Example 8.3 Goodness of Fit Test

A bag of Skittles was bought (see **Figure 8.4**) and the five colors counted. Before any count, it was theorized that the five colors would be in proportion of 17% red, 23% green, 25% brown, 20% orange, and 15% yellow. **Determine whether the proportion chosen fits at the .05 level of confidence**. A total of 386 skittles were counted as 56 red, 91 green, 91 brown, 87 orange, and 61 yellow.

Solution:

Step 1. **State the hypotheses**.

H_0: The proportion of Skittles will be 17% red, 23% green, 25% brown, 20% orange, and 15% yellow.

H_A: The proportion of Skittles will *not* be 17% red, 23% green, 25% brown, 20% orange, and 15% yellow.

Step 2. **Find the critical value**. Degrees of freedom (df) (k – 1) is 5 -1 or 4. At alpha .05, $\chi^2_{cv} = \mathbf{9.488}$.

Step 3. **Compute the test value**.

	Red	Green	Brown	Orange	Yellow
Expected	0.17•386 = 65.6	0.23•386 = 88.8	0.25•386 = 96.5	0.20•386 = 77.2	0.15•386 = 57.9
Observed	56	91	91	87	61

$$\chi^2_{obs} = \frac{(56-65.6)^2}{65.6} + \frac{(91-88.8)^2}{88.8} + \frac{(91-96.5)^2}{96.5} + \frac{(87-77.2)^2}{77.2} + \frac{(61-57.9)^2}{57.9}$$

$$= \frac{92.16}{65.6} + \frac{4.84}{88.8} + \frac{30.25}{96.5} + \frac{96.04}{77.2} + \frac{9.61}{57.9}$$

$$= 1.4 + 0.055 + 0.3 + 1.24 + 0.17 = \mathbf{2.865} = \chi^2_{obs}$$

Step 4. **Make the decision**. **2.865 < 9.488**

Accept the null hypothesis; the proportions are as stated with 95% confidence.

Summary

The Chi Square tests involve using frequency counts involving one or more variables. The formula is basically the same for all the tests but a Yates Continuity of Correction may be used when the degrees of freedom (df) is one when two or more variables are tested. When there is only one row, degrees of freedom are (c -1), or the number of columns minus one.

Three Chi Square tests are presented: (a) Test of Independence, (b) Test of Equality, and (c) Goodness of Fit. The last two tests use one row and expected values are either determined by an equal proportion or theorized percentage while the first is a test of the independence of two or more variables. It is always a right-tailed test. Expected values are determined by row total times column total divided by grand total for a Test of Independence. A Chi-Square Distribution table is used for determining critical values. See **Figure 8.5** for Chapter Summary Mind Map.