mice altered to have low levels with normal mice. Of 33 mice with lowered levels of DNA methylation, 23 developed tumors. None of the control group of 18 normal mice developed tumors in the same time period.

(a) Explain why we cannot safely use either the large-sample confidence interval or the test for comparing the proportions of normal and altered mice that develop tumors.
(b) The plus four method adds two observations, a success and a failure, to each sample. What are the sample sizes and the numbers of mice with tumors after you do this? Give a plus four 99% confidence interval for the difference in the proportions of the two populations that develop tumors.
(c) Based on your confidence interval, is the difference between normal and altered mice significant at the 1% level?

20.21 Is there a significant difference in the proportions of papers with and without statistical help that are rejected without review? State hypotheses, find the test statistic, use software or the bottom row of Table C to get a *P*-value, and give your conclusion. (This observational study does not establish causation, because studies that include statistical help may also be better in other ways than those that do not.)

20.23 Give a 95% confidence interval for the difference between the proportions of papers rejected without review when a statistician is and is not involved in the research.

20.27 The North Carolina State University study in the previous exercise also looked at possible differences in the proportions of female and male students who succeeded in the course. They found that 23 of the 34 women and 60 of the 89 men succeeded. Is there evidence of a difference between the proportions of women and men who succeed?

20.29 Nicotine patches are often used to help smokers quit. Does giving medicine to fight depression help? A randomized double-blind experiment assigned 244 smokers who wanted to stop to receive nicotine patches and another 245 to receive both a patch and the antidepression drug bupropion. After a year, 40 subjects in the nicotine patch group and 87 in the patch-plus-drug group had abstained from smoking. Give a 99% confidence interval for the difference (treatment minus control) in the proportion of smokers who quit.

20.31 Do our emotions influence economic decisions? One way to examine the issue is to have subjects play an "ultimatum game" against other people and against a computer. Your partner (person or computer) gets $10, on the condition that it be shared with you. The partner makes you an offer. If you refuse, neither of you gets anything. So it's to your advantage to accept even the unfair offer of $2 out of the $10. Some people get mad and refuse unfair offers. Here are data on the responses of 76 subjects randomly assigned to receive an offer of $2 from either a person they were introduced to or a computer:

	Accept	**Reject**
Human offers	20	18
Computer offers	32	6

We suspect that emotion will lead to offers from another person being rejected more often than offers from an impersonal computer. Do a test to assess the evidence for this conjecture.

20.35 Are shoppers more or less likely to use credit cards for "impulse purchases" that they decide to make on the spot, as opposed to purchases that they had in mind when they went to the store? Stop every third person leaving a department store with a purchase. (This is in effect a random sample of people who buy at that store.) A few questions allow us to classify the purchase as impulse or not. Here are the data on how the customer paid:

	Credit card?	
	Yes	**No**
Impulse purchases	13	18
Planned purchases	35	31

Estimate with 95% confidence the percent of all customers at this store who use a credit card. Give numerical summaries to describe the difference in credit card use between impulse and planned purchases. Is this difference statistically significant?

Chapter 21 Problem Statements

21.1 A sample survey of 1497 adult Internet users found that 36% consult the online collaborative encyclopedia Wikipedia. Give a 95% confidence interval for the proportion of all adult Internet users who refer to Wikipedia.

21.3 When the new euro coins were introduced throughout Europe in 2002, curious people tried all sorts of things. Two Polish mathematicians spun a Belgian euro (one side of the coin has a different design for each country) 250 times. They got 140 heads. Newspapers reported this result widely. Is it significant evidence that the coin is not balanced when spun?

21.5 Ask young men to estimate their own degree of body muscle by choosing from a set of 100 photos. Then ask them to choose what they think women prefer. The researchers know the actual degree of muscle, measured as kilograms per square meter of fat-free mass, for each of the photos. They can therefore measure the difference between what a subject thinks women prefer and the subject's own self-image. Call this difference the "muscle gap." Here are summary statistics for the muscle gap from two samples, one of American and European young men and the other of Chinese young men from Taiwan:

Group	n	$\bar{x}$	s
American/European	200	2.35	2.5
Chinese	55	1.20	3.2

Give a 95% confidence interval for the mean size of the muscle gap for all American and European young men. On the average, men think they need this much more muscle to match what women prefer.

21.7 Here's how butterflies mate: a male passes to a female a packet of sperm called a spermatophore. Females may mate several times. Will they remate sooner if the first spermatophore they receive is small? Among 20 females who received a large spermatophore (greater than 25 milligrams), the mean time to the next mating was 5.15 days, with standard deviation 0.18 day. For 21 females who received a small spermatophore (about 7 milligrams), the mean was 4.33 days and the standard deviation was 0.31 day. Is the observed difference in means statistically significant?

21.9 Give a 90% confidence interval for the difference between the proportions of all Hispanic and all white young people who listen to rap every day.

21.11 A study of the inheritance of speed and endurance in mice found a trade-off between these two characteristics, both of which help mice survive. To test endurance, mice were made to swim in a bucket with a weight attached to their tails. (The mice were rescued when exhausted.) Here are data on endurance in minutes for female and male mice:

Group	*n*	Mean	Standard Deviation
Female	162	11.4	26.09
Male	135	6.7	6.69

(a) Both sets of endurance data are skewed to the right. Why are *t* procedures nonetheless reasonably accurate for these data?
(b) Do the data show that female mice have significantly higher endurance on the average than male mice?

21.13 Use the information in Exercise 21.11 to give a 95% confidence interval for the mean difference (female minus male) in endurance times.

21.15 Use the information in the previous exercise to give a 99% confidence interval for the proportion of all students in 2004 who had at least one parent who graduated from college. (The sample excludes 17-year-olds who had dropped out of school, so your estimate is valid for students but is probably too high for all 17-year-olds.)

21.19 Starting in the 1970s, medical technology allowed babies with very low birth weight (VLBW, less than 1500 grams, about 3.3 pounds) to survive without major handicaps. It was noticed that these children nonetheless had difficulties in school and as adults. A long-term study has followed 242 VLBW babies to age 20 years, along with a control group of 233 babies from the same population who had normal birth weight.

(a) Is this an experiment or an observational study? Why?
(b) At age 20, 179 of the VLBW group and 193 of the control group had graduated from high school. Is the graduation rate among the VLBW group significantly lower than for the normal-birth-weight controls?

21.21 Of the 126 women in the VLBW group, 37 said they had used illegal drugs; 52 of the 124 control group women had done so. The IQ scores for the VLBW women had mean 86.2 (standard deviation 13.4), and the normal-birth-weight controls had mean IQ 89.8 (standard deviation 14.0). Is there a statistically significant difference between the two groups in either proportion using drugs or mean IQ?

21.23 The Women's Health Initiative is a randomized, controlled clinical trial designed to see if a low-fat diet reduces the incidence of breast cancer. In all, 19,541 women were assigned at random to a low-fat diet and a control group of 29,294 women were assigned to a normal diet. All the subjects were between ages 50 and 79 and had no prior breast cancer. After 8 years, 655 of the women in the low-fat group and 1072 of the women in the control group had developed breast cancer. Does this clinical trial give evidence that a low-fat diet reduces breast cancer?

21.25 High levels of cholesterol in the blood are not healthy in either humans or dogs. Because a diet rich in saturated fats raises the cholesterol level, it is plausible that dogs owned as pets have higher cholesterol levels than dogs owned by a veterinary research clinic. "Normal" levels of cholesterol based on the clinic's dogs would then be misleading. A clinic compared healthy dogs it owned with healthy pets brought to the clinic to be neutered. The summary statistics for blood cholesterol levels (milligrams per deciliter of blood) appear below.

Group	***n***	$\bar{x}$	***s***
Pets	26	193	68
Clinic	23	174	44

Is there strong evidence that pets have a higher mean cholesterol level than clinic dogs?

21.27 Continue your work with the information in Exercise 21.25. Give a 95% confidence interval for the mean cholesterol level in pets.

21.39 Dogs are big and expensive. Rats are small and cheap. Might rats be trained to replace dogs in sniffing out illegal drugs? A first study of this idea trained rats to rear up on their hind legs when they smelled simulated cocaine. To see how well rats performed after training, they were let loose on a surface with many cups sunk in it, one of which contained simulated cocaine. Four out of six trained rats succeeded in 80 out of 80 trials. How should we estimate the long-term success rate p of a rat that succeeds in every one of 80 trials?

(a) What is the rat's sample proportion $\hat{p}$? What is the large-sample 95% confidence interval for p? It's not plausible that the rat will *always* be successful, as this interval says.
(b) Find the plus four estimate $\tilde{p}$ and the plus four 95 % confidence interval for p. These results are more reasonable.

21.41 At what age do infants speak their first word of English? Here are data on 20 children (ages in months):

15	26	10	9	15	20	18	11	8	20
7	9	10	11	11	10	12	17	11	10

(In fact, the sample contained one more child, who began to speak at 42 months. Child development experts consider this abnormally late, so we dropped the outlier to get a sample of ``normal" children. The investigators are willing to treat these data as an SRS.) Is there good evidence that the mean age at first word among all normal children is greater than one year?

21.45 The color of a fabric depends on the dye used and also on how the dye is applied. This matters to clothing manufacturers, who want the color of the fabric to be just right. The study discussed in the previous exercise went on to dye fabric made of ramie with the same "procion blue" dye applied in two different ways. Here are the lightness scores for 8 pieces of identical fabric dyed in each way:

Method B	40.98	40.88	41.30	41.28	41.66	41.50	41.39	41.27
Method C	42.30	42.20	42.65	42.43	42.50	42.28	43.13	42.45

(a) This is a randomized comparative experiment. Outline the design.
(b) A clothing manufacturer wants to know which method gives the darker color (lower lightness score). Use sample means to answer this question. Is the difference between the two sample means statistically significant? Can you tell from just the *P*-value whether the difference is large enough to be important in practice?

21.47 We wonder what proportion of female students have at least one parent who allows them to drink around him or her. Table 21.1 contains information about a sample of 94 students. Use this sample to give a 95% confidence interval for this proportion.

21.49 We don't like to find broken crackers when we open the package. How can makers reduce breaking? One idea is to microwave the crackers for 30 seconds right after baking them. Analyze the following results from two experiments intended to examine this idea. Does microwaving significantly improve indicators of future breaking? How large is the improvement? What do you conclude about the idea of microwaving crackers?

(a) The experimenter randomly assigned 65 newly baked crackers to be microwaved and another 65 to a control group that is not microwaved. Fourteen days after baking, 3 of the 65 microwaved crackers and 57 of the 65 crackers in the control group showed visible checking, which is the starting point for breaks.

(b) The experimenter randomly assigned 20 crackers to be microwaved and another 20 to a control group. After 14 days, he broke the crackers. Here are summaries of the pressure needed to break them, in pounds per square inch:

	Microwave	**Control**
Mean	139.6	77.0
Standard deviation	33.6	22.6

Chapter 22 Problem Statements

22.1 The Pennsylvania State University has its main campus in University Park and more than 20 smaller "commonwealth campuses" around the state. The Penn State Division of Student Affairs polled a random sample of undergraduates about their use of online social networking. (The response rate was only about 20%, which casts some doubt on the usefulness of the data.) Facebook was the most popular site, with more than 80% of students having an account. Here is a comparison of Facebook use by undergraduates at the University Park and commonwealth campuses:

	University Park	**Commonwealth**
Do not use Facebook	68	248
Several times a month or less	55	76
At least once a week	215	157
At least once a day	640	394

(a) What percent of University Park students fall in each Facebook category? What percent of commonwealth campus students fall in each category? Each column should add to 100% (up to roundoff error). These are the conditional distributions of Facebook use given campus setting.
(b) Make a bar graph that compares the two conditional distributions. What are the most important differences in Facebook use between the two campus settings?

22.3 In the setting of Exercise 22.1, we might do several significance tests to compare University Park with the commonwealth campuses.

(a) Is there a significant difference between the proportions of students in the two locations who do not use Facebook? Give the *P*-value.
(b) Is there a significant difference between the proportions of students in the two locations who are in the "At least once a week" category? Give the *P*-value.
(c) Explain clearly why *P*-values for individual outcomes like these can't tell us whether the two distributions for all four outcomes in the two locations differ significantly.

22.5 The two-way table in Exercise 22.1 displays data on use of Facebook by two groups of Penn State students. It's clear that nonusers are much more frequent at the commonwealth campuses. Let's look just at students who have Facebook accounts:

Use Facebook	**University Park**	**Commonwealth**
Several times a month or less	55	76
At least once a week	215	157
At least once a day	640	394
Total Facebook Users	910	627

The null hypothesis is that there is no relationship between campus and Facebook use.

(a) If this hypothesis is true, what are the expected counts for Facebook use among commonwealth campus students? This is one column of the two-way table of expected counts. Find the column total and verify that it agrees with the column total for the observed counts.
(b) Commonwealth campus students as a group are older and more likely to be married and employed than University Park students. What does comparing the observed and expected counts in this column show about Facebook use by these students?

22.13 Many birds are injured or killed by flying into windows. It appears that birds don't see windows. Can tilting windows down so that they reflect earth rather than sky reduce bird strikes? Place six windows at the edge of a woods: two vertical, two tilted 20 degrees, and two tilted 40 degrees. During the next four months, there were 53 bird strikes, 31 on the vertical windows, 14 on the 20-degree windows, and 8 on the 40-degree windows. If the tilt has no effect, we expect strikes on windows with all three tilts to have equal probability. Test this null hypothesis. What do you conclude?

22.15 Police may use minor violations such as not wearing a seat belt to stop motorists for other reasons. A large study in Michigan first studied the population of drivers not wearing seat belts during daylight hours by observation at more than 400 locations around the state. Here is the population distribution of seat belt violators by age group:

Age group	16 to 29	30 to 59	60 or older
Proportion	0.328	0.594	0.078

The researchers then looked at court records and called a random sample of 803 drivers who had actually been cited by police for not wearing a seat belt. Here are the counts:

Age group	16 to 29	30 to 59	60 or older
Count	401	382	20

Does the age distribution of people cited differ significantly from the distribution of ages of all seat belt violators? Which age groups have the largest contributions to chi-square? Are these age groups cited more or less frequently than is justified? (The study found that males, blacks, and younger drivers were all over-cited.)

22.17 For reasons known only to social scientists, the General Social Survey (GSS) regularly asks its subjects their astrological sign. Here are the counts of responses for the most recent GSS:

Sign	Aries	Taurus	Gemini	Cancer	Leo	Virgo
Count	321	360	367	374	383	402

Sign	Libra	Scorpio	Sagittarius	Capricorn	Aquarius	Pisces
Count	392	329	331	354	376	355

If births are spread uniformly across the year, we expect all 12 signs to be equally likely. Are they? Follow the four-step process in your answer.

22.29 The General Social Survey (GSS) asked this question: "Consider a person who believes that Blacks are genetically inferior. If such a person wanted to make a speech in your community claiming that Blacks are inferior, should he be allowed to speak, or not?" Here are the responses, broken down by the race of the respondent:

	Black	**White**	**Other**
Allowed	140	976	121
Not allowed	129	480	131

(a) Because the GSS is essentially an SRS of all adults, we can combine the races in these data and give a 99% confidence interval for the proportion of all adults who would allow a racist to speak. Do this.
(b) Find the column percents and use them to compare the attitudes of the three racial groups. How significant are the differences found in the sample?

22.43 The nonprofit group Public Agenda conducted telephone interviews with a stratified sample of parents of high school children. There were 202 black parents, 202 Hispanic parents, and 201 white parents. One question asked was "Are the high schools in your state doing an excellent, good, fair or poor job, or don't you know enough to say?" Here are the survey results:

	Black Parents	Hispanic Parents	White Parents
Excellent	12	34	22
Good	69	55	81
Fair	75	61	60
Poor	24	24	24
Don't know	22	28	14
Total	202	202	201

Are the differences in the distributions of responses for the three groups of parents statistically significant? What departures from the null hypothesis "no relationship between group and response" contribute most to the value of the chi-square statistic? Write a brief conclusion based on your analysis.

22.45 Before bringing a new product to market, firms carry out extensive studies to learn how consumers react to the product and how best to advertise its advantages. Here are data from a study of a new laundry detergent. The subjects are people who don't currently use the established brand that the new product will compete with. Give subjects free samples of both detergents. After they have tried both for a while, ask which they prefer. The answers may depend on other facts about how people do laundry.

	Laundry Practices			
	Soft water, warm wash	Soft water, Hot wash	Hard water, Warm wash	Hard water, Hot wash
Prefer standard product	53	27	42	30
Prefer new product	63	29	68	42

How do laundry practices (water hardness and wash temperature) influence the choice of detergent? In which settings does the new detergent do best? Are the differences between the detergents statistically significant?

22.47 Make a 2 x 5 table by combining the counts in the three rows that mention Democrat and in the three rows that mention Republican and ignoring strict independents and supporters of other parties. We might think of this table as comparing all adults who lean Democrat and all adults who lean Republican. How does support for the two major parties differ among adults with different levels of education?

Chapter 23 Problem Statements

23.1 An outbreak of the deadly Ebola virus in 2002 and 2003 killed 91 of the 95 gorillas in 7 home ranges in the Congo. To study the spread of the virus, measure "distance" by the number of home ranges separating a group of gorillas from the first group infected. Here are data on distance and number of days until deaths began in each later group:

Distance *x*	1	3	4	4	4	5
Days *y*	4	21	33	41	43	46

(a) Examine the data. Make a scatterplot with distance as the explanatory variable and find the correlation. There is a strong linear relationship.
(b) Explain in words what the slope β of the population regression line would tell us if we knew it. Based on the data, what are the estimates of β and the intercept α of the population regression line?
(c) Calculate by hand the residuals for the six data points. Check that their sum is 0 (up to roundoff error). Use the residuals to estimate the standard deviation σ that measures variation in the responses (days) about the means given by the population regression line. You have now estimated all three parameters.

23.3 One effect of global warming is to increase the flow of water into the Arctic Ocean from rivers. Such an increase may have major effects on the world's climate. Six rivers (Yenisey, Lena, Ob, Pechora, Kolyma, and Severnaya Dvina) drain two-thirds of the Arctic in Europe and Asia. Several of these are among the largest rivers on earth. Table 23.2 presents the total discharge from these rivers each year from 1936 to 1999. Discharge is measured in cubic kilometers of water. Use software to analyze these data.

(a) Make a scatterplot of river discharge against time. Is there a clear increasing trend? Calculate r^2 and briefly interpret its value. There is considerable year-to-year variation, so we wonder if the trend is statistically significant.
(b) As a first step, find the least-squares line and draw it on your plot. Then find the regression standard error *s*, which measures scatter about this line. We will continue the analysis in later exercises.

23.5 The most important question we ask of the data in Table 23.2 is this: is the increasing trend visible in your plot (Exercise 23.3) statistically significant? If so, changes in the Arctic may already be affecting the earth's climate. Use software to answer this question. Give a test statistic, its *P*-value, and the conclusion you draw from the test.

23.7 Exercise 23.1 gives data showing that the delay in deaths from an Ebola outbreak in groups of gorillas increases linearly with distance from the origin of the outbreak. There are only 6 observations, so we worry that the apparent relationship may be just chance. Is the correlation significantly greater than 0? Answer this question in two ways.

(a) Return to your *t* statistic from Exercise 23.4. What is the one-sided *P*-value for this *t*? Apply your result to test the correlation.
(b) Find the correlation *r* and use Table E to approximate the *P*-value of the one-sided test.

23.9 Exercise 23.1 presents data on distance and days until an Ebola outbreak reached six groups of gorillas. Software tells us that the least-squares slope is $b = 11.263$ with standard error $SE_b = 1.591$. Because there are only 6 observations, the observed slope *b* may not be an accurate estimate of the population slope β. Give a 90% confidence interval for β.

23.11 Use the data in Table 23.2 to give a 90% confidence interval for the slope of the population regression of Arctic river discharge on year. Does this interval convince you that discharge is actually increasing over time? Explain your answer.

23.29 We know that there is a strong linear relationship. Let's check the other conditions for inference. Figure 23.14 includes a table of the two variables, the predicted values $\hat{y}$ for each *x* in the data, the residuals, and related quantities. (This table is stored as *ex23-29.dat* on the text CD and Web site.)

(a) Round the residuals to the nearest whole number and make a stemplot. The distribution is single-peaked and symmetric and appears close to Normal.
(b) Make a residual plot, residuals against boats registered. Use a vertical scale from –25 to 25 to show the pattern more clearly. Add the "residual = 0" line. There is no clearly nonlinear pattern. The spread about the line may be a bit greater for larger values of the explanatory variable, but the effect is not large.
(c) It is reasonable to regard the number of manatees killed by boats in successive years as independent. The number of boats grew over time. Someone says that pollution also grew over time and may explain more manatee deaths. How would you respond to this idea?

23.33 Exercise 5.53 (page 158) gives data on William Gray's predictions of the number of named tropical storms in Atlantic hurricane seasons from 1984 to 2007. Use these data for regression inference as follows.

(a) Does Professor Gray do better than random guessing? That is, is there a significantly positive correlation between his forecasts and the actual number of storms? (Report a t statistic from regression output and give the one-sided P-value.)
(b) Give a 95% confidence interval for the mean number of storms in years when Professor Gray forecasts 16 storms.

23.41 Exercise 7.25 (page 188) gives data on the abundance of the pine cones that red squirrels feed on and the mean number of offspring per female squirrel over 16 years. The strength of the relationship is remarkable because females produce young before the food is available. How significant is the evidence that more cones leads to more offspring? (Use a vertical scale from –2 to 2 in your residual plot to show the pattern more clearly.)

23.43 Exercise 5.51 (page 157) describes a study that found that the number of stumps from trees felled by beavers predicts the abundance of beetle larvae. Is there good evidence that more beetle larvae clusters are present when beavers have left more tree stumps? Estimate how many more clusters accompany each additional stump, with 95% confidence.

Chapter 24 Problem Statements

24.9 Bromeliads are tropical flowering plants. Many are epiphytes that attach to trees and obtain moisture and nutrients from air and rain. Their leaf bases form cups that collect water and are home to the larvae of many insects. As a preliminary to a study of changes in the nutrient cycle, Jacqueline Ngai and Diane Srivastava examined the effects of adding nitrogen, phosphorus, or both to the cups. They randomly assigned 8 bromeliads growing in Costa Rica to each of four treatment groups, including an unfertilized control group. A monkey destroyed one of the plants in the control group, leaving 7 bromeliads in that group. Here are the numbers of new leaves on each plant over the 7 months following fertilization:

Nitrogen	Phosphorus	Both	Neither
15	14	14	11
14	14	16	13
15	14	15	16
16	11	14	15
17	13	14	15
18	12	13	11
17	15	17	12
13	15	14	

Analyze these data and discuss the results. Does nitrogen or phosphorus have a greater effect on the growth of bromeliads? Follow the four-step process as illustrated in Example 24.4.

24.13 What conditions help overweight people exercise regularly? Subjects were randomly assigned to three treatments: a single long exercise period 5 days per week; several 10-minute exercise periods 5 days per week; and several 10-minute periods 5 days per week on a home treadmill that was provided to the subjects. The study report contains the following information about weight loss (in kilograms) after six months of treatment:

Treatment	n	$\bar{x}$	s
Long exercise periods	37	10.2	4.2
Short exercise periods	36	9.3	4.5
Short periods with equipment	42	10.2	5.2

(a) Do the standard deviations satisfy the rule of thumb for safe use of ANOVA?
(b) Calculate the overall mean response $\bar{x}$. the mean squares MSG and MSE, and the F statistic.
(c) Which F distribution would you use to find the P-value of the ANOVA F test? Software says that $P = 0.634$. What do you conclude from this study?

24.33 Our bodies have a natural electrical field that helps wounds heal. Might higher or lower levels speed healing? An experiment with newts investigated this question. Newts were randomly assigned to five groups. In four of the groups, an electrode applied to one hind limb (chosen at random) changed the natural field, while the other hind limb was not manipulated. Both limbs in the fifth (control) group remained in their natural state.

Table 24.5 gives data from this experiment. The "Group" variable shows the field applied as a multiple of the natural field for each newt. For example, "0.5" is half the natural field, "1" is the natural level (the control group), and "1.5" indicates a field 1.5 times natural. "Diff" is the response variable, the difference in the healing rate (in micrometers per hour) of cuts made in the experimental and control limbs of that newt. Negative values mean that the experimental limb healed more slowly. The investigators conjectured that nature heals best, so that changing the field from the natural state (the "1" group) will slow healing.

Do a complete analysis to see whether the groups differ in the effect of the electrical field level on healing. Follow the four-step process in your work.

Table 24.5 Effect of electrical field on healing rate in newts

Group	Diff	Group	Diff	Group	Diff	Group	Diff	Group	Diff
0	-10	0.5	-1	1	-7	1.25	1	1.5	-13
0	-12	0.5	10	1	15	1.25	8	1.5	-49
0	-9	0.5	3	1	-4	1.25	-15	1.5	-16
0	-11	0.5	-3	1	-16	1.25	14	1.5	-8
0	-1	0.5	-31	1	-2	1.25	-7	1.5	-2
0	6	0.5	4	1	-13	1.25	-1	1.5	-35
0	-31	0.5	-12	1	5	1.25	11	1.5	-11
0	-5	0.5	-3	1	-4	1.25	8	1.5	-46
0	13	0.5	-7	1	-2	1.25	11	1.5	-22
0	-2	0.5	-10	1	-14	1.25	-4	1.5	2
0	-7	0.5	-22	1	5	1.25	7	1.5	10
0	-8	0.5	-4	1	11	1.25	-14	1.5	-4
		0.5	-1	1	10	1.25	0	1.5	-10
		0.5	-3	1	3	1.25	5	1.5	2
				1	6	1.25	-2	1.5	-5
				1	-1				
				1	13				
				1	-8				

24.35 "Durable press" cotton fabrics are treated to improve their recovery from wrinkles after washing. Unfortunately, the treatment also reduces the strength of the fabric. A study compared the breaking strength of untreated fabric with that of fabrics treated by three commercial durable press processes. Five specimens of the same fabric were assigned at random to each group. Here are the data, in pounds of pull needed to tear the fabric:

Untreated	60.1	56.7	61.5	55.1	59.4
Permafresh 55	29.9	30.7	30.0	29.5	27.6
Permafresh 48	24.8	24.6	27.3	28.1	30.3
Hylite LF	28.8	23.9	27.0	22.1	24.2

The untreated fabric is clearly much stronger than any of the treated fabrics. We want to know if there is a significant difference in breaking strength among the three durable press treatments. Analyze the data for the three processes and write a clear summary of your findings. Which process do you recommend if breaking strength is a main concern? Use the four-step process to guide your discussion. (Although the standard deviations do not quite satisfy our rule of thumb, that rule is conservative and many statisticians would use ANOVA for these data.)

24.39 Your work in Exercise 24.30 shows that there were significant differences in mean plant biomass among the three treatments in 2003. Do a complete analysis of the data for 2001 and report your conclusions.

Chapter 25 Problem Statements

25.1 Our lead example for the two-sample *t* procedures in Chapter 18 concerned a study comparing the level of physical activity of lean and mildly obese people who don't exercise. Here are the minutes per day that the subjects spent standing or walking over a 10-day period:

Lean subjects		Obese subjects	
511.100	543.388	260.244	416.531
607.925	677.188	464.756	358.650
319.212	555.656	367.138	267.344
584.644	374.831	413.667	410.631
578.869	504.700	347.375	426.356

The data are a bit irregular but not distinctly non-Normal. Let's use the Wilcoxon test for comparison with the two-sample *t* test.

(a) Find the median minutes spent standing or walking for each group. Which group appears more active?
(b) Arrange all 20 observations in order and find the ranks.
(c) Take W to be the sum of the ranks for the lean group. What is the value of W? If the null hypothesis (no difference between the groups) is true, what are the mean and standard deviation of W?
(d) Does comparing W with the mean and standard deviation suggest that the lean subjects are more active than the obese subjects?

25.3 In Exercise 25.1, you found the Wilcoxon rank sum W and its mean and standard deviation. We want to test the null hypothesis that the two groups don't differ in activity against the alternative hypothesis that the lean subjects spend more time standing and walking.

(a) What is the probability expression for the *P*-value of W if we use the continuity correction?
(b) Find the *P*-value. What do you conclude?

25.7 Use your software to carry out the one-sided Wilcoxon rank sum test that you did by hand in Exercise 25.3. Use the exact distribution if your software will do it. Compare the software result with your result in Exercise 25.3.

25.11 Exercise 18.8 (text page 482) compares the breaking strength of polyester strips buried for 16 weeks with that of strips buried for 2 weeks. The breaking strengths in pounds are

2 weeks	118	126	126	120	129
16 weeks	124	98	110	140	110

(a) What are the null and alternative hypotheses for the Wilcoxon test? For the two-sample *t* test?
(b) There are two pairs of tied observations. What ranks do you assign to each observation, using average ranks for ties?
(c) Apply the Wilcoxon rank sum test to these data. Compare your result with the P = 0.1857 obtained from the two-sample *t* test in Figure 18.5.

25.13 The data in Exercise 25.5 for a story told without pictures (Story 1) have tied observations. Is there good evidence that high-progress readers score higher than low-progress readers when they retell a story they have heard without pictures?

(a) Make a back-to-back stemplot of the 5 responses in each group. Are any major deviations from Normality apparent?
(b) Carry out a two-sample *t* test. State hypotheses and give the two sample means, the *t* statistic and its *P*-value, and your conclusion.
(c) Carry out the Wilcoxon rank sum test. State hypotheses and give the rank sum *W* for high-progress readers, its *P* -value, and your conclusion. Do the *t* and Wilcoxon tests lead you to different conclusions?

25.15 Exercise 7.41 (text page 193) gives data from an experiment in which some bellflower plants in a forest were “fertilized” with dead cicadas and other plants were not disturbed. The data record the mass of seeds produced by 39 cicada plants and 33 undisturbed (control) plants. Do the data show that dead cicadas increase seed mass? Do data analysis to compare the two groups, explain why you would be reluctant to use the two-sample *t* test, and apply the Wilcoxon test. Follow the four-step process in your report.

25.19 Lymphocytes (white blood cells) play an important role in defending our bodies against tumors and infections. Can lymphocytes be genetically modified to recognize and destroy cancer cells? In one study of this idea, modified cells were infused into 11 patients with metastatic melanoma (serious skin cancer) that had not responded to existing treatments. Here are data for an “ELISA” test for the presence of cells that trigger an immune response, in counts per 100,000 cells before and after infusion. High counts suggest that infusion had a beneficial effect.

Patient	1	2	3	4	5	6	7	8	9	10	11
Pre	14	0	1	0	0	0	0	20	1	6	0
Post	41	7	1	215	20	700	13	530	35	92	108

(a) Examine the differences (post minus pre). Why can't we use the matched pairs *t* test to see if infusion raised the ELISA counts?
(b) We will apply the Wilcoxon signed rank test. What are the ranks for the absolute values of the differences in counts? What is the value of W^+?
(c) What would be the mean and standard deviation of W^+ if the null hypothesis (infusion makes no difference) were true? Compare W^+ with this mean (in standard deviation units) to reach a tentative conclusion about significance.

25.23 Exercise 17.7 (text page 449) reports the following data on the percent of nitrogen in bubbles of ancient air trapped in amber:

63.4 65.0 64.4 63.3 54.8 64.5 60.8 49.1 51.0

We wonder if ancient air differs significantly from the present atmosphere, which is 78.1% nitrogen.

(a) Graph the data, and comment on skewness and outliers. A rank test is appropriate.
(b) We would like to test hypotheses about the median percent of nitrogen in ancient air (the population):

$$H_0: \text{ median} = 78.1$$
$$H_a: \text{ median} \neq 78.1$$

To do this, apply the Wilcoxon signed rank statistic to the differences between the observations and 78.1. (This is the one-sample version of the test.) What do you conclude?

25.25 Cola makers test new recipes for loss of sweetness during storage. Trained tasters rate the sweetness before and after storage. Here are the sweetness losses (sweetness before storage minus sweetness after storage) found by 10 tasters for one new cola recipe:

2.0 0.4 0.7 2.0 –0.4 2.2 –1.5 1.2 1.1 2.3

Are these data good evidence that the cola lost sweetness?

(a) These data are the differences from a matched pairs design. State hypotheses in terms of the median difference in the population of all tasters, carry out a test, and give your conclusion.
(b) The output in Figure 17.6 (text page 454) showed that the one-sample *t* test had *P*-value $P = 0.0123$ for these data. How does this compare with your result from (a)? What are the hypotheses for the *t* test? What conditions must be met for each of the *t* and Wilcoxon tests?

25.27 Exercise 24.30 describes an experiment that examines the effect on plant biomass in plots of California grassland randomly assigned to receive added water in the winter, added water in the spring, or no added water. The experiment continued for several years. Here are data for 2004 (mass in grams per square meter):

Winter	Spring	Control
254.6453	517.6650	178.9988
233.8155	342.2825	205.5165
253.4506	270.5785	242.6795
228.5882	212.5324	231.7639
158.6675	213.9879	134.9847
212.3232	240.1927	212.4862

The sample sizes are small and the data contain some possible outliers. We will apply a nonparametric test.

(a) Examine the data. Show that the conditions for ANOVA (text page 644) are not met. What appear to be the effects of extra rain in winter or spring?
(b) What hypotheses does ANOVA test? What hypotheses does Kruskal-Wallis test?
(c) What are I, the n_i, and N? Arrange the counts in order and assign ranks.
(d) Calculate the Kruskal-Wallis statistic H. How many degrees of freedom should you use for the chi-square approximation to its null distribution? Use the chi-square table to give an approximate P-value. What does the test lead you to conclude?

25.29 Here are the breaking strengths (in pounds) of strips of polyester fabric buried in the ground for several lengths of time:

2 weeks	118	126	126	120	129
4 weeks	130	120	114	126	128
8 weeks	122	136	128	146	140
16 weeks	124	98	110	140	110

Breaking strength is a good measure of the extent to which the fabric has decayed. Do a complete analysis that compares the four groups. Give the Kruskal-Wallis test along with a statement in words of the null and alternative hypotheses.

25.45 Investigators compared the number of tree species in unlogged plots in the rain forest of Borneo with the number of species in plots logged 8 years earlier. Here are the data:

Unlogged	22	18	22	20	15	21	13	13	19	13	19	15
Logged	17	4	18	14	18	15	15	10	12			

Does logging significantly reduce the number of species in a plot after 8 years?

25.51 “Second, we found that species richness within tributaries exceeded that within their adjacent upstream mainstem stations.” Again, do a test to confirm significance and report your finding.

Chapter 26 Problem Statements

26.13 Exercise 26.10 concerns process control data on the hardness of tablets (measured in kilograms) for a pharmaceutical product. Table 26.4 gives data for 20 new samples of size 4, with the $\bar{x}$ and s for each sample. The process has been in control with mean at the target value $\mu = 11.5$ kg and standard deviation $\sigma = 0.2$ kg.

Table 26.4 Twenty samples of size 4, with $\bar{x}$ and s

Sample	Hardness (kilograms)				Mean	StDev
1	11.432	11.35	11.582	11.184	11.387	0.1660
2	11.791	11.323	11.734	11.512	11.590	0.2149
3	11.373	11.807	11.651	11.651	11.620	0.1806
4	11.787	11.585	11.386	11.245	11.501	0.2364
5	11.633	11.212	11.568	11.469	11.470	0.1851
6	11.648	11.653	11.618	11.314	11.558	0.1636
7	11.456	11.270	11.817	11.402	11.486	0.2339
8	11.394	11.754	11.867	11.003	11.504	0.3905
9	11.349	11.764	11.402	12.085	11.650	0.3437
10	11.478	11.761	11.907	12.091	11.809	0.2588
11	11.657	12.524	11.468	10.946	11.649	0.6564
12	11.820	11.872	11.829	11.344	11.716	0.2492
13	12.187	11.647	11.751	12.026	11.903	0.2479
14	11.478	11.222	11.609	11.271	11.395	0.1807
15	11.750	11.520	11.389	11.803	11.616	0.1947
16	12.137	12.056	11.255	11.497	11.736	0.4288
17	12.055	11.730	11.856	11.357	11.750	0.2939
18	12.107	11.624	11.727	12.207	11.916	0.2841
19	11.933	10.658	11.708	11.278	11.394	0.5610
20	12.512	12.315	11.671	11.296	11.948	0.5641

(a) Make both $\bar{x}$ and s charts for these data based on the information given about the process.

(b) At some point, the within-sample process variation increased from $\sigma = 0.2$ kg to $\sigma = 0.4$ kg. About where in the 20 samples did this happen? What is the effect on the s chart? On the $\bar{x}$ chart?

(c) At that same point, the process mean changed from $\mu = 11.5$ kg to $\mu = 11.7$ kg. What is the effect of this change on the s chart? On the $\bar{x}$ chart?

26.15 Figure 26.10 reproduces a data sheet from the floor of a factory that makes electrical meters. The sheet shows measurements on the distance between two mounting holes for 18 samples of size 5. The heading informs us that the measurements are in multiples of 0.0001 inch above 0.6000 inch. That is, the first measurement, 44, stands for

0.6044 inch. All the measurements end in 4. Although we don't know why this is true, it is clear that in effect the measurements were made to the nearest 0.001 inch, not to the nearest 0.0001 inch.

Calculate $\bar{x}$ and *s* for the first two samples. The data file *ex26-15.dat* contains $\bar{x}$ and *s* for all 18 samples. Based on long experience with this process, you are keeping control charts based on $\mu = 43$ and $\sigma = 12.74$. Make *s* and $\bar{x}$ charts for the data in Figure 26.10 and describe the state of the process.

26.21 Table 26.6 gives data on the losses (in dollars) incurred by a hospital in treating major joint replacement (DRG 209) patients. The hospital has taken from its records a random sample of 8 such patients each month for 15 months.

Table 26.6 Hospital losses for 15 samples of DRG 209 patients

Sample	Loss (Dollars)								$\bar{x}$	s
1	6835	5843	6019	6731	6362	5696	7193	6206	6360.6	521.7
2	6452	6764	7083	7352	5239	6911	7479	5549	6603.6	817.1
3	7205	6374	6198	6170	6482	4763	7125	6241	6319.8	749.1
4	6021	6347	7210	6384	6807	5711	7952	6023	6556.9	736.5
5	7000	6495	6893	6127	7417	7044	6159	6091	6653.2	503.7
6	7783	6224	5051	7288	6584	7521	6146	5129	6465.8	1034.3
7	8794	6279	6877	5807	6076	6392	7429	5220	6609.2	1104
8	4727	8117	6586	6225	6150	7386	5674	6740	6450.6	1033
9	5408	7452	6686	6428	6425	7380	5789	6264	6479.0	704.7
10	5598	7489	6186	5837	6769	5471	5658	6393	6175.1	690.5
11	6559	5855	4928	5897	7532	5663	4746	7879	6132.4	1128.6
12	6824	7320	5331	6204	6027	5987	6033	6177	6237.9	596.6
13	6503	8213	5417	6360	6711	6907	6625	7888	6828.0	879.8
14	5622	6321	6325	6634	5075	6209	4832	6386	5925.5	667.8
15	6269	6756	7653	6065	5835	7337	6615	8181	6838.9	819.5

(a) Make an *s* control chart using center lines and limits calculated from these past data. There are no points out of control.
(b) Because the *s* chart is in control, base the $\bar{x}$ chart on all 15 samples. Make this chart. Is it also in control?

26.27 If the mesh tension of individual monitors follows a Normal distribution, we can describe capability by giving the percent of monitors that meet specifications. The old specifications for mesh tension are 100 to 400 mV. The new specifications are 150 to 350 mV. Because the process is in control, we can estimate that tension has mean 275 mV and standard deviation 38.4 mV.

(a) What percent of monitors meet the old specifications?
(b) What percent meet the new specifications?

26.29 Figure 26.10 (page 26-20) displays a record sheet for 18 samples of distances between mounting holes in an electrical meter. The data file *ex26-15.dat* adds $\bar{x}$ and *s* for each sample. In Exercise 26.15, you found that Sample 5 was out of control on the process-monitoring *s* chart. The special cause responsible was found and removed. Based on the 17 samples that were in control, what are the natural tolerances for the distance between the holes?

26.35 Here are data from an urban school district on the number of eighth-grade students with three or more unexcused absences from school during each month of a school year. Because the total number of eighth-graders changes a bit from month to month, these totals are also given for each month.

Month	**Sep.**	**Oct.**	**Nov.**	**Dec.**	**Jan.**	**Feb.**	**Mar.**	**Apr.**	**May**	**Jun.**
Students	911	947	939	942	918	920	931	925	902	883
Absent	291	349	364	335	301	322	344	324	303	344

(a) Find $\bar{p}$. Because the number of students varies from month to month, also find $\bar{n}$, the average per month.
(b) Make a *p* chart using control limits based on $\bar{n}$ students each month. Comment on control.
(c) The exact control limits are different each month because the number of students *n* is different each month. This situation is common in using *p* charts. What are the exact limits for October and June, the months with the largest and smallest *n*? Add these limits to your *p* chart, using short lines spanning a single month. Do exact limits affect your conclusions?

26.43 Painting new auto bodies is a multistep process. There is an "electrocoat" that resists corrosion, a primer, a color coat, and a gloss coat. A quality study for one paint shop produced this breakdown of the primary problem type for those autos whose paint did not meet the manufacturer's standards:

Problem	Percent
Electrocoat uneven---redone	4
Poor adherence of color to primer	5
Lack of clarity in color	2
"Orange peel" texture in color	32
"Orange peel" texture in gloss	1
Ripples in color coat	28
Ripples in gloss coat	4
Uneven color thickness	19
Uneven gloss thickness	5
Total	100

Make a Pareto chart. Which stage of the painting process should we look at first?

26.51 Calculate control limits for s, make an s chart, and comment on control of short-term process variation.

Chapter 27 Problem Statements

27.15 The table below shows the progress of world record times (in seconds) for the 10,000-meter run for both men and women.

Men				Women	
Record year	Time (seconds)	Record year	Time (seconds)	Record year	Time (seconds)
1912	1880.8	1962	1698.2	1967	2286.4
1921	1840.2	1963	1695.6	1970	2130.5
1924	1835.4	1965	1659.3	1975	2100.4
1924	1823.2	1972	1658.4	1975	2041.4
1924	1806.2	1973	1650.8	1977	1995.1
1937	1805.6	1977	1650.5	1979	1972.5
1938	1802.0	1978	1642.4	1981	1950.8
1939	1792.6	1984	1633.8	1981	1937.2
1944	1775.4	1989	1628.2	1982	1895.3
1949	1768.2	1993	1627.9	1983	1895.0
1949	1767.2	1993	1618.4	1983	1887.6
1949	1761.2	1994	1612.2	1984	1873.8
1950	1742.6	1995	1603.5	1985	1859.4
1953	1741.6	1996	1598.1	1986	1813.7
1954	1734.2	1997	1591.3	1993	1771.8
1956	1722.8	1997	1587.8		
1956	1710.4	1998	1582.7		
1960	1698.8	2005	1577.5		

(a) Make a scatterplot of world record time against year, using separate symbols for men and women. Describe the pattern for each gender. Then compare the progress of men and women.
(b) Fit the model with two regression lines, one for women and one for men, and identify the estimated regression lines.
(c) Women began running this long distance later than men, so we might expect their improvement to be more rapid. Moreover, it is often said that men have little advantage over women in distance running as opposed to sprints, where muscular strength plays a greater role. Do the data appear to support these claims?

27.19 An experiment was conducted using a Geiger-Mueller tube in a physics lab. Geiger-Mueller tubes respond to gamma rays and to beta particles (electrons). A pulse that corresponds to each detection of a decay product is produced, and these pulses were counted using a computer-based nuclear counting board. Elapsed time (in seconds) and counts of pulses for a short-lived unstable isotope of silver are shown in Table 27.5 (see page 27-36).

(a) Create a scatterplot of the counts versus time and describe the pattern.
(b) Since some curvature is apparent in the scatterplot, you might want to consider the quadratic model for predicting counts based on time. Fit the quadratic model and identify the estimated mean response.
(c) Add the estimated mean response to your scatterplot. Would you recommend the use of the quadratic model for predicting radioactive decay in this situation? Explain.
(d) Transform the counts using the natural logarithm and create a scatterplot of the transformed variable versus time.
(e) Fit a simple linear regression model using the natural logarithm of the counts. Provide the estimated regression line, a scatterplot with the estimated regression line, and appropriate residual plots.
(f) Does the simple linear regression model for the transformed counts fit the data better than the quadratic regression model? Explain.

27.23 Suppose that the couple shopping for a diamond in Example 27.15 had used a quadratic regression model for the other quantitative variable, *Depth*. Use the data in the file *ta27-04.dat* to answer the following questions.

(a) What is the estimated quadratic regression model for mean total price based on the explanatory variable *Depth*?
(b) As you discovered in part (a), it is always possible to fit quadratic models, but we must decide if they are helpful. Is this model as informative to the couple as the model in Example 27.15? What percent of variation in the total price is explained by using the quadratic regression model with *Depth*?

27.25 Table 27.8 contains data on the size of perch caught in a lake in Finland. Use statistical software to help you analyze these data.

(a) Use the multiple regression model with two explanatory variables, length and width, to predict the weight of a perch. Provide the estimated multiple regression equation.
(b) How much of the variation in the weight of perch is explained by the model in part (a)?
(c) Does the ANOVA table indicate that at least one of the explanatory variables is helpful in predicting the weight of perch? Explain.
(d) Do the individual *t* tests indicate that both β_1 and β_2 are significantly different from zero? Explain.
(e) Create a new variable, called interaction, that is the product of length and width. Use the multiple regression model with three explanatory variables, length, width, and interaction, to predict the weight of a perch. Provide the estimated multiple regression equation.

(f) How much of the variation in the weight of perch is explained by the model in part (e)?
(g) Does the ANOVA table indicate that at least one of the explanatory variables is helpful in predicting the weight of perch? Explain.
(h) Describe how the individual *t* statistics changed when the interaction term was added.

27.27 Use explanatory variables length, width, and interaction from Exercise 27.25 (page 27-49) on the 56 perch to provide 95% confidence intervals for the mean and prediction intervals for future observations. Interpret both intervals for the 10th perch in the data set. What *t* distribution is used to provide both intervals?

27.41 A multimedia statistics learning system includes a test of skill in using the computer's mouse. The software displays a circle at a random location on the computer screen. The subject clicks in the circle with the mouse as quickly as possible. A new circle appears as soon as the subject clicks the old one. Table 5.3 (text page 159) gives data for one subject's trials, 20 with each hand. Distance is the distance from the cursor location to the center of the new circle, in units whose actual size depends on the size of the screen. Time is the time required to click in the new circle, in milliseconds.

(a) Specify the population multiple regression model for predicting time from distance separately for each hand. Make sure you include the interaction term that is necessary to allow for the possibility of having different slopes. Explain in words what each β in your model means.
(b) Use statistical software to find the estimated multiple regression equation for predicting time from distance separately for each hand. What percent of variation in the distances is explained by this multiple regression model?
(c) Explain how to use the estimated multiple regression equation in part (b) to obtain the least-squares line for each hand. Draw these lines on a scatterplot of time versus distance.

27.45 The Sanchez household is about to install solar panels to reduce the cost of heating their house. In order to know how much the solar panels help, they record their consumption of natural gas before the solar panels are installed. Gas consumption is higher in cold weather, so the relationship between outside temperature and gas consumption is important. Here are the data for 16 consecutive months:

	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June
Degree-days	24	51	43	33	26	13	4	0
Gas used	6.3	10.9	8.9	7.5	5.3	4.0	1.7	1.2

	July	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.
Degree-days	0	1	6	12	30	32	52	30
Gas used	1.2	1.2	2.1	3.1	6.4	7.2	11.0	6.9

Outside temperature is recorded in degree-days, a common measure of demand for heating. A day's degree-days are the number of degrees its average temperature falls below 65ºF. Gas used is recorded in hundreds of cubic feet.

(a) Create an indicator variable, say *INDwinter*, which is 1 for the months of November, December, January, and February. Make a plot of all the data using a different symbol for winter months.
(b) Fit the model with two regression lines, one for winter months and one for other months, and identify the estimated regression lines.
(c) Do you think that two regression lines were needed to explain the relationship between gas used and degree-days? Explain.

Chapter 28 Problem Statements

28.1 The full data for the logging study appear in Table 24.2 (text page 640). The data for counts of individual trees in the plots studied also appear in the data file *ex28-01.dat*. Carry out data analysis and ANOVA to determine whether logging affects the mean count of individual trees in a plot.

28.3 If you are a dog lover, perhaps having your dog along reduces the effect of stress. To examine the effect of pets in stressful situations, researchers recruited 45 women who said they were dog lovers. The EESEE story "Stress among Pets and Friends" describes the results. Fifteen of the subjects were randomly assigned to each of three groups to do a stressful task alone (the control group), with a good friend present, or with their dog present. The subject's mean heart rate during the task is one measure of the effect of stress. Table 28.2 displays the data. Are there significant differences among the mean heart rates under the three conditions?

28.7 Using the Minitab output in Figure 28.4, verify the values for the sample contrast $\hat{L}_2$ and its standard error given in Example 28.7. Give a 95% confidence interval for the population contrast L_2. Carry out a test of the hypothesis $H_0 : L_2 = 0$ against the two-sided alternative. Be sure to state your conclusions in the setting of the study.

28.11 A student project measured the increase in the heart rates of fellow students when they stepped up and down for three minutes to the beat of a metronome. The explanatory variables are step height (Lo = 5.75 inches, Hi = 11.5 inches) and metronome beat (Slow = 14 steps/minute, Med = 21 steps/minute, Fast = 28 steps/minute). The subject's heart rate was measured for 20 seconds before and after stepping. The response variable is the increase in heart rate during exercise. The data appear in Table 28.3.

(a) Display the 6 treatments in a two-way layout.
(b) Find the group means, plot the means, and discuss the interaction and the two main effects.

Table 28.3 Increase in heart rate due to exercise in six treatment groups

Lo/Slow	Lo/Med	Lo/Fast	Hi/Slow	Hi/Med	Hi/Fast
15	21	24	39	45	66
9	24	42	33	27	60
6	15	27	15	24	51
9	15	48	15	39	30
0	18	18	16	6	57

28.13 The researchers who conducted the study in the previous exercise also recorded the number of times each of three types of behavior (object play, locomotor play, and social play) occurred. The file *ex28-13.dat* contains the counts of social play episodes by each rat during the observation period. Use two-way ANOVA to analyze the effects of gender and housing.

28.29 Exercise 24.33 (text page 661) describes a study on the rate at which the skin of newts heals under the body's natural electrical field (the 1 group in Table 24.5) and under four levels of electric field that differ from the natural level (the 0, 0.5, 1.25, and 1.5 groups). Carry out a one-way ANOVA to compare the mean healing rates. Then perform Tukey multiple comparisons for the 10 pairs of population means. Use the "underline" method illustrated in Example 28.14 to display the complicated results. What do you conclude?

28.31 The data file *ex28-31.dat* has resting and final heart rates as well as the increase in heart rate for the study described in Exercise 28.11. If the randomization worked well, there should be no significant differences among the 6 groups in mean resting heart rate (variable HRrest in the data file).

(a) How many pairwise comparisons are there among the means of 6 populations?
(b) Use Tukey's method to compare these means at the overall 10% significance level.

28.33 The researchers who conducted the study in the previous exercise also recorded the number of times each of three types of behavior (object play, locomotor play, and social play) occurred. The file *ex28-33.dat* contains the counts of object play episodes for each rat during the observation period. Carry out a complete analysis of the effects of gender and housing type.