Calculator Tips Chapters 9, 10, & 11

9.2 Inferences about Two Proportions

We will be using 2-PropZTest for the hypothesis testing of 2 population proportions and 2-PropZInt for the confidence intervals for 2 population proportions.

How to use 2-PropZTest
Press [STAT] → select TESTS → Select 2-PropZTest
- x1: number of successes in first sample
- n1: size of first sample
- x2: number of successes in second sample
- n2: size of second sample
- p1: what your alternative hypothesis is

Select Calculate and hit enter.

If the problem gives you the values for $\hat{p}_1$ *and* $\hat{p}_2$ but not x_1 *and* x_2, you can find x_1 *and* x_2 by using the formulas

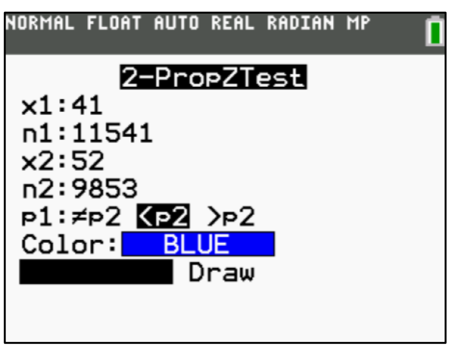
$$x_1 = (\hat{p}_1) \cdot (n_1)\,, \quad x_2 = (\hat{p}_2) \cdot (n_2)$$

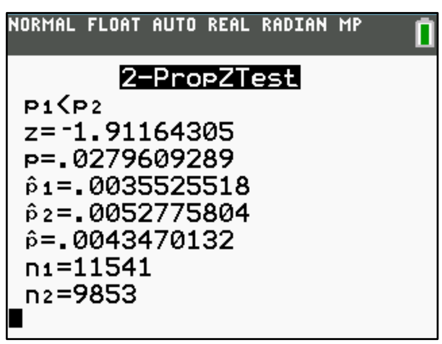
IMPORTANT: x_1 and x_2 must be whole numbers (round to the nearest whole number if you end up with a decimal).

Example 1 (page 452)
Do Airbags Save Lives? The table below lists results from a simple random sample of front-seat occupants involved in car crashes (based on data from “Who Wants Airbags?” by Meyer and Finney, *Chance*, Vol. 18, No. 2). Use a 0.05 significance level to test the claim that the fatality rate of occupants is lower for those in cars equipped with airbags.

	Airbag Available	No Airbag Available
Occupant Fatalities	41	52
Total number of occupants	11,541	9,853

Note that we are using the Airbag Available column as sample 1 and the No Airbag Available column as sample 2. Since The claim that the fatality rate of occupants is lower for those in cars equipped with airbags, this means that p1 < p2 for our alternative hypothesis (we view p as the proportion of the sample being fatal, so we want the population proportion of those with airbags to be smaller than those without).





z= -1.91164305 is the test statistic
p=.0279609289 is the p-value (not the proportion)
$\hat{p}$=.0043470132 is the pooled sample proportion
Note that this will not help you find the critical value. You will need to do that by finding the critical value on the z-scores tables or by using invNorm(

How to use 2-PropZInt
Press [STAT] → select TESTS → Select 2-PropZInt
- x1: number of successes in first sample
- n1: size of first sample
- x2: number of successes in second sample
- n2: size of second sample
- C-Level: Confidence Level (As a decimal)

Select Calculate and hit enter.

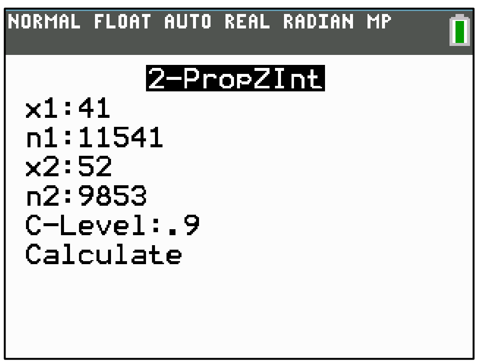
If the problem gives you the values for $\hat{p}_1$ and $\hat{p}_2$ but not x_1 and x_2, you can find x_1 and x_2 by using the formulas

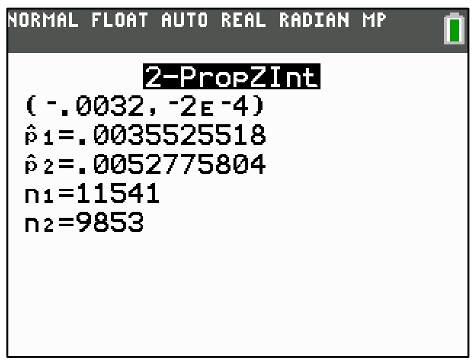
$$x_1 = (\hat{p}_1) \cdot (n_1)\,, \quad x_2 = (\hat{p}_2) \cdot (n_2)$$

IMPORTANT: x_1 and x_2 must be whole numbers (round to the nearest whole number if you end up with a decimal).

Example 2 (page 452)
Use the sample data given in Example 1 to construct a 90% confidence interval estimate of the difference between the two population proportions. What does the result suggest about the effectiveness of airbags in an accident?

	Airbag Available	No Airbag Available
Occupant Fatalities	41	52
Total number of occupants	11,541	9,853





The confidence interval is (-.0032, -.00024).

Because the interval does not contain 0, and that both endpoints are negative, then $p_1 - p_2 < 0$. Solving the inequality for p_1 by $adding$ p_2 on $both$ $sides$ and we $have$ $p_1 < p_2$. So the fatality rate is lower for occupants in cars with airbags than for the occupants in cars with no airbags.

9.3 Inferences about Two Means: Independent Samples

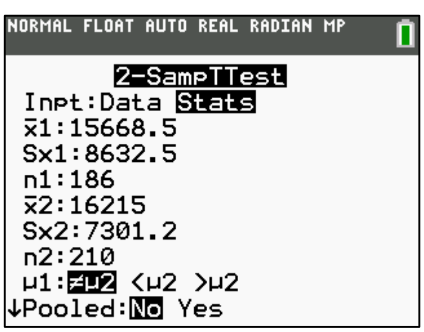
If the samples are **independent**, we will be using 2-SampTTest for hypothesis testing and 2-SampTInt for confidence intervals. We will not be using 2-SampZTest or 2-SampZInt because σ is unknown for all of the homework and quizzes for Chapter 9.

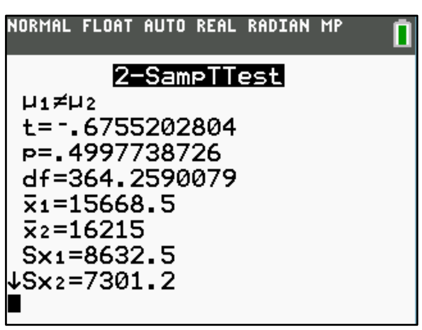
2-SampTTest

1. Press [STAT] → TESTS → 2-SampTTest
2. You select Data if you are given the data values for both samples but not the statistics (i.e. you were not given the sample mean or sample standard deviation for both samples) or select Stats if you are given the sample mean and sample standard deviation for both samples.
    a. If you are given a list of data values for both samples, you need to type in the first sample into L1, and then type in the second sample into L2. You can edit the lists by pressing [STAT] → Edit and press enter.
        i. Then going back to the 2-SampTTest, select Data. Then make sure you have L1 set for List1:, L2 for List2:, Freq1: and Freq2: should both be set to 1, and then select your alternative hypothesis.
        ii. Pooled should be set to no if $\sigma_1 \neq \sigma_2$. If $\sigma_1 = \sigma_2$, then Pooled should be set to yes. Sometimes the problem will also state to set pooled to yes.
    b. If you select Stats (meaning you are given the sample means and sample standard deviations, you type in the sample means, sample standard deviations, sample size, your alternative hypothesis, and then whether or not it should be Pooled.
        i. Pooled should be set to no if $\sigma_1 \neq \sigma_2$. If $\sigma_1 = \sigma_2$, then Pooled should be set to yes. Sometimes the problem will also state to set pooled to yes.
    c. Select Calculate and press enter.

Example 4 (page 463)
Are Men and Women Equal Talkers? A headline in *USA Today* proclaimed that "Men, women are equal talkers." That headline referred to a study of the numbers of words that samples of men and women spoke in a day. Given below are the results from the study. Use a 0.05 significance level to test the claim that men and women speak the same mean number of words in a day. Does there appear to be a difference?

Number of Words Spoken in a Day	
Men	Women
$n_1 = 186$	$n_2 = 210$
$\bar{x}_1 = 15{,}668.5$	$\bar{x}_2 = 16{,}215.0$
$s_1 = 8632.5$	$s_2 = 7301.2$





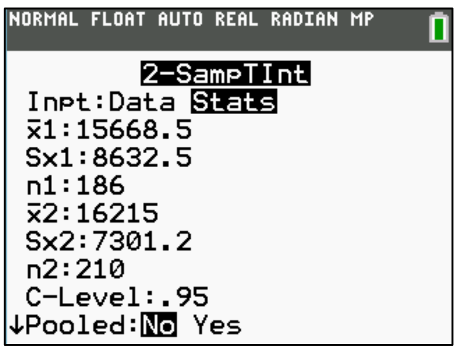
t= -.6755202804 is the test statistic
p=.4997738726 is the p-value
df=364.2590079 is the degrees of freedom (formula 9-1, page 463).

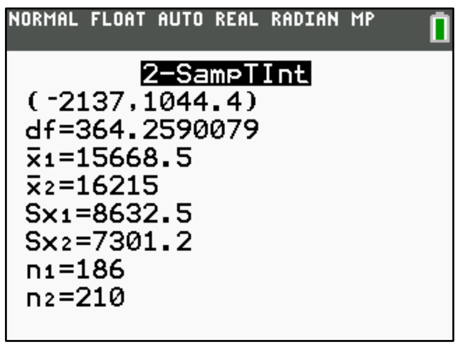
Alternatively, you can use df = smaller of $n_1 - 1$ *and* $n_2 - 1$

Example 5 (page 465)

Using the sample data given in Example 4 (page 463), construct a 95% confidence interval estimate of the difference between the mean number of words spoken by men and the mean number of words spoken by women.

Number of Words Spoken in a Day	
Men	Women
$n_1 = 186$	$n_2 = 210$
$\bar{x}_1 = 15{,}668.5$	$\bar{x}_2 = 16{,}215.0$
$s_1 = 8632.5$	$s_2 = 7301.2$





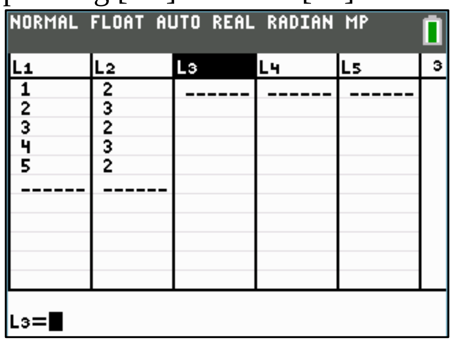
(-2137, 1044.4) is the confidence interval.

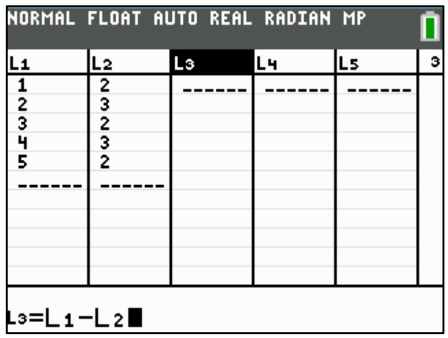
Note: df=364.2590079 is the degrees of freedom (formula 9-1, page 463).

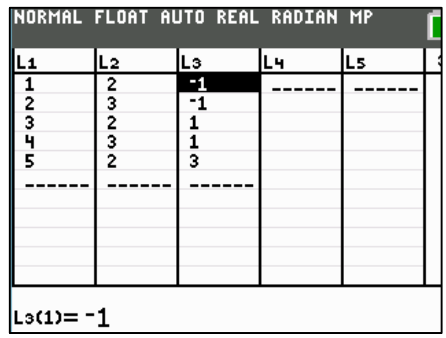
Alternatively, you can use df = smaller of $n_1 - 1$ *and* $n_2 - 1$

9.4 Inferences from Dependent Samples

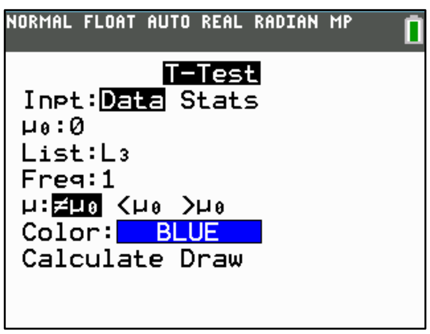
If the samples are **dependent**, we will be using T-Test for hypothesis testing and T-Interval for confidence intervals. With dependent samples, we first calculate each individual difference between the two values in each pair. This can easily be done by entering the sample 1 values in L1, entering the sample 2 values in L2, then subtracting L1-L2 to make a new list L3. To do all of this, Press [STAT] → select Edit… and press enter. Just for practice, type in 1, 2, 3, 4, 5 into L1 and then type 2, 3, 2, 3, 2 into L2. Once you've done this, move the cursor to highlight L3 at the top. Press enter and you should now see L3= at the bottom of the screen with a blinking cursor (see the first image). Now type L1-L2 and press enter again. Remember, you type in L1 by pressing [2nd] and then [1]. L2 can be typed in by pressing [2nd] and then [2].







When using the T-Test or the T-Interval, you select Data and make sure to change the List: so it is using the list of the differences (L3). You use T-Test and T-Interval the same as you did in Chapter 8.

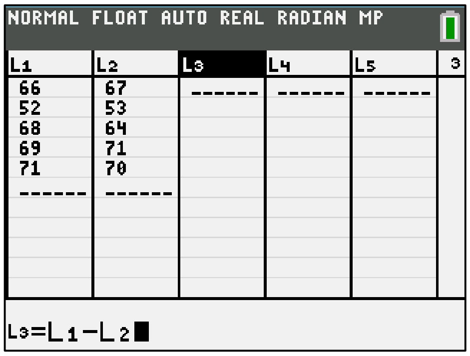


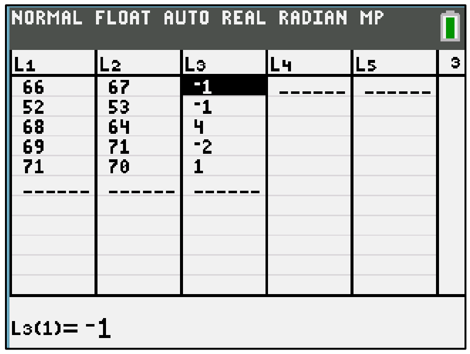
Example 1 (page 476)

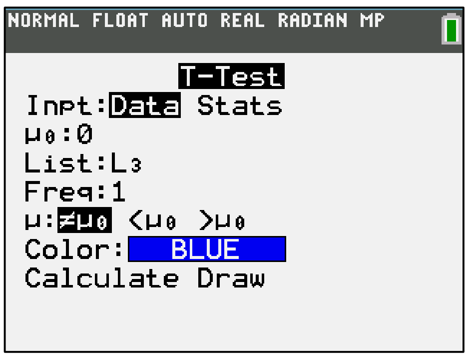
Hypothesis Test of Claimed Freshman Weight Gain. Data Set 3 in Appendix B includes measured weights of college students in September and April of their freshman year. The table below lists a small portion of those sample values to better illustrate the method of hypothesis testing. Use the sample data with a 0.05 significance level to test the claim that for the population of students, the mean change in weight from September to April is equal to 0 kg.

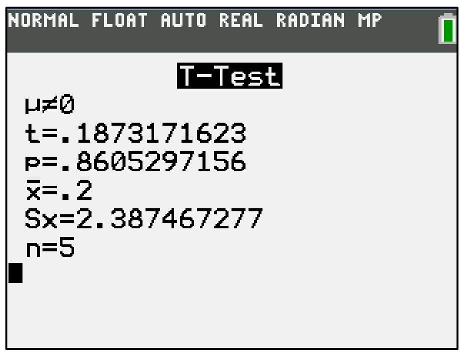
Weight (kg) Measurements of Students in their Freshman Year					
April weight	66	52	68	69	71
September weight	67	53	64	71	70
Difference d = (April weight) – (September weight)	-1	-1	4	-2	1

Because this table already found the differences, we don't have to use the April weight or the September weight. For demonstration purposes, to find the differences between the April weight (sample 1) and the September weight (sample 2), enter the sample 1 on L1 and then enter sample 2 in L2. Then take L1-L2 to get L3.



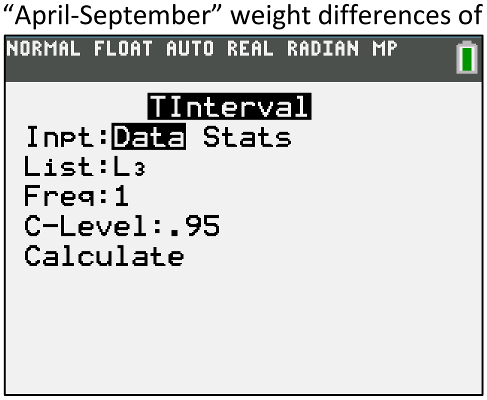


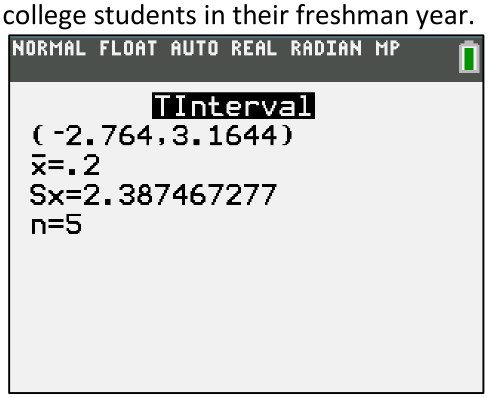




t=.1873171623 is the test statistic
p=.8605297156 is the p-value

Example 3 (page 479)
Using the same paired sample data, construct a 95% confidence interval estimate of μ_d, which is the mean of the "April-September" weight differences of college students in their freshman year.



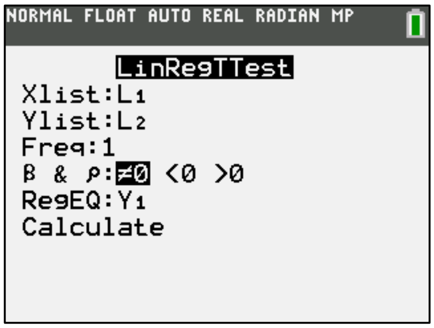


The confidence interval is (-2.764, 3.1644).

Chapter 10 Correlation and Regression

LinRegTTest can be used to find the linear correlation coefficient r as well as the linear regression equation.

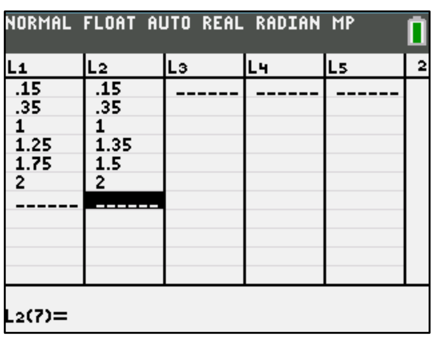
1. Enter the x-values in L1 and enter the y-values in L2
2. Go to [STAT] → TESTS → LinRegTTest
3. Xlist: L1
   Ylist: L2
   Freq: 1
   ≠ 0
   ReqEQ: Y_1 (you can find Y_1 by Pressing [VARS] → Y-Vars → Function → Y_1)

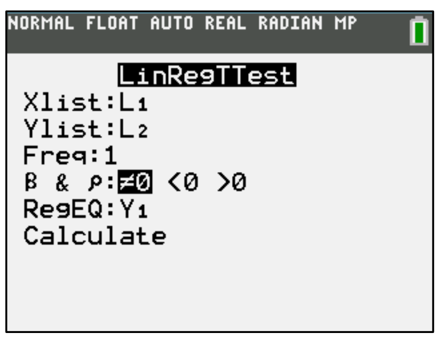


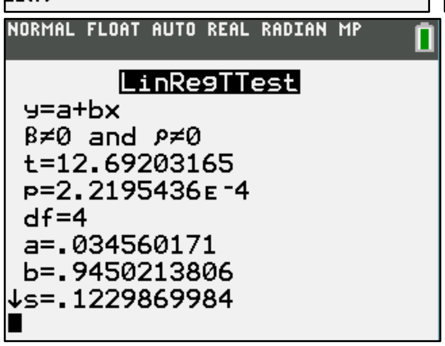
Example 1 (page 499)

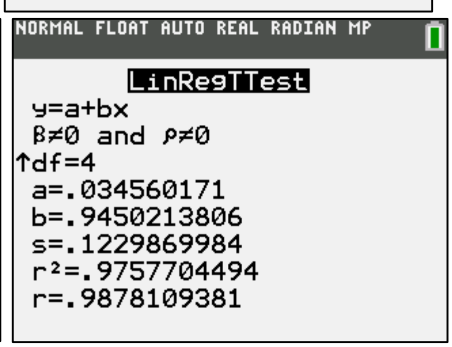
The paired pizza/subway fare costs are shown below. Use the calculator to find the value of the linear correlation coefficient r for the paired sample data.

Costs of a Slice of Pizza and Subway Fare (in dollars)						
Cost of Pizza	0.15	0.35	1.00	1.25	1.75	2.00
Subway Fare	0.15	0.35	1.00	1.35	1.50	2.00







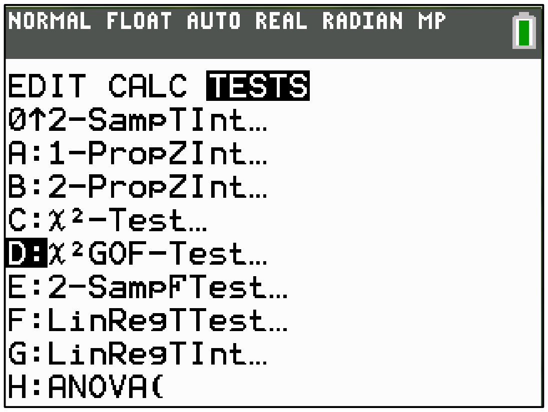


t is the test statistic
p is the p-value
df is the degrees of freedom (df = n-2)
a = .034560171 is the y-intercept of the regression equation and b = .9450213806 is the slope of the regression line.
So the regression line is y = a + bx =.034560171 + (.9450213806)x
r = .9878109381 is the linear correlation coefficient
$r^2 = .9757704494$ is the proportion of the variation in y that is explained by the linear relationship between x and y.

Chapter 11

11-2 Goodness-of-Fit

Most TI-83 calculator's do not have $\chi^2 GOF - Test$ on their calculator. To determine if your calculator has it, press [STAT] → TESTS and then look for $\chi^2 GOF - Test$

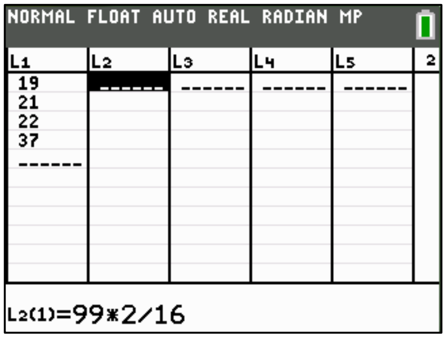


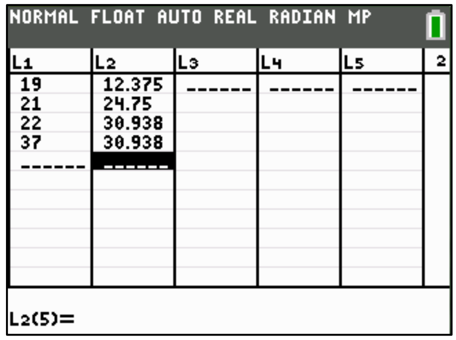
If it is not on your calculator, you will need to program it on there. Refer to the How to Program the Chi-Square Goodness of Fit document for directions on how to do this.

First you type the observed frequency into L1 and the expected frequency into L2. Once you do that, go to [STAT] → TESTS → $\chi^2 GOF - Test$. After selecting $\chi^2 GOF - Test$, you will need to set Observed: to L1 and Expected to L2. df: should equal one less the number of categories

Example 2 (page 560)
World Series Games. The table below lists the numbers of games played in the baseball World Series, as of the writing of the text book. The table includes the expected proportions for the numbers of games in a World Series, assuming that in each series, both teams have about the same chance of winning. Use a 0.05 significance level to test the claim that the actual numbers of games fit the distribution indicated by the probabilities.

Numbers of Games in World Series Contests				
Games Played	4	5	6	7
Actual World Series Contests	19	21	22	37
Expected Proportion	2/16	4/16	5/16	5/16

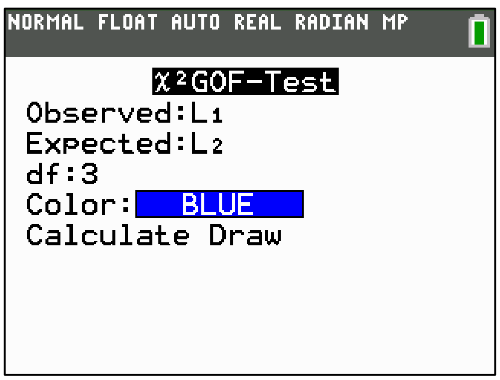


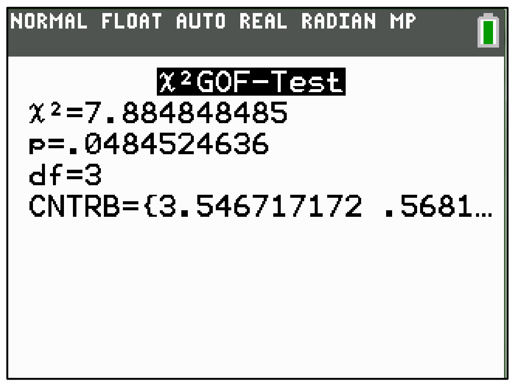


Observed Frequency O	Expected Frequency $E = np$
19	$99 \cdot \frac{2}{16} = 12.3750$
21	$99 \cdot \frac{4}{16} = 24.7500$
22	$99 \cdot \frac{5}{16} = 30.9375$
37	$99 \cdot \frac{5}{16} = 30.9375$

<u>L1 is ALWAYS the observed frequency</u> and <u>L2 is the expected frequency</u>. 99 comes from the total number of World Series Contests (19+21+22+37=99).

After selecting $\chi^2 GOF - Test$, you will need to set Observed: to L1 and Expected to L2. df: should equal one less the number of categories, (so df = 4 – 1 = 3).





$\chi^2 = 7.884848485$ is the test statistic

p = .0484524636 is the p-value

11.3 Contingency Tables

For contingency tables, we will be using the χ^2- Test

1. Open the matrix menu by pressing [2ND] and then [x^{-1}]
2. Select EDIT, then press [ENTER].
3. Enter the dimensions of the matrix
    - first number is the number of rows, second number is the number of columns
4. Now enter all of the observed data values into the matrix exactly as it appears in the table
5. When done, press [STAT], then go to TESTS > χ^2- Test. For Observed, be sure to select the same matrix that you entered the data (you do this by pressing [2ND] and then [x^{-1}], then select the matrix you want). Expected: can be set to any other matrix as the calculator will automatically fill in the expected values into that matrix.

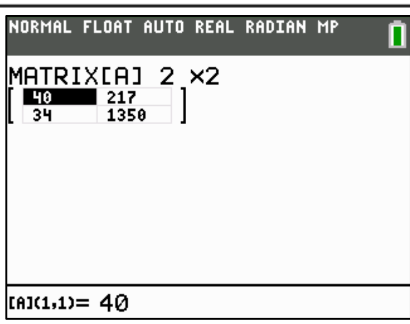
Example 4 (page 572)

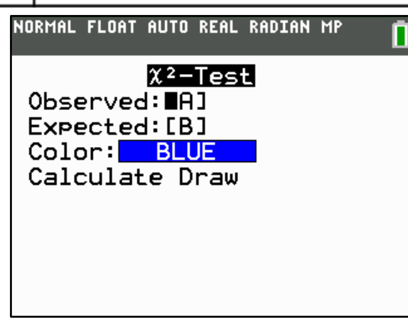
Is the Nurse a Serial Killer?

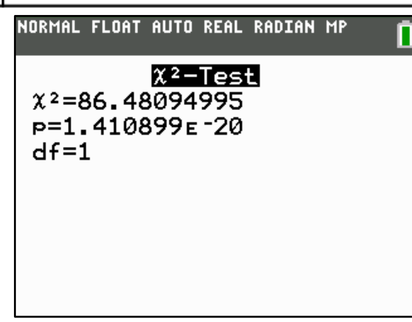
Three alert nurses at the Veteran's Affairs Medical Center in Northampton, Massachusetts noticed an unusually high number of deaths at times when another nurse, Kristen Gilbert, was working. Those same nurses also noticed missing supplies of the drug epinephrine (a synthetic adrenaline that stimulates the heart). This eventually led to an investigation on whether or not Kristen Gilbert was a serial killer.

We are going to test the claim that whether Gilbert was on duty for a shift is independent of whether a patient died during the shift. Because this is a serious analysis, we will use a significance level of 0.01. What does the result suggest about the charge that Gilbert killed her patients?

	Shifts with a death	Shifts without a death
Gilbert Was Working	40	217
Gilbert Was Not Working	34	1350







χ^2 is the test statistic

p is the p-value

df is the degrees of freedom

11-4 Analysis of Variance

When you are testing the equality of three or more population means by analyzing sample variances, we will use the ANOVA(on the calculator.

1. Enter each sample into a list (L1, L2, L3, etc…)
2. Then press [STAT] → TESTS → ANOVA(
3. Enter each list of data being used, separating each by a comma
    a. ANOVA(L1, L2, L3, ….)

Example 1 (page 583)

Use the chest deceleration measurements listed in the table below to test the claim that the three samples come from populations with means that are all equal.

Chest Deceleration Measurements (in g) from Car Crash Tests										
Small Cars	44	43	44	54	38	43	42	45	44	50
Medium Cars	41	49	43	41	47	42	37	43	44	54
Large Cars	32	37	38	45	37	33	38	45	43	42