AP Statistics Notes – Unit Eleven: Inference for Categorical Data
Goodness of Fit Tests

Syllabus Objectives: 3.24 – The student will describe the properties of the Chi-square distribution. 3.25 – The student will solve problems using tables of the Chi-square distribution.

In the previous units, we discussed inference procedures for means and proportions. In some cases, we want to examine the distribution of proportions for a population or determine whether the distribution of one variable has been influenced by another. Chi-square procedures help us in these situations.

- **Univariate Categorical Data**
  - Univariate (one variable) categorical data is best summarized in a **one-way frequency table.**
  - Example: Consider the following observations of sample of faculty status for faculty in a large university system.
  -

	Category				
	Full Professor	Associate Professor	Assistant Professor	Instructor	Adjunct/ Part time
Frequency	22	31	25	35	41

  - A local newsperson might be interested in testing hypotheses about the proportion of the population that fall in each of the categories. For example, the newsperson might want to test to see if the five categories occur with equal frequency throughout the whole university system.
  - For each category, the **expected count** for that category is the product of the total number of observations with the hypothesized proportion for that category.
  - Consider the sample of faculty from the large university system above. If we believe that the five categories will occur with equal frequency, then each frequency would be 0.2. To find the total number of observations, add up the five frequencies to obtain the total → $22+31+25+35+41=154$. Multiply 154 by the frequency to obtain the expected count for each category.
  -

	Category				
	Full Professor	Associate Professor	Assistant Professor	Instructor	Adjunct/ Part time
Frequency	22	31	25	35	41
Hypothesized Proportion	0.2	0.2	0.2	0.2	0.2
Expected Count	30.8	30.8	30.8	30.8	30.8

- **Test for Goodness of Fit**

  - According to the M&M/Mars Company, in 1995, "..*the new mix of colors of M&M's plain chocolate candies will contain 30 percent browns, 20 percent yellows and reds, and 10 percent each of oranges, greens, and blues*." However, the mix of colors has been known to change every few years. We want to see if there is sufficient evidence to reject the company's 1995 claim. To do this, we need the **Chi-square** (χ^2) **Goodness of Fit Test**.
  - A **Goodness of Fit Test** is used to determine whether a population has a certain hypothesized distribution. The data must be **counts** for the categories of a categorical variable. We look at the counts and expected counts and ask, "Are the differences just natural sampling variability, or are they so large that they indicate something important?"
  - Test Procedures:
    - **Hypotheses:** The null hypothesis is that the population proportions are equal to the hypothesized proportions. The alternative is that at least two of the proportions differ from the hypothesized proportions.

      $H_0 : p_1 =$ hypothesized proportion for category 1

      $p_2 =$ hypothesized proportion for category 2

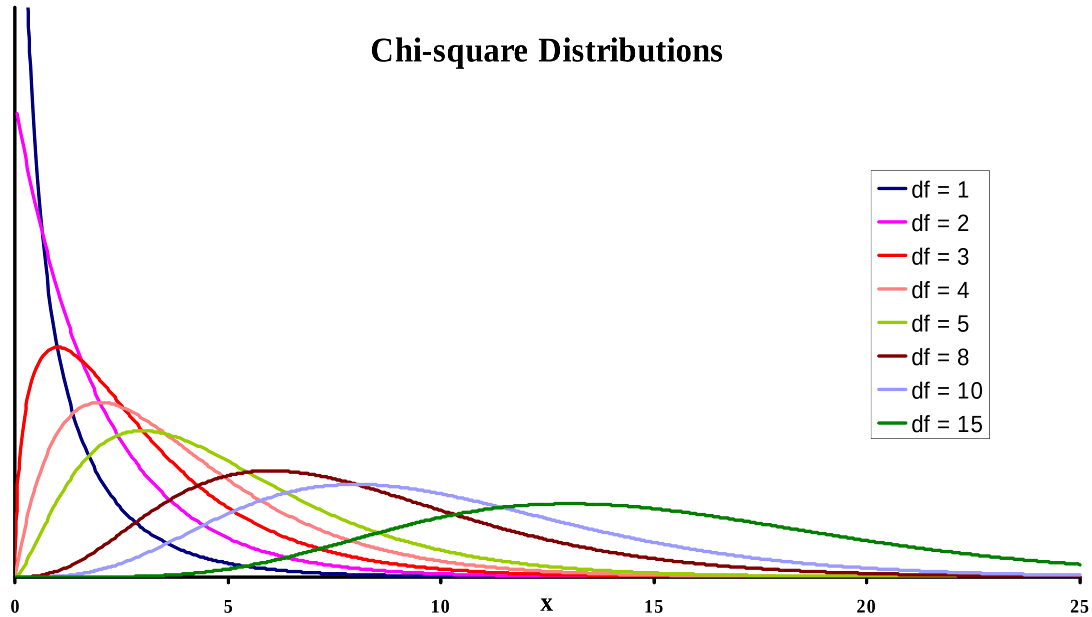
      $\vdots$

      $p_k =$ hypothesized proportin for category k

      $H_a :$ at least one of these proportions is incorrect
    - **Assumptions:**
      1. The data must be *counts* for the categories for a categorical variable.
      2. The counts in the cells should be independent of each other.
      3. Observed cell counts are based on a *random sample*.
      4. The *sample size is large.* The sample size is large enough for the Chi-square test to be appropriate as long as every **expected** cell count is at least 5.
    - **Test Statistic:** The test statistic is found by adding up the sum of the squares of the deviations between the observed and expected counts divided by the expected counts. This test statistic will determine if the observed sample distribution is significantly different from the hypothesized population distribution.

      $$\chi^2 = \sum \left[\frac{(\text{observed cell count - expected cell count})^2}{\text{expected cell count}} \right]$$

      The test statistic has approximately a χ^2 distribution with (*k*-1) degrees of freedom. The test statistic is used only for testing hypotheses, not for constructing confidence intervals. If the observed counts don't match the expected, the statistic will be large. It can't be "too small". That would just mean that our model really fit the data well. So, the Chi-square test is always one sided.
    - **P-values:** The P-value associated with the computed test statistic value is the area to the right of χ^2 under the df = *k*-1 Chi-square curve.

- **Properties of the Chi-square distribution**
  - The **Chi-square** distributions are a family of distributions that take only positive values and are skewed to the right. A specific chi-square distribution is specified by one parameter, called the **degrees of freedom.**
  - Below shows the density curves for several members of the chi-square family of distributions. Notice as the degrees of freedom increase, the density curves become less skewed and larger values become more probable.
  - 

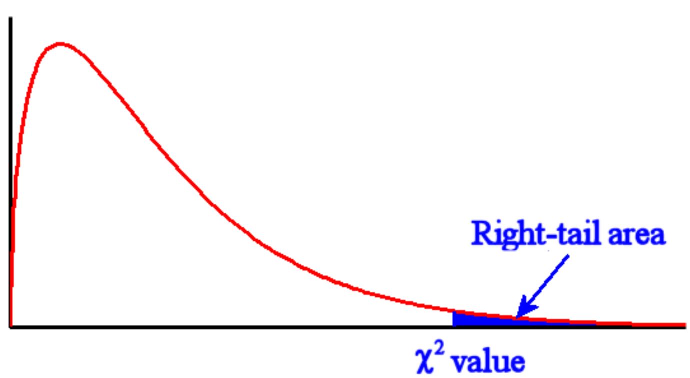
  - The chi-square density curves have the following properties:
    1. The total area under a chi-square curve is equal to 1.
    2. Each chi-square curve (except when df = 1) begins at 0 on the horizontal axis, increases to a peak, and then approaches the horizontal axis asymptotically from above.
    3. Each chi-square curve is skewed to the right.
    4. As the number of degrees of freedom increase, the curve becomes more and more symmetrical and looks more like a normal curve.
    5. Chi-square distributions take only positive values.

- The Chi-square table gives critical values for chi-square distributions. These upper-tail areas give you P-values for a chi-square test.
-

Right-tail area	df = 1	df = 2	df = 3	df = 4	df = 5
> .100	< 2.70	< 4.60	< 6.25	< 7.77	< 9.23
0.100	2.70	4.60	6.25	7.77	9.23
0.095	2.78	4.70	6.36	7.90	9.37
0.090	2.87	4.81	6.49	8.04	9.52
0.085	2.96	4.93	6.62	8.18	9.67
0.080	3.06	5.05	6.75	8.33	9.83
0.075	3.17	5.18	6.90	8.49	10.00
0.070	3.28	5.31	7.06	8.66	10.19
0.065	3.40	5.46	7.22	8.84	10.38
0.060	3.53	5.62	7.40	9.04	10.59
0.055	3.68	5.80	7.60	9.25	10.82
0.050	3.84	5.99	7.81	9.48	11.07
0.045	4.01	6.20	8.04	9.74	11.34
0.040	4.21	6.43	8.31	10.02	11.64
0.035	4.44	6.70	8.60	10.34	11.98
0.030	4.70	7.01	8.94	10.71	12.37
0.025	5.02	7.37	9.34	11.14	12.83
0.020	5.41	7.82	9.83	11.66	13.38
0.015	5.91	8.39	10.46	12.33	14.09
0.010	6.63	9.21	11.34	13.27	15.08
0.005	7.87	10.59	12.83	14.86	16.74
0.001	10.82	13.81	16.26	18.46	20.51
< .001	> 10.82	> 13.81	> 16.26	> 18.46	> 20.51

Right-tail area	df = 6	df = 7	df = 8	df = 9	df = 10
> .100	< 10.64	< 12.01	< 13.36	< 14.68	< 15.98
0.100	10.64	12.01	13.36	14.68	15.98
0.095	10.79	12.17	13.52	14.85	16.16

.



Syllabus Objectives: 4.15 – The student will perform a Chi-square test for goodness of fit, homogeneity of proportions, and independence.

- **Goodness of Fit Test example:**
  - **Example 1** – Let us revisit the example introduced on page one. Consider the newsperson's desire to determine if the faculty of a large university system was equally distributed. Let us test this hypothesis at a significance level of 0.05.
    - **Solution:** We will follow the same six steps stated previously in units 9 and 10. Step 1: First, we must state the population parameter and the population of interest: Let $p_i =$ the true proportions of all faculty in this university system that are full professors, associate professors, assistant professors, instructors and adjunct/part time, so that $p_1 =$ true proportion of full professors, $p_2 =$ true proportion of associate professors, $p_3 =$ true proportion of assistant professors, $p_4 =$ true proportion of instructors, and $p_5 =$ true proportion of adjunct/part timers.
    - Step 2: Write the hypotheses for the test. Because the newsperson wants to see if the proportions are equally distributed among the 5 categories, each category should contain 20%.
      $H_0 : p_1 = 0.2, p_2 = 0.2, p_3 = 0.2, p_4 = 0.2, p_5 = 0.2 \quad H_a$: at least one of these proportions is incorrect.
    - Step 3: Choose the appropriate inference procedure and verify the conditions for using the procedure. We are doing a Chi-square Goodness of Fit Test. Although we do not know for sure how the sample was obtained, we shall assume selection procedure generated a simple random sample. We saw earlier in the table on page one, that the expected counts were all 30.8 which is greater than 5. All conditions are met.
    - Step 4: Carry out the inference procedure.

	Category					
	Full Professor	Associate Professor	Assistant Professor	Instructor	Adjunct/ Part time	Total
Frequency	22	31	25	35	41	154
Hypothesized Proportion	0.2	0.2	0.2	0.2	0.2	1
Expected Count	30.8	30.8	30.8	30.8	30.8	154

$$\chi^2 = \frac{(22-30.8)^2}{30.8} + \frac{(31-30.8)^2}{30.8} + \frac{(25-30.8)^2}{30.8} + \frac{(35-30.8)^2}{30.8} + \frac{(41-30.8)^2}{30.8}$$
$$= 2.514 + 0.001 + 1.092 + 0.573 + 3.378$$
$$= 7.56$$

The P-value is based on a Chi-square distribution with df = 5 – 1 = 4. We look up the computed value of 7.56 under the df = 4 column on the table. We see that 7.56 is smaller than 7.77, the lowest value of χ^2 in the table for df = 4, so that the P-value is greater than 0.100.

- Step 5: Make your decision. Our p-value is large and our significance level is 0.05. Since 0.100 > 0.05, our decision is to Fail to Reject H_0.
- Step 6: Interpret your results in the context of the problem. At a level of significance of 0.05, there is insufficient evidence to refute the claim that the proportion of faculty in each of the different categories is the same.

- **Example 2 –** Let's refer back to the M&M's example. According to the company, "..*the new mix of colors of M&M's plain chocolate candies will contain 30 percent browns, 20 percent yellows and reds, and 10 percent each of oranges, greens, and blues*." Let's test this by purchasing a plain bag of M&M's and counting the different colors. For this example, let's say we got 20 yellows, 30 red, 12 orange, 22 blue, 7 green and 15 brown. Is this sample consistent with the company's stated proportions? Test an appropriate hypothesis and state your conclusions.

  - **Solution:** Step 1: $p_i =$ the true proportion of colors of M&M's where $p_b =$ proportion of brown, $p_y =$ proportion of yellow, $p_r =$ proportion of red, $p_o =$ proportion of orange, $p_u =$ proportion of blue, and $p_g =$ proportion of green
  - Step 2: $H_0 : p_b = 0.30, p_y = p_r = 0.20, p_o = p_u = p_g = 0.10$ and $H_a :$ at least one of these is incorrect.
  - Step 3: We are doing a Chi-square Goodness of Fit Test. We shall assume selection procedure generated a simple random sample. There are a total of $15 + 20 + 30 + 12 + 22 + 7 = 106$ M&M's in the bag. To find the expected counts, we need to multiply each hypothesized proportion by 106. $E_b = 0.30(106) = 31.8, E_y = E_r = 0.20(106) = 21.2,$ $E_o = E_u = E_g = 0.10(106) = 10.6$. Since all expected counts are greater than 5, all assumptions are met and we can proceed with the test.
  - Step 4: Calculate the Chi-square test statistic.

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(15-31.8)^2}{31.8} + \frac{(20-21.2)^2}{21.2} + \frac{(30-21.2)^2}{21.2}$$

$$+ \frac{(12-10.6)^2}{10.6} + \frac{(22-10.6)^2}{10.6} + \frac{(7-10.6)^2}{10.6}$$

$$\chi^2 = 8.875 + 0.679 + 3.653 + 0.185 + 12.260 + 1.222$$

$$\chi^2 = 26.874$$

The P-value is based on a Chi-square distribution with df = 6 – 1 = 5. We look up the computed value of 26.874 under the df = 5 column on the table. We see that 26.874 is greater than 20.51, the highest value of χ^2 in the table for df = 5, so that the P-value is less than 0.001.

- Step 5: The problem does not state a significance level, so we will use $\alpha = 0.05$. Since our p-value is smaller than the significance level, we will Reject H_0 at the 0.05 level of significance.
- Step 6: At the 0.05 level of significance, there is sufficient evidence to reject M&M's claim and conclude that the population distribution IS different from the stated proportions in 1995. By looking at the components of the Chi-square test statistic, we can see that the greatest contributors to our large Chi-square value were 8.875 and 12.260. This tells us that the proportion of browns and the proportion of blues were very different from what was expected.

Syllabus Objectives: 1.23 – The student will examine two-way tables for association between categorical variables. 4.15 – The student will perform a Chi-square test for goodness of fit, homogeneity of proportions, and independence.

- **Two-Way Tables**
  - Data resulting from observations made on two different categorical variables can be summarized using a tabular format. A two-way table gives counts that show a relationship between two different categorical variables. Each of the counts occupies a **cell** in the table.
  - For example, consider the student data set giving information on 79 students. This dataset was obtained from a sample of 79 students taking elementary statistics. Below is the table containing this data.

	Contacts	Glasses	None
Female	**5**	**9**	**11**
Male	**5**	**22**	**27**

  - This is an example of a **two-way frequency table,** or **contingency table.** The numbers in the 6 cells are the observed cell counts. **Marginal totals** are obtained by adding the observed cell counts in each row and also in each column.

	Contacts	Glasses	None	**Row Marginal Total**
Female	**5**	**9**	**11**	**25**
Male	**5**	**22**	**27**	**54**
Column Marginal Total	**10**	**31**	**38**	**79**

- The sum of the column marginal total (or the row marginal totals) is called the **grand total**. The grand total in the table above is 79.

- Expected cell counts for a Two-Way Table
  - When the row indicates the population, the expected count for a cell is simply the overall proportion (over all populations) that have the category times the number in the population.
  - To illustrate:

	Contacts	Glasses	None	Row Marginal Total
Female	**5**	**9**	**11**	**25**
Male	**5**	**22**	**27**	**54**
Column Marginal Total	**10**	**31**	**38**	**79**

$\frac{10}{79}$ = overall proportion of students using contacts

54 = total number of male students

$\frac{10}{79} \bullet 54 = 6.83$ = expected number of males that use contacts as primary vision correction

  - The expected values for each cell represent what would be expected if there is no difference between the groups under study and can be found easily by using the following formula.
  - $$\text{Expected cell count} = \frac{(\text{Row total})(\text{Column total})}{\text{Grand total}}$$
  -

	Contacts	Glasses	None	Row Marginal Total
Female	**5** $\frac{25 \times 10}{79}$	**9** $\frac{25 \times 31}{79}$	**11** $\frac{25 \times 38}{79}$	**25**
Male	**5** $\frac{54 \times 10}{79}$	**22** $\frac{54 \times 31}{79}$	**27** $\frac{54 \times 38}{79}$	**54**
Column Marginal Total	**10**	**31**	**38**	**79**

The expected counts are in parentheses.

	Contacts	Glasses	None	Row Marginal Total
Female	**5 (3.16)**	**9 (9.81)**	**11 (12.03)**	**25**
Male	**5 (6.84)**	**22 (21.19)**	**27 (25.97)**	**54**
Column Marginal Total	**10**	**31**	**38**	**79**

- **Inference for Two-Way Tables**

  - There are two tests that can be used for two-way tables – Chi-square test for homogeneity of populations and Chi-square test of association/independence.
  - **Chi-square test for homogeneity of populations**
    - Typically, with a two-way table used to test homogeneity, the rows indicate different populations and the columns indicate different categories or vice versa.
    - For a test of homogeneity, the central question is whether the category proportions are the same for all of the proportions.
  - **Test steps:**
    1. Select independent SRS's from each of the *c* populations. Classify each individual in a sample according to a categorical response variable with *r* possible values. There are *c* different sets of proportions to be compared, one for each population.
    2. You can safely use the chi-square test with critical values from the chi-square distribution when all of the individual expected counts are 5 or greater.
    3. The null hypothesis is that the distribution of the response variable is the same in all *c* populations. The alternative hypothesis says that these *c* distributions are not all the same.
    4. If H_o is true, the chi-square statistic χ^2 has degrees of freedom (df) of $(r-1)(c-1)$.
  - **Chi-square test of association/independence**
    - The two-way table arises by classifying observations from a single population in two ways.
    - For a test of association/independence, the central question is whether the row and column variables are related to each other. The test assesses whether this observed association is statistically significant.
  - **Test steps:**
    1. Select a single SRS, with each individual classified according to both of two categorical variables.
    2. You can safely use the chi-square test with critical values from the chi-square distribution when all of the individual expected counts are 5 or greater.
    3. The null hypothesis is that the row and column variables are not related to each other. The alternative hypothesis says that the variables are related to each other or that there is a relationship. Language in the hypotheses can reference association or independence: Null: The two variables are independent or there is no association between the two variables. The null hypothesis always references no effect. Alternative: The two variables are not independent or dependent or there is an association between the two variables.
    4. The χ^2 test is based on the following degrees of freedom: df = $(r-1)(c-1)$, where *r* and *c* are the two categorical variables.

- **Chi-square test for Homogeneity example:**

  - **Example 1** – The following data come from a clinical trial of a drug regime used in treating a type of cancer, lymphocytic lymphomas. Patients (273) were randomly divided into two groups, with one group of patients receiving cytoxan plus prednisone (CP) and the other receiving a BCNU plus prednisone (BP). The responses to treatment were graded on a qualitative scale. The two-way table summary of the results is below. Setup up and perform an appropriate hypothesis test at the 0.05 level of significance.

Treatment \ Response	Complete Response	Partial Response	No Change	Progression	Row Marginal Total
BP	**26**	**51**	**21**	**40**	**138**
CP	**31**	**59**	**11**	**34**	**135**
Column Marginal Total	**57**	**110**	**32**	**74**	**273**

    - **Solution:** We will follow the same steps we did with the Goodness of Fit Chi-square test except that step 1 is no longer needed. We proceed to step 2.
    - Step 2: Write the hypotheses for the test. H_0 : The true response to treatment proportions are the same for both treatments. H_a : The true response to treatment proportions are not all the same for both treatments.
    - Step 3: Choose the appropriate inference procedure and verify the conditions for using the procedure. We are doing a Chi-square test for the homogeneity of proportions. We assume that the samples were chosen randomly and independently. We must also check that all expected cell counts are at least 5. The smallest expected count is 15.82, so all assumptions are met and we can carry out the test.

Treatment \ Response	Complete Response	Partial Response	No Change	Progression	Row Marginal Total
Female	**26 (28.81)**	**51 (55.60)**	**21 (16.18)**	**40 (37.41)**	**138**
Male	**31 (28.19)**	**59 (54.40)**	**11 (15.82)**	**34 (36.59)**	**135**
Column Marginal Total	**57**	**110**	**32**	**74**	**273**

- Step 4: Carry out the inference procedure.

$$\chi^2 = \frac{(26-28.81)^2}{28.81} + \frac{(51-55.60)^2}{55.60} + \frac{(21-16.18)^2}{16.18}$$
$$+\frac{(40-37.41)^2}{37.41} + \frac{(31-28.19)^2}{28.19}$$
$$+\frac{(59-54.40)^2}{54.40} + \frac{(11-15.82)^2}{15.82} + \frac{(34-36.59)^2}{36.59}$$

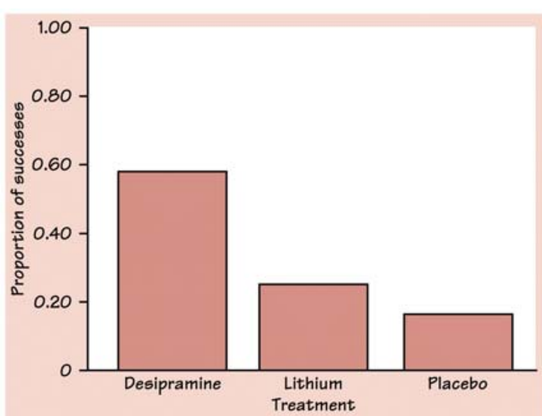
$$= 0.275+0.381+1.439+0.180+0.281+0.390+1.471+0.184 = 4.60$$

The two-way table for this example has 2 rows and 4 columns, so the appropriate df is (2-1)(4-1) = 3. We look up the computed value of 4.60 under the df = 3 column on the table. We see that 4.60 < 6.25, the lowest value of χ^2 in the table for df = 3, so that the P-value is greater than 0.100.

- Step 5: Make your decision. Our p-value is large and our significance level is 0.05. Since 0.100 > 0.05, our decision is to Fail to Reject H_0.
- Step 6: Interpret your results in the context of the problem. At a level of significance of 0.05, there is insufficient evidence to conclude that the responses are different for the two treatments.

- **Example 2** – Chronic users of cocaine need the drug to feel pleasure. Perhaps giving them a medication that fights depression will help them stay off cocaine. A three-year study compared an antidepressant called desipramine with lithium and a placebo. The subjects were 72 chronic users of cocaine who wanted to break their drug habit. Twenty-four of the subjects were randomly assigned to each treatment. Below are the counts and proportions of the subjects who avoided relapse into cocaine use during the study along with a bar graph.

Group	Treatment	Subjects	No Relapse	**Proportion**
1	Desipramine	24	14	**0.583**
2	Lithium	24	6	**0.250**
3	Placebo	24	4	**0.167**



The sample proportions of subjects who stayed off cocaine are quite different, but this is not enough evidence that the proportions of successes for the three treatments differ in the population of all cocaine users. To answer that question, we must perform the Chi-square test. Setup up and perform an appropriate hypothesis test at the 0.05 level of significance.

- **Solution:** We will follow the same steps we did in Example 1.
- Step 2: Write the hypotheses for the test. H_0 : The true proportions of cocaine addicts who avoid relapse are the same for the three treatments. H_a : The true proportions are not the same.
- Step 3: Choose the appropriate inference procedure and verify the conditions for using the procedure. We are doing a Chi-square test for the homogeneity of proportions. We assume that the samples were chosen randomly and independently. We must also check that all expected cell counts are at least 5. Below is the contingency table and expected values. The expected values are in parentheses. The smallest expected count is 8.0, so all assumptions are met and we can carry out the test.

Treatment	No Relapse	Relapse	Total
1	14 (8)	10 (16)	24
2	6 (8)	18 (8)	24
3	4 (8)	20 (16)	24
Total	24	48	72

- Step 4: Carry out the inference procedure. Below is a minitab output for this example.

```
Expected counts are printed below observed counts
              C1          C2        Total
    1         14          10           24
            8.00       16.00

    2          6          18           24
            8.00       16.00

    3          4          20           24
            8.00       16.00

Total         24          48           72

ChiSq =    4.500  + 2.250 +
           0.500  + 0.250 +
           2.000  + 1.000  = 10.500
df = 2

   Chisquare 2.
   10.5000     0.9948
```

The two-way table for this example has 3 rows and 2 columns, so the appropriate df is (3-1)(2-1) = 2. Note the χ^2 value in the printout is 10.500 and the last line gives 0.9948 as the probability of a value *less than* 10.500 if the null hypothesis is true. The P-value is therefore $1 - 0.9948 = 0.0052$.

- Step 5: Make your decision. Since our p-value is small and .0052 < 0.05, our decision is to Reject H_0.
- Step 6: Interpret your results in the context of the problem. At a level of significance of 0.05, there is sufficient evidence to conclude that there are differences among the three proportions.

- **Chi-square test of association/independence example:**
  - **Example 1** – Consider the two categorical variables, gender and principle form of vision correction for the sample of students presented earlier on page seven. We shall now test to see if the gender and the principal form of vision correction are independent. Setup up and perform an appropriate hypothesis test at the 0.05 level of significance.

	Contacts	Glasses	None	Row Marginal Total
Female	**5**	**9**	**11**	**25**
Male	**5**	**22**	**27**	**54**
Column Marginal Total	**10**	**31**	**38**	**79**

- **Solution:** We follow the same steps that we used for the Chi-square test for homogeneity.
- Step 2: Write the hypotheses for the test. H_0: Gender and principle method of vision correction are independent. H_a: Gender and principle method of vision correction are not independent.
- Step 3: Choose the appropriate inference procedure and verify the conditions for using the procedure. We are doing a Chi-square test of independence. We assume that the sample was chosen randomly. We must also check that all expected cell counts are at least 5. The smallest expected count is 3.16, making one of the expected counts less than 5. We will proceed with caution as we carry out the test.

	Contacts	Glasses	None	Row Marginal Total
Female	**5 (3.16)**	**9 (9.81)**	**11 (12.03)**	**25**
Male	**5 (6.84)**	**22 (21.19)**	**27 (25.97)**	**54**
Column Marginal Total	**10**	**31**	**38**	**79**

- Step 4: Carry out the inference procedure.

$$\chi^2 = \frac{(5-3.16)^2}{3.16} + \frac{(9-9.81)^2}{9.81} + \frac{(11-12.03)^2}{12.03} + \frac{(5-6.84)^2}{6.84} + \frac{(22-21.19)^2}{21.19} + \frac{(27-25.97)^2}{25.97}$$

$$\chi^2 = 1.071 + 0.067 + 0.088 + 0.495 + 0.031 + 0.041$$

$$\chi^2 = 1.793$$

The two-way table for this example has 2 rows and 3 columns, so the appropriate df is (2-1)(3-1) = 2. We look up the computed value of 1.795 under the df = 2 column on the table. We see that 1.795 < 4.60, the lowest value of χ^2 in the table for df = 2, so that the P-value is greater than 0.100.

- Step 5: Make your decision. Our p-value is large and our significance level is 0.05. Since 0.100 > 0.05, our decision is to Fail to Reject H_0.
- Step 6: Interpret your results in the context of the problem. At a level of significance of 0.05, there is not sufficient evidence to conclude that the gender and vision correction are related. For all practical purposes, one would find it reasonable to assume that gender and need for vision correction are independent.

o **Example 2** – A rural county hospital offers several health services. The hospital administrators conducted a poll to determine whether the residents' satisfaction with the available services depends on their gender. A random sample of 1,000 adult county residents was selected. The gender of each respondent was recorded and each was asked whether he or she was satisfied with the services offered by the hospital. The resulting data are shown in the table below. Using a significance level of 0.05, conduct an appropriate test to determine if, for adult residents of this county, there is an association between gender and whether or not they were satisfied with services offered by the hospital.

	Male	Female	Total
Satisfied	384	416	800
Not Satisfied	80	120	200
Total	464	536	1,000

- **Solution:** We follow the same steps that we used for example 1.
- Step 2: Write the hypotheses for the test. H_0: Gender and satisfaction with health services offered by the hospital are independent; OR There is no association between gender and satisfaction. H_a: Gender and satisfaction with health services offered by the hospital are dependent; OR There is an association between gender and satisfaction.

- Step 3: Choose the appropriate inference procedure and verify the conditions for using the procedure. We are doing a Chi-square test of independence. A random sample has been taken. We must also check that all expected cell counts are at least 5. The smallest expected count is 92.8, which is greater than 5. We can perform the chi-square test.

$$\text{Expected counts: } E_{(1,1)} = \frac{800 \times 464}{1000}, \quad E_{(1,2)} = \frac{800 \times 536}{1000},$$

$$E_{(2,1)} = \frac{200 \times 464}{1000}, \quad E_{(2,2)} = \frac{200 \times 536}{1000}$$

$$\text{Expected: } E_{(1,1)} = 371.20, \; E_{(1,2)} = 428.80, \; E_{(2,1)} = 92.80, \; E_{(2,2)} = 107.20$$

- Step 4: Carry out the inference procedure.

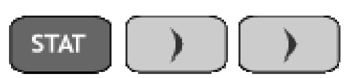
$$\chi^2 = \frac{(384 - 371.20)^2}{371.20} + \frac{(416 - 428.80)^2}{428.80} + \frac{(80 - 92.80)^2}{92.80} + \frac{(120 - 107.20)^2}{107.20}$$

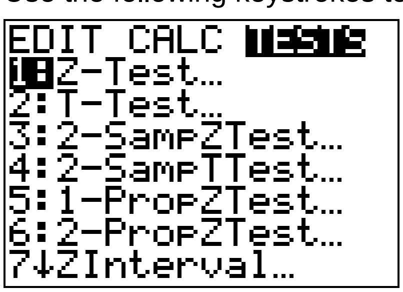
$$\chi^2 = 0.441 + 0.382 + 1.766 + 1.528$$

$$\chi^2 = 4.117$$

The two-way table for this example has 2 rows and 2 columns, so the appropriate df is (2-1)(2-1) = 1. We look up the computed value of 4.117 under the df = 1 column on the table. We see that 4.117 is between 4.01 and 4.21, making our P-value between 0.40 and 0.45.

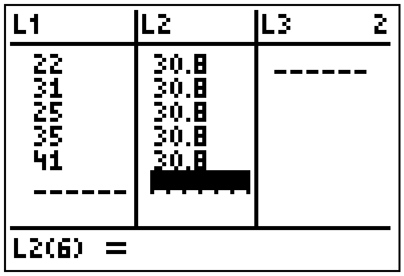
- Step 5: Make your decision. Because our p-value is 0.40<p<0.45, which is less than 0.05, we can Reject H_0.
- Step 6: Interpret your results in the context of the problem. At a level of significance of 0.05, there is evidence of an association between gender and satisfaction with health services offered by the hospital for adult residents of this county.

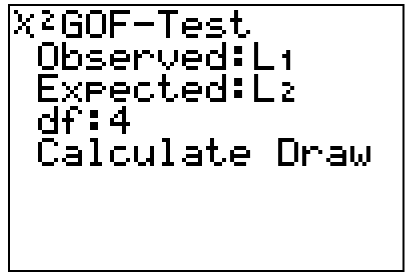
- **Using the TI-84:**
  - Chi-square Goodness of Fit Tests, Homogeneity, and Association/Independence Tests can be done on the TI-84 calculator.
  - Note that the calculator will only help on the calculation step. All other steps for the tests must still be done.
  - Use the following keystrokes to get to the test menu: 

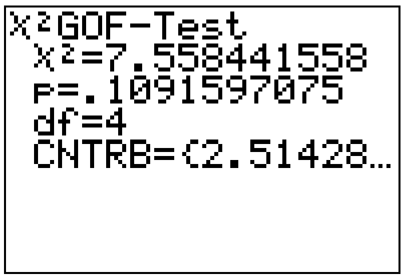


- Select the inference procedure you wish to perform and input the necessary information. Test C is used for the Homogeneity and Association Tests (two-way table) and Test D is used only for the Goodness of Fit Test.

- **Example 1 – Revisit Goodness of Fit Example 1:** You must input the observed values in List 1 and input the expected values in List 2. Then go to the test menu, input the degrees of freedom, place the cursor on "Calculate" and hit "enter".

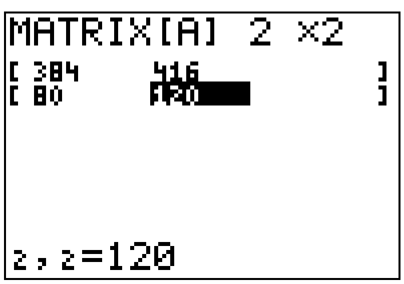


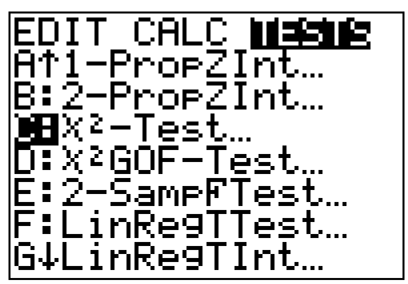


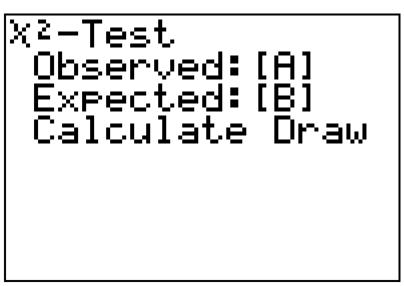


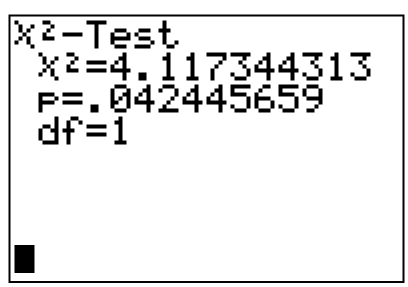
We receive the Chi-square value of 7.56 and a P-value of 0.109. CNTRB gives us the contributing factors and we can see how each category contributed to the Chi-square statistic.

- **Example 2 – Revisit the Association/Independence Example 2:** This test does not use lists, but instead uses matrices. You must first put in the observed counts into Matrix A. That is the only information that needs to be inputted. The TI-84 will then calculate the expected values and store them in Matrix B for you after the test is run.









The Chi-square test statistic is 4.117, df = 1, and the P-value is 0 .042.