A Chocolate Candy Color Distribution: An Enumerative Statistical Experiment

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 meaning of the steps being developed.

One such activity involves using the Chi-Square statistical test and its applications to counting M&M's of different colors. All students are aware that M&M chocolate candies come in different colors. For instance, a package of the plain M&M candy (nonholiday) contains a mixture of six colors: brown, blue, green, orange, red and yellow.

According to the information provided by Mars Incorporated, the manufacturer of M&M's, the following should be the color distribution for the plain chocolate M&M's:

Brown	30%
Yellow	20%
Red	20%
Orange	10%
Green	10%
Blue	10%

We shall test this distribution hypothesis, called the null hypothesis, with a randomly selected set of plain M&M's.

Experiment 1

We combined the contents of seven 1.69 oz packages. Results of counting the different colors in our sample are as follows:

Color	Number
Brown	104
Yellow	73
Red	92
Orange	38
Green	32
Blue	43
TOTAL	382

The expected distribution:

Color	Expected Number
Brown	30% of 382 = 114.6
Yellow	20% of 382 = 76.4
Red	20% of 382 = 76.4
Orange	10% of 382 = 38.2
Green	10% of 382 = 38.2
Blue	10% of 382 = 38.2

To test the null hypothesis, we shall use the Chi-Square statistic. Let us construct Table 1 with column entries as follows:

O = The observed frequencies, the numbers of each color of M&M's actually present in our package.

E = The expected frequencies (if the null hypothesis were true).

$(O - E)^2/E$ = A measure of the discrepancy between O and E.

Table I

Color	O	E	$(O - E)^2/E$
Brown	104	114.6	0.98
Yellow	73	76.4	0.15
Red	92	76.4	3.19
Orange	38	38.2	0.00
Green	32	38.2	1.01
Blue	43	38.2	0.60
TOTAL	382	382	5.93

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 yellow. A larger number indicates that O and E are relatively far apart, as is the case for red.

The sum of this discrepancy column, 5.93, is called the computed Chi-Square Statistic (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 Chi-Square Statistic (TCSS).

To read a Chi-Square table, the degrees of freedom must first be determined; that is, the number of categories (colors) −1. In our case, the degrees of freedom is 6 − 1 = 5. This means that if the total number of M&M's were known, and the number in each of five categories were known, the number in the sixth category could be calculated.

The significance level is the probability of rejecting a null hypothesis which 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%	9.236
5%	11.070
1%	15.085

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.
- If CCSS < TCSS, then CCSS is small in the "judgment of the referee." If this is true, *accept* the null hypothesis.

For Experiment 1, our CCSS of 5.93 is less than any of the TCSS values; for each level of significance, we do *not* reject the null hypothesis. In other words, we retain the assumption that the packaging process includes the percent of M&M's of each color as claimed by the manufacturer.

There are many different types of M&M's besides the plain chocolate, nonholiday variety used in Experiment 1. Distribution information from Mars Incorporated predicts the following percents for other types of M&M's.

Nonholiday

Peanut Butter or Almond	Peanut
Brown 20%	Brown 20%
Yellow 20%	Yellow 20%
Red 20%	Red 20%
Green 20%	Orange 20%
Blue 20%	Green 10%
	Blue 10%

Holiday

Easter (Plain Chocolate, Peanut or Almond)

Yellow	25%
Blue	25%
Green	25%
Pink	25%

Experiment 2: Peanut Butter (Nonholiday)

We used seven 1.63 oz packages for our sample set.

Color	Predicted %	*O*	*E*	$(O-E)^2/E$
Brown	20	48	37.2	3.14
Yellow	20	30	37.2	1.39
Red	20	58	37.2	11.63
Green	20	32	37.2	0.73
Blue	20	18	37.2	9.91
TOTAL		186	186	26.80

For 4 degrees of freedom (5 − 1), a Chi-Square table yielded the following values:

Significant Level	TCCS
10%	7.779
5%	9.488
1%	13.277

Since 26.80 is greater than any of the above TCCS statistics, the null hypothesis is rejected for all significance levels. In other words, we *reject* the assumption that the packaging process places equal numbers of M&M's of each color in our set.

Experiment 3: Almonds (Nonholiday)

We used seven 1.31 oz packages for our sample set.

Color	Predicted %	*O*	*E*	$(O-E)^2/E$
Brown	20	31	20.4	5.51
Yellow	20	21	20.4	0.02
Red	20	16	20.4	0.95
Green	20	19	20.4	0.10
Blue	20	15	20.4	1.43
TOTAL		102	102	8.01

Using the same degrees of freedom and Chi-Square table as in Experiment 2 we

- reject the null hypothesis at the 10 percent significance level, since 8.01 > 7.779, but

- do not reject the null hypothesis at the 5 percent and 1 percent significance level, since 8.01 < 9.488 and 8.01 < 13.277.

There is enough evidence to cause doubts that the distribution percents are correctly described, but not enough evidence to conclusively prove it. In a legal setting, this is similar to having enough evidence to indict but not convict.

Experiment 4: Peanut (Nonholiday)

We used seven 1.74 oz packages for our sample set.

Color	Predicted %	O	E	$(O-E)^2/E$
Brown	20	25	31.4	1.30
Yellow	20	41	31.4	2.94
Red	20	18	31.4	5.72
Orange	20	20	31.4	4.14
Green	10	26	15.7	6.76
Blue	10	27	15.7	8.13
TOTAL		157	157	28.99

For 5 degrees of freedom, the Chi-Square table entries are the same as Experiment 1. We reject all significance levels since 28.99 is greater than any of the TCSS. The predicted percentages are not confirmed in our sample set.

For Experiments 5, 6 and 7, the Easter pastel colors are used. For each Easter type, the predicted color distributions are 25 percent for each of the colors yellow, blue, green and pink.

For Experiments 5–7, the TCSS values are:

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

Experiment 5: Easter (Plain Chocolate)

We used one 16 oz package.

Color	Predicted %	O	E	$(O-E)^2/E$
Yellow	25	122	128.5	0.33
Blue	25	117	128.5	1.03
Green	25	142	128.5	1.42
Pink	25	133	128.5	0.16
TOTAL		514	514	2.94

We do not reject the null hypothesis at any significance level.

Experiment 6: Easter (Peanuts)

We used one 16 oz package.

Color	Predicted %	O	E	$(O-E)^2/E$
Yellow	25	31	49	6.61
Blue	25	50	49	0.02
Green	25	37	49	2.94
Pink	25	78	49	17.16
TOTAL		196	196	26.73

A resounding rejection of the null hypothesis at all the significance levels is in order!

Experiment 7: Easter (Almonds)

We used one 12 oz package.

Color	Predicted %	O	E	$(O-E)^2/E$
Yellow	25	29	27.5	0.08
Blue	25	23	27.5	0.74
Green	25	16	27.5	4.81
Pink	25	42	27.5	7.65
TOTAL		110	110	13.28

Rejection of the null hypothesis at all levels is again in order.

A variety of conclusions resulted in the different experiments. Sometimes the results were consistent with the distribution predictions, leading to nonrejection of the null hypothesis; sometimes the results were inconsistent with the distribution predictions, leading to rejection of the null hypothesis. On other occasions the results were mixed—"inconsistent enough" to yield null hypothesis rejections at some significance levels but not at others.

Challenges for Readers and Their Students

1. Redo experiments with larger numbers of M&M's. How do your results compare with ours?
2. Investigate Christmas (red and green) and Valentine (red, pink and white) M&M distributions.
3. Find other candies for which predicted color distributions are known and replicate our process.
4. Find other real-world enumerative data for which the Chi-Square Method can be used.