

Practicum 5: Testing Mars Chocolate  
January 1, 2017  
Ole J. Forsberg

The question we are to answer is this: Is the stated distribution of Milk Chocolate M&Ms reasonable? According to research, the distribution of Milk Chocolate M&Ms by color is

<b>Color</b>	<b>Blue</b>	<b>Orange</b>	<b>Green</b>	<b>Yellow</b>	<b>Red</b>	<b>Brown</b>
<b>Proportion</b>	0.24	0.20	0.16	0.14	0.13	0.13

To test if this distribution is reasonable, I will collect data. I purchased a 20-ounce bag, which contained 249 Milk Chocolate M&Ms. This size was a trade-off between power (more M&Ms) and cost (fewer M&Ms).

The table below gives the observed frequency distribution of M&M colors in that bag. It also provides the expected counts, given that the provided distribution is correct.

<b>Color</b>	<b>Blue</b>	<b>Orange</b>	<b>Green</b>	<b>Yellow</b>	<b>Red</b>	<b>Brown</b>
<b>Observed</b>	60	54	46	33	23	33
<b>Expected</b>	59.76	49.80	38.84	34.86	32.37	32.37

As we are testing a categorical distribution, we should use the chi-square goodness-of-fit test. There are  $6 - 1 = 5$  degrees of freedom. The Chi-Square table gives a critical value of 11.07.

To calculate the test statistic, I used the formula

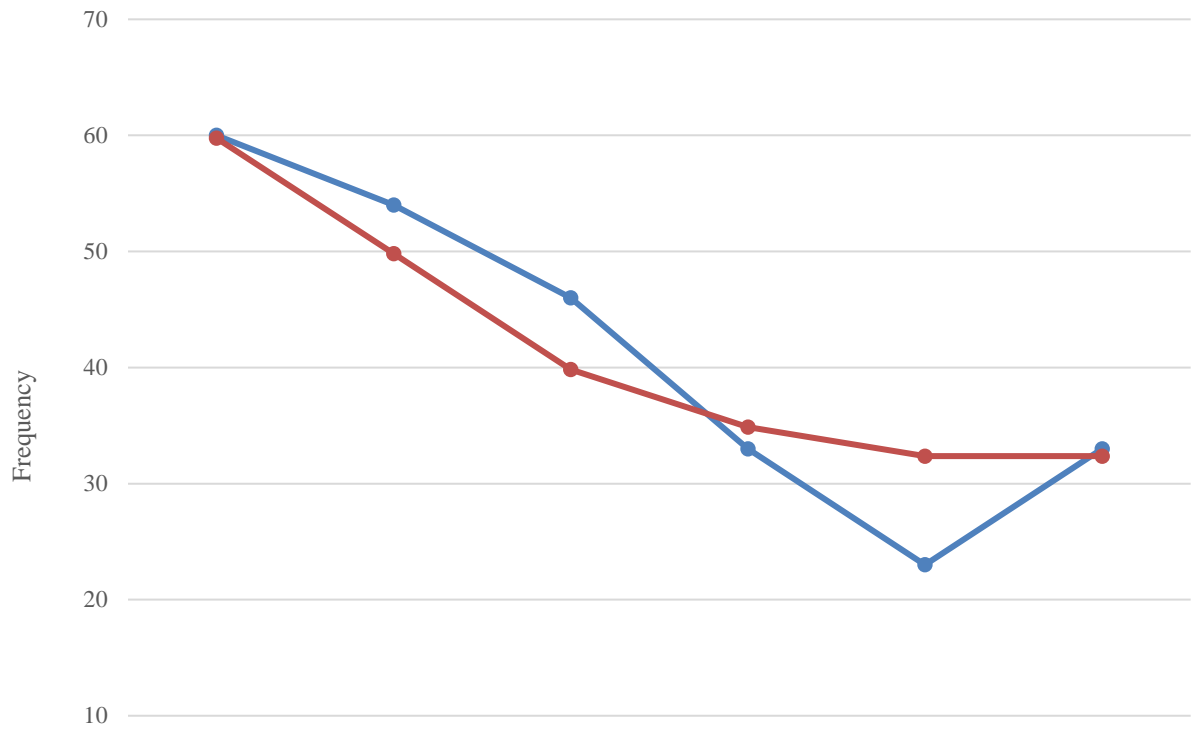
$$X^2 = \sum_{i=1}^6 \frac{(\text{Obs} - \text{Exp})^2}{\text{Exp}}$$

According to that formula, the value of the test statistic is 4.131. Since this value is less than the critical value, we cannot reject the null hypothesis. There is no significant evidence that the stated distribution is not the true color distribution of Milk Chocolate M&Ms.

The p-value of 0.5307 agrees with this conclusion. Since the p-value is greater than our alpha level of 0.05, we fail to reject the null hypothesis. Again, we do not have the necessary evidence to conclude that the posted distribution is wrong

I was unable to reject the null hypothesis. This is due either to the posted distribution being correct or the sample size is too small to detect the difference. Were I to do this experiment again, I would purchase more M&Ms. Doing so will increase the ability to determine if the stated distribution is correct.

An illustration of the divergence between the observed and the expected is given in the line graphic on the next page.



	Blue	Orange	Green	Yellow	Red	Brown
Obs	60	54	46	33	23	33
Exp	59.76	49.8	39.84	34.86	32.37	32.37