A Chi-Square Investigation With Tums Antacid

By Bonnie H. Litwiller and David R. Duncan

In teaching statistical processes, it is important that there be applications to real world settings and activities. When this is done, students are more likely to see the significance of the steps being used in the statistics and the practical application of the results.

One such activity involves the use of the Chi-Square statistical test and its applications to counting the different colored tablets in a bottle of "Tums" antacids.

Most people know that *Assorted Fruit Extra Strength* Tums come in four different colors: green, orange, pink, and yellow. Let us conjecture that there are, in the long run, equal numbers of Tums for each of the four colors per container. Is this conjecture true? We shall test this conjecture, called the **null hypothesis**, with a randomly selected container of Tums. Results of counting the different colors in our container are:

Green	23
Orange	22
Pink	35
Yellow	17
Total	97

If our null hypothesis is correct, how many of each color should be present? Since

$$\frac{97 \text{ (number of Tums per bottle)}}{4 \text{ (number of different colors)}} = 24.25,$$

we would expect approximately 24.25 of each color in our container.

To test the null hypothesis we shall use the Chi-Square statistic. Let us construct Table I with column entries as follows:

O	=	the observed frequencies, the numbers of each color of Tums actually present in our container
E	=	the expected frequencies (assuming that the null hypothesis is true); in our case, these are all 24.25
$\frac{(\mathbf{O}-\mathbf{E})^2}{\mathbf{E}}$	=	a measure of the discrepancy between **O** and **E**

Table I

Color	**O**	**E**	$\frac{(\mathbf{O}-\mathbf{E})^2}{\mathbf{E}}$
Green	23	24.25	0.064
Orange	22	24.25	0.209
Pink	35	24.25	4.765
Yellow	17	24.25	2.168
	97		7.206

In the last column (a measure of discrepancy), a small number indicates that **O** and **E** are relatively close together, as is the case for the green tablets. This means that we had about as many greens as we would have expected to have. A large number indicates that **O** and **E** are relatively far apart, as is the case for the pink tablets. This means that there was quite a disparity between the number of pink tablets that we expected to have, and the the number of pink tablets that were actually in the bottle.

The sum of this discrepancy column, 7.206, is called the *Computed Chi-Square Statistic*; abbreviate it as **CCSS**. A determination must be made as to whether the **CCSS** is large enough to cause us to reject the null hypothesis. To make this decision a "referee" is needed. This "referee" is found in the Table of the Chi-Square Statistic; abbreviate it as **TCSS**.

To read this table, the degrees of freedom must first be determined; that is, the number of categories (colors) minus one. In our case the degrees of freedom is $4 - 1 = 3$. The interpretation of this is that if the total number of Tums is known, and the number in each of three categories is known, then the number in the fourth category can be calculated.

Roughly speaking, the **significance level** is the probability of rejecting a null hypothesis, when the null hypothesis is,in fact, true. This could occur because the sample is not representative of the population. From a Chi-Square table, we find:

Significance Level	**TCSS**
10%	6.251
5%	7.815
1%	11.344

The decision mechanism for the null hypothesis is:

- If **CCSS** > **TCSS**, then **CCSS** is large in the "judgment of the referee." If this is true, reject the null hypothesis. Thus, we may not expect to find equal numbers of colored tablets in each bottle Tums.
- If **CCSS** < **TCSS**, then **CCSS** is small in the "judgment of the referee." If this is true, accept the null hypothesis. Thus, we may expect to find equal numbers of colored tablets in each bottle Tums.

For our container of Tums, our **CCSS** of 7.206 is greater than 6.251 but less than either 7.815 or 11.344. We consequently reject our null hypothesis at the 10% significance level, but accept it at the 5% and 1% significance levels.

At the 10% level we reject the assumption that the packaging process generally places equal numbers of Tums of each color in each container. At the more demanding 5% and 1% significance levels we do not have enough evidence to reject the assumption.

Performing further experiments, we replicated this activity for four additional containers of Tums. **Table II** reports the counts and CCSS calculations for all five containers.

Table II

Container	Green	Orange	Pink	Yellow	Total	**CCSS**
1	23	22	35	17	97	7.206
2	21	29	24	22	96	1.584
3	24	31	24	18	97	3.496
4	27	26	27	18	98	2.326
5	22	27	26	22	97	0.856

For each of the containers 2, 3, 4, and 5, the null hypothesis was accepted at each significance level. That is, there was insufficient evidence to reject the assumption of generally equal numbers of colors in each container.

For one final application, suppose the five containers were combined to form a giant supply of Tums. **Table III** reports the Chi-Square results for the situation:

Table III

Color	**O**	**E**	$\frac{(\mathbf{O}-\mathbf{E})^2}{\mathbf{E}}$
Green	117	121.25	0.149
Orange	135	121.25	1.559
Pink	136	121.25	1.794
Yellow	97	121.25	4.850
	485		8.352

The null hypothesis is rejected at both the 10% and 5% significance levels. Only at the much stricter 1% level is there insufficient evidence to reject the null hypothesis.

Challenges to the Reader and His/Her Students:

(1) Form a new null hypothesis based on the analysis of the five combined containers. Test this hypothesis with additional supplies of Tums.
(2) Redo this experience with more containers of Tums.
(3) Use this same process with M & M's or jelly beans.
(4) Find other situations in which items can be placed into categories and counted, and hypotheses can be formulated and tested.

University of Northern Iowa
Cedar Falls, Iowa 50614-0506