Chapter 13
Categorical Data Analysis

Testing Categorical Probabilities: One-Way Table

Example 13.2 A One-Way χ^2 Test – Effectiveness of a TV Program on Marijuana

Problem: Suppose an educational television station has broadcast a series of programs on the physiological and psychological effects of smoking marijuana. Now that the series is finished, the station wants to see whether citizens within the viewing area have changed their minds about how the possession of marijuana should be considered legally. Before the series was shown, it was determined the 7% of the citizens favored legalization, 18% favored decriminalization, 65% favored the existing law (an offender could be fined or imprisoned) and 10% had no opinion.

A summary of the opinions after the series was shown of a random sample of 500 people in the viewing area is given in Table 13 – 1. Test at the $\alpha = 0.01$ level to see whether these data indicate that the distribution of opinions differs significantly from the proportions that existed before the educational series was aired.

Table 13 – 1

Distribution of Opinions about Marijuana Possession			
Legalization	Decriminalization	Existing Law	No Opinion
39	99	336	26

Solution: The TI-84 Plus has a built in procedure to run a χ^2 goodness of fit test. However, the TI-83 Plus does not have a built in procedure.

Let p_1 = Proportion of all citizens favoring legalization
Let p_2 = Proportion of all citizens favoring decriminalization
Let p_3 = Proportion of all citizens favoring existing law
Let p_4 = Proportion of all citizens with no opinion

Then the null and alternative hypotheses are:

H_0 $p_1 = 0.07$, $p_2 = 0.18$, $p_3 = 0.65$, $p_4 = 0.10$
H_a: at least one of the proportions differs from its null hypothesis value

Using the TI-84 Plus

1. Enter the observed frequencies into list L1. The observed values are given in Table 13 – 1.

2. Enter the expected frequencies into list L2. Note: The expected frequencies are based on the total sample size. In this example, the sample size is 500. You must multiply each of the relative frequencies in the null hypothesis by 500. The expected frequencies are:

$E_1 = 500(0.07) = 35$
$E_2 = 500(0.18) = 90$
$E_3 = 500(0.65) = 325$
$E_4 = 500(0.10) = 50$

Once both the observed and expected frequencies are entered, your screen should look like Figure 13 – 1.

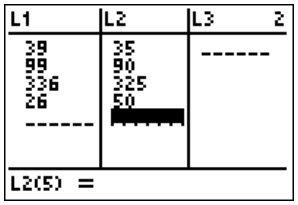

Figure 13 – 1

3. Press [STAT] and arrow over to the TESTS menu. Arrow down to D: χ^2 GOF or press [ALPHA] D. Then press [ENTER]. See Figure 13 – 2.

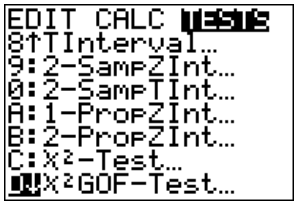

Figure 13 – 2

4. Enter the observed list by pressing [2nd] L1 [ENTER]. See Figure 13 – 3.

5. Enter the observed list by pressing [2nd] L2 [ENTER]. See Figure 13 – 3.

6. Enter the degrees of freedom which is defined as df = k-1 where k is the number of possible categories for the variable. In this example, df = 4 - 1 =3. Press [ENTER]. See Figure 13 – 3 for the completed screen.

7. Highlight Calculate and press ENTER. The screen will appear as in Figure 13 – 4.

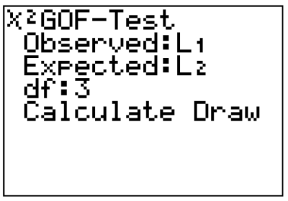

Figure 13 – 3

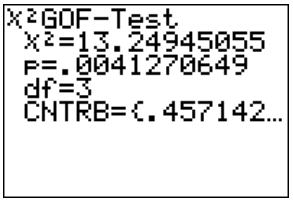

Figure 13 – 4

From the screen in Figure 13 – 4, we see the test statistic is $\chi^2 = 13.25$ and the p-value is 0.004. Since the p-value is less than $\alpha = 0.01$ we reject H_0 and conclude there is sufficient evidence that the opinions on the legalization of marijuana have changed since the series was aired.

Testing Categorical Probabilities: Two-Way (Contingency) Table

Example 13.3 Conducting a Two-Way Analysis: Marital Status and Religion

Problem: A social scientist wants to determine whether the marital status (divorced or not divorced) of U.S. men is independent of their religious affiliation (or lack thereof). A sample of 500 men is surveyed, and the results are shown in Table 13 – 2. Test to see whether there is sufficient evidence to indicate that the marital status of men who have been or are currently married is dependent on religious affiliation. Use $\alpha = 0.01$.

Table 13 – 2

	Religious Affiliation					
	A	B	C	D	None	Total
Divorced	39	19	12	28	18	116
Married, Never Divorced	172	61	44	70	37	384
Total	211	80	56	98	55	500

Solution: The hypotheses are:

H_0: The marital status of U.S. men and their religious affiliations are independent
H_a: The marital status of U.S. men and their religious affiliations are dependent

In order to run the chi-square test, the table of observed frequencies must be placed in a matrix in the calculator. Only the observed frequencies are entered, not the totals.

1. Press [2nd] MATRX, arrow over to EDIT, and press [ENTER]. See Figure 13 – 5.

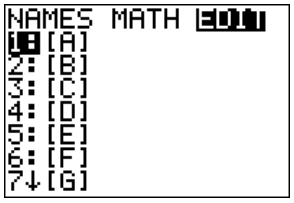

Figure 13 – 5

2. Enter the number of rows, here 2, and press [ENTER]. See Figure 13 – 6.

3. Enter the number of columns, here 5, and press [ENTER]. See Figure 13 – 6.

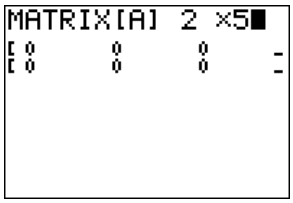

Figure 13 – 6

4. Begin entering your data values going across the rows and pressing [ENTER] after <u>each</u> entry. Be sure you do not enter the totals, and that you press [ENTER] after the last entry. See Figure 13 – 7 for final screen.

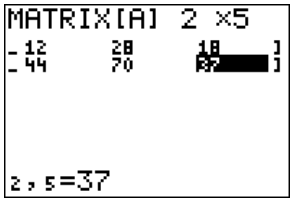

Figure 13 – 7

5. Exit the matrix edit mode by pressing [2nd] QUIT.

Now we are ready to run the test. The TI-83/84 Plus has several matrices and we have used matrix A for this example.

6. Press [STAT] and arrow over to TESTS.

7. Press [ALPHA] C or arrow down to **C:** χ^2-Test and press [ENTER]. See Figure 13 – 8.

Figure 13 – 8

8. For Observed enter the matrix containing your observed frequencies. Here that is matrix A so press [2nd] MATRX [1] or press [2nd] MATRX and press [ENTER]. See Figure 13 – 9.

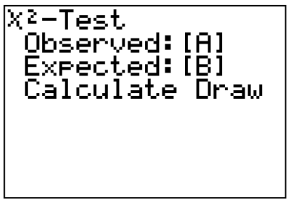

Figure 13 – 9

9. Your calculator will return to the chi-square test menu. Press [ENTER].

10. Enter the matrix where you would like your expected frequencies stored. Here we will use matrix B. Press [2nd] MATRX [2]. See Figure 13 – 9.

11. Your calculator will return to the chi-square test menu. Press [ENTER].
See Figure 13 – 9.

12. Highlight Calculate and press [ENTER]. Your screen will appear as in Figure 13 – 10.

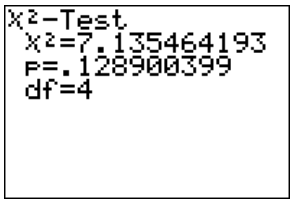

Figure 13 – 10

As can be seen from the screen in Figure 13 – 10, the test statistic is $\chi^2 = 7.1355$ and the p-value is 0.1289; which is not below our α-value of 0.01. We therefore we have sufficient evidence to conclude that marital status and religious affiliation are not dependent events.

Exercise 13.35 IQ and Mental Retardation

Problem: A person is diagnosed with a mental deficiency if, before the age of 18, his or her score on a standard IQ test is no higher than 70 (two standard deviations below the mean of 100). Researchers at Cornell and West Virginia Universities examined the impact of rising IQ scores on diagnoses of mental deficiency (MD) (*American Psychologist*, October, 2003). IQ data were collected from different school districts across the U.S., and the students were tested with either the Wechsler Intelligence Scale for Children – Revised (WISC – R) or the Wechsler Intelligence Scale for Children – Third Revision (WISC – III) IQ tests. The researchers focused on those students with IQs just above the mental deficiency cutoff (between 70 and 85), based on the original IQ test. These "borderline" MD students were then retested one year later with one of the IQ tests. Table 13 – 3 gives the number of students diagnosed with mental deficiency on the basis of the retest. Conduct a chi-square test for independence to determine whether the proportion of students diagnosed with MD depends on the IQ test/retest method. Use $\alpha = 0.01$.

Table 13 – 3
IQ and Mental Deficiency

Test/Retest	Diagnosed with MD	Above MD cutoff IQ	Total
WISC-R/WISC-R	25	167	192
WISC-R/WISC-III/	54	103	157
WISC-III/WISC-III	36	141	177
Total	115	411	526

Solution: The hypotheses are:

H_0: The diagnosis of MD and the IQ test/retest method are independent
H_a: The diagnosis of MD and the IQ test/retest method are dependent

In order to run the chi-square test, the table of observed frequencies must be placed in a matrix in the calculator. Only the observed frequencies are entered, not the totals.

1. Press 2nd MATRX, arrow over to EDIT, and press ENTER. See Figure 13 – 11.

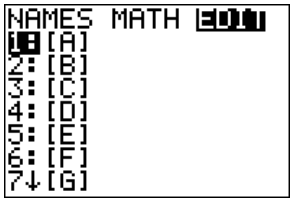

Figure 13 – 11

2. Enter the number of rows, here 3, and press ENTER. See Figure 13 – 12.

3. Enter the number of columns, here 2, and press ENTER. See Figure 13 – 12.

4. Begin entering your data values going across the rows and pressing [ENTER] after each entry. Be sure you do not enter the totals, and that you press [ENTER] after the last entry. See Figure 13 – 12 for final screen.

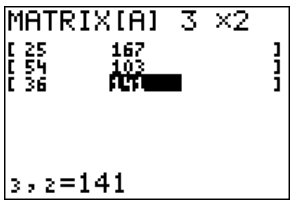

Figure 13 – 12

5. Exit the matrix edit mode by pressing [2nd] QUIT.

Now we are ready to run the test. The TI-83/84 Plus has several matrices and we have used matrix A for this example.

6. Press [STAT] and arrow over to TESTS.

7. Press [ALPHA] C or arrow down to **C:** χ^2-Test and press [ENTER]. See Figure 13 – 13.

Figure 13 – 13

8. For Observed enter the matrix containing your observed frequencies. Here that is matrix A so press [2nd] MATRX [1] or press [2nd] MATRX and press [ENTER]. See Figure 13 – 14.

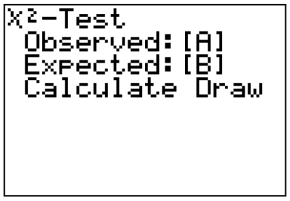

Figure 13 – 14

9. Your calculator will return to the chi-square test menu. Press [ENTER].

10. Enter the matrix where you would like your expected frequencies stored. Here we will use matrix B. Press [2nd] MATRX [2]. See Figure 13 – 14.

11. Your calculator will return to the chi-square test menu. Press [ENTER].
See Figure 13 – 14.

12. Highlight Calculate and press [ENTER]. Your screen will appear as in Figure 13 – 15.

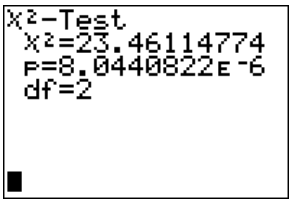

Figure 13 – 15

As can be seen from the screen in Figure 13 – 15, the test statistic is $\chi^2 = 23.46$ and the p-value is 0.0000080440822 which is less than 0.01 so we reject H_0. We therefore conclude that there is sufficient evidence that the Test/retest method and Test score are not independent events.