CHAPTER 11 CHI-SQUARE AND *F* DISTRIBUTIONS

CHI-SQUARE TEST OF INDEPENDENCE (SECTION 11.1 OF *UNDERSTANDABLE STATISTICS*)

The TI-83 Plus and TI-84 Plus calculators support tests for independence. Press the STAT key, select **TESTS,** and option **C:** χ^2 **–Test.** The original observed values need to be entered in a matrix.

Example

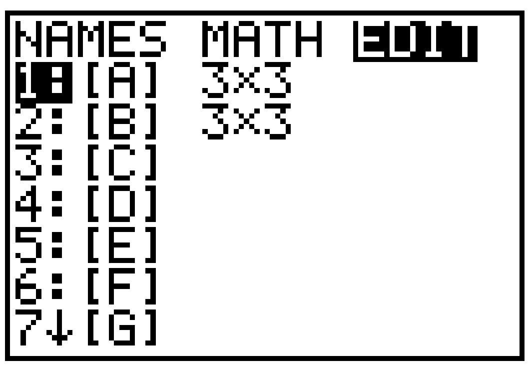
A computer programming aptitude test has been developed for high school seniors. The test designers claim that scores on the test are independent of the type of school the student attends: rural, suburban, urban. A study involving a random sample of students from each of these types of institutions yielded the following information, where aptitude scores range from 200 to 500 with 500 indicating the greatest aptitude and 200 the least. The entry in each cell is the observed number of students achieving the indicated score on the test.

School Type

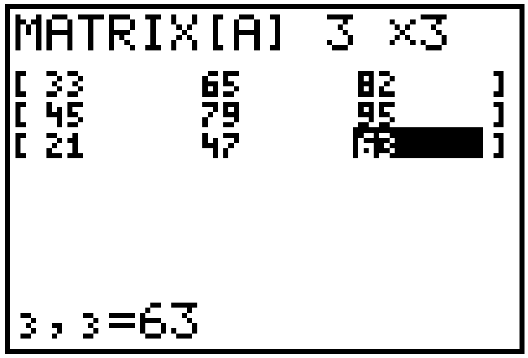
Score	Rural	Suburban	Urban
200–299	33	65	82
300–399	45	79	95
400–500	21	47	63

Using the option **C:** χ^2 **-Test…,** test the claim that the aptitude test scores are independent of the type of school attended at the 0.05 level of significance.

We need to use a matrix to enter the data. Press 2nd **[MATRX].** Highlight **EDIT,** and select **1:[A].** Press ENTER.

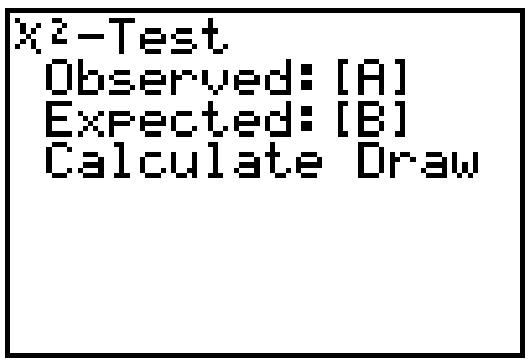


Since the contingency table has 3 rows and 3 columns, type 3 (for number of rows), press ENTER, type 3 (for number of columns), press ENTER. Then type in the observed values. Press ENTER after each entry. Notice that you enter the table by rows. When the table is entered completely, press 2nd **[MATRX]** again.

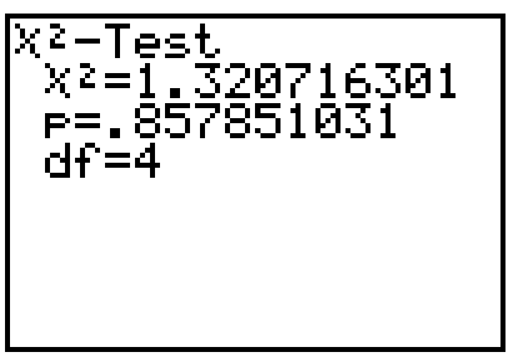


This time, select **2:[B].** Set the dimensions to be the same as for matrix [A]. For this example, [B] should be 3 rows × 3 columns.

Now press the STAT key, highlight TESTS, and use option **C:** χ^2 **–Test.** This tells you that the observed values of the table are in matrix [A]. The expected values will be placed in matrix [B].



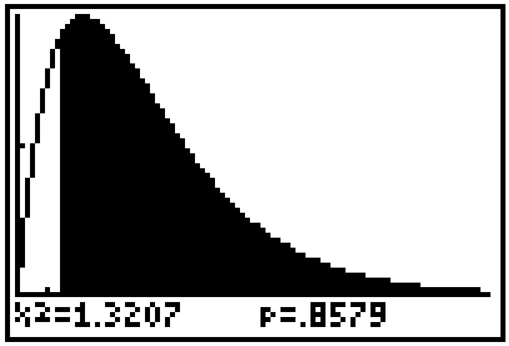
Highlight **Calculate** and press ENTER. We see that the sample value of χ^2 is 1.32. The *P* value is 0.858. Since the *P* value is larger than 0.05, we do not reject the null hypothesis.



If you want to see the expected values in matrix [B], type 2nd **[MATRX]** and select **[B]** under **NAMES.**

Graph

Notice that one of the options of the χ^2 test is to graph the χ^2 distribution and show the sample test statistic on the graph. Highlight **Draw** and press ENTER. Before you do this, be sure that all STAT PLOTS are set to Off, and that you have cleared all the entries in the Y= menu.



LAB ACTIVITIES FOR CHI-SQUARE TEST OF INDEPENDENCE

1. We Care Auto Insurance had its staff of actuaries conduct a study to see if vehicle type and loss claim are independent. A random sample of auto claims over six months gives the information in the contingency table.

Total Loss Claims per Year per Vehicle

Type of vehicle	\$0–999	\$1000–2999	\$3000–5999	\$6000+
Sports car	20	10	16	18
Truck	16	25	33	9
Family Sedan	40	68	17	7
Compact	52	73	48	12

Test the claim that car type and loss claim are independent. Use $\alpha = 0.05$.

2. An educational specialist is interested in comparing three methods of instruction:

SL – standard lecture with discussion

TV – video taped lectures with no discussion

IM – individualized method with reading assignments and tutoring, but no lectures.

The specialist conducted a study of these three methods to see if they are independent. A course was taught using each of the three methods and a standard final exam was given at the end. Students were put into the different method sections at random. The course type and test results are shown in the next contingency table.

Final Exam Score

Course Type	< 60	60–69	70–79	80–89	90–100
SL	10	4	70	31	25
TV	8	3	62	27	23
IM	7	2	58	25	22

Test the claim that the instruction method and final test scores are independent, using $\alpha = 0.01$.

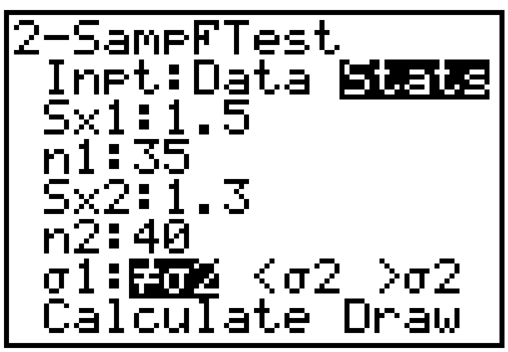
TESTING TWO VARIANCES (SECTION 11.5 OF *UNDERSTANDABLE STATISTICS*)

Under the **TESTS** menu option **D:2-SampFTest** provides hypothesis testing for two variances. Again there is a choice of data entry styles: raw data in lists, or summary statistics for which you provide the sample standard deviations and sample sizes.

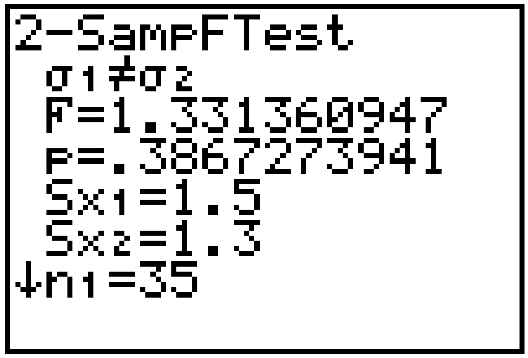
Example

Two paint manufacturing processes are under study. A random sample of 35 applications of the paint produced under the first method shows that the paint's life has a standard deviation of 1.5 years. For the second method, a random sample of 40 applications shows a standard deviation of 1.3 years. Use a 5% level of significance to test if the variances are equal.

Use the **D:2-StampFTest** option. Select **Stats** and enter the data as requested. Note that we follow the convention of always entering the larger standard deviation as S_{x1}.



Highlight **Calculate** and press ENTER.



The sample *F* statistic is 1.33 with *P* value 0.387. We do not reject H_0.

ONE-WAY ANOVA (SECTION 11.5 OF *UNDERSTANDABLE STATISTICS*)

The TI-83 Plus and TI-84 Plus support one-way ANOVA tests. Press the STAT key, and under **TESTS** select option **F:ANOVA(.** The format of the command for ANOVA is **ANOVA** $\mathbf{(L_1, L_2, \ldots)}$ where all the treatment lists are given.

Example

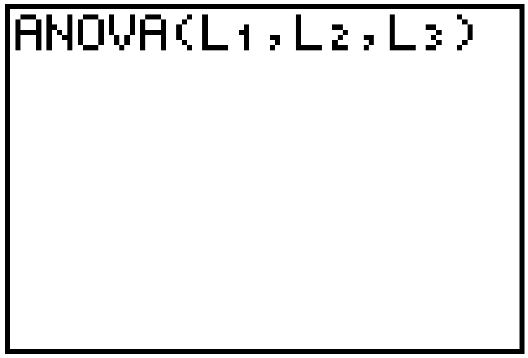
A psychologist has developed a series of tests to measure a person's level of depression. The composite scores range from 50 to 100, with 100 representing the most severe depression level. A random sample of 12 patients with approximately the same depression level (as measured by the tests) was divided into 3 different treatment groups. Then, one month after treatment was completed, the depression level of each patient was again evaluated. The after-treatment depression levels are given.

Treatment 1	70	65	82	83	71
Treatment 2	75	62	81		
Treatment 3	77	60	80	75	

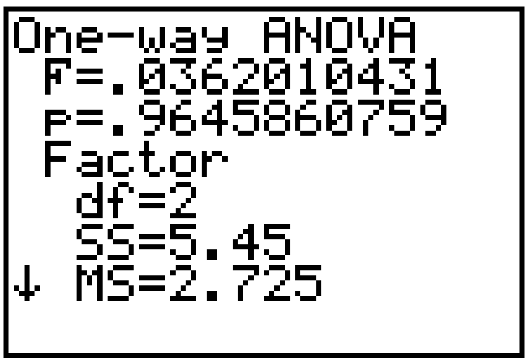
Use the option **F: ANOVA** to test the claim that the population means are all the same, at the 5% level of significance.

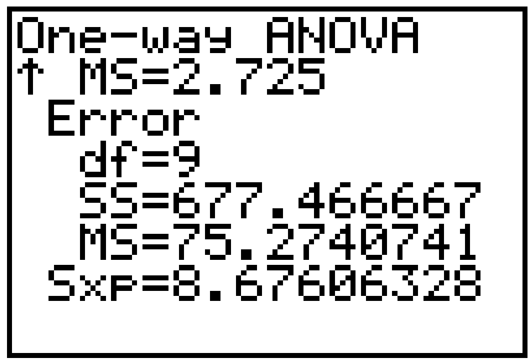
Enter the treatment data in lists L_1, L_2, and L_3 respectively.

Select the **F:ANOVA (** option from the **TESTS** menu. Then type the names of the 3 lists containing the data. Separate the list names by commas.



Press [ENTER]. The output is on two screens.





We see that the sample F is $F = 0.036$ with P value 0.9646. We do not reject H_0 and conclude that there is no evidence that some of the means are different.

LAB ACTIVITIES FOR ANALYSIS OF VARIANCE

1. A random sample of 20 overweight adults was randomly divided into 4 groups, and each group was given a different diet plan. The weight loss for each individual (in pounds) after 3 months follows.

Plan 1	18	10	20	25	17
Plan 2	28	12	22	17	16
Plan 3	16	20	24	8	17
Plan 4	14	17	18	5	16

Test the claim that the population mean weight loss is the same for the four diet plans, at the 5% level of significance.

2. A psychologist is studying the time it takes rats to respond to stimuli after being given doses of different tranquilizing drugs. A random sample of 18 rats was divided into 3 groups. Each group was given a different drug. The response time to stimuli was measured (in seconds). The results follow.

Drug A	3.1	2.5	2.2	1.5	0.7	2.4
Drug B	4.2	2.5	1.7	3.5	1.2	3.1
Drug C	3.3	2.6	1.7	3.9	2.8	3.5

Test the claim that the population mean response times for the three drugs are the same, at the 5% level of significance.

3. A research group is testing various chemical combinations designed to neutralize and buffer the effects of acid rain on lakes. Eighteen lakes of similar size in the same region have all been affected in the same way by acid rain. The lakes are divided into four groups and each group of lakes is sprayed with a different chemical combination. An acidity index is then taken after treatment. The index ranges form 60 to 100 with 100 indicating the greatest acid rain pollution. The results follow.

Combination I	63	55	72	81	75
Combination II	78	56	75	73	82
Combination III	59	72	77	60	
Combination IV	72	81	66	71	

Test the claim that the population mean acidity index after each of the four treatments is the same, at the 0.01 level of significance.