Project #1
The Great M&M Count

By
Scott Hogan

For
ENGR 315
Dr. Jennifer Turns

Report Due: April 7, 1999.

Problem statement

This project was designed to become familiar with some terms and use of descriptive statistics. Given about 100 bags of M&M's, the task is to determine how much variation occurred between each bag in terms of total number of M&M's, numbers of each color of M&M and number of broken M&M's. In addition to quantifying the variation between bags, a sample bag was compared to the population to see how closely it matched the characteristics of the population.

Procedure

In class, each person was given a bag of M&M's presumably purchased together and from the same lot. Each person was to count the number of each of the colors of candies, the total number of candies, and the number of defective candies of each color. This data was then tabulated for analysis. In order to determine how much one bag of candy varied from the next, the mean and standard deviation were calculated for the number of each color in a bag. The individual bag was then compared to the population.

The Collected Class Sample – Aggregate Data Analysis

Figure 1 gives the mean and standard deviation for the number of each color of candy in a package as well as the total number of candies. From the standard deviation of the total number of candies, it is apparent that there is a fair amount of variance in the total number of candies. The "Total" column translates as follows: each bag has a 68%

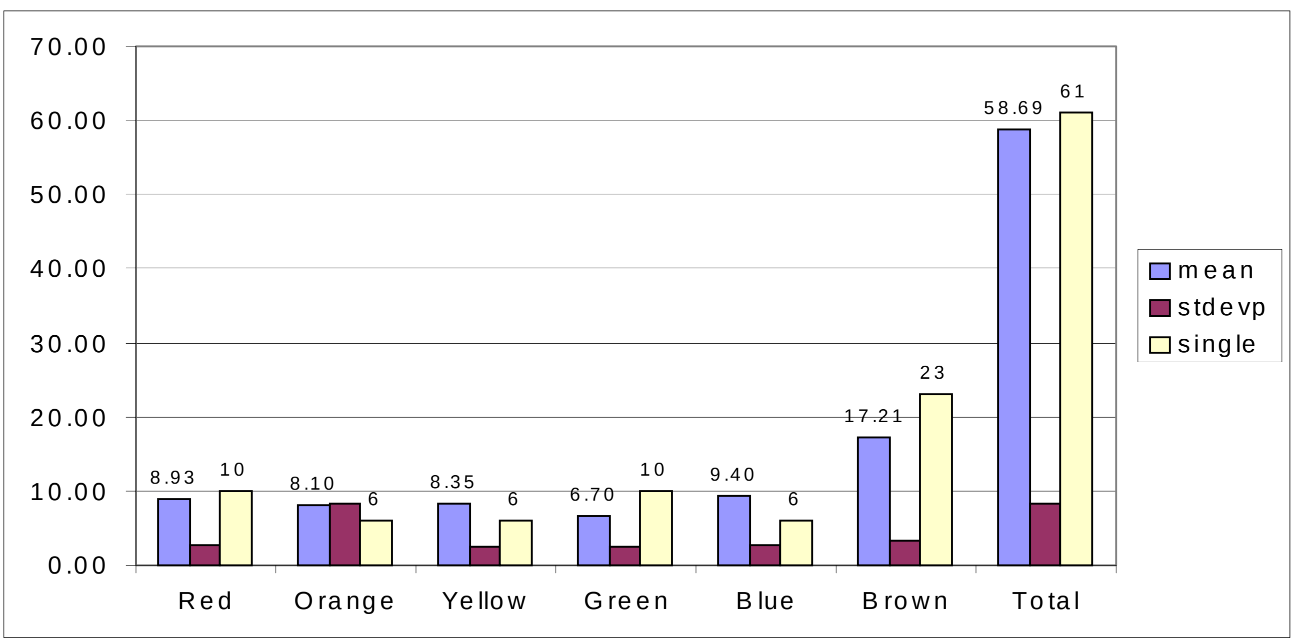

Figure 1. Class Data.

likely-hood of having between 50 and 67 candies where 50 is the mean minus the standard deviation and 67 is the mean plus the standard deviation. This brings up an interesting point. Do the candies vary in weight significantly? Most packages are filled very accurately because they cannot have less than the amount of product sold in the package. In addition, they do not typically waste a lot of product in the form of over filling because it costs money.

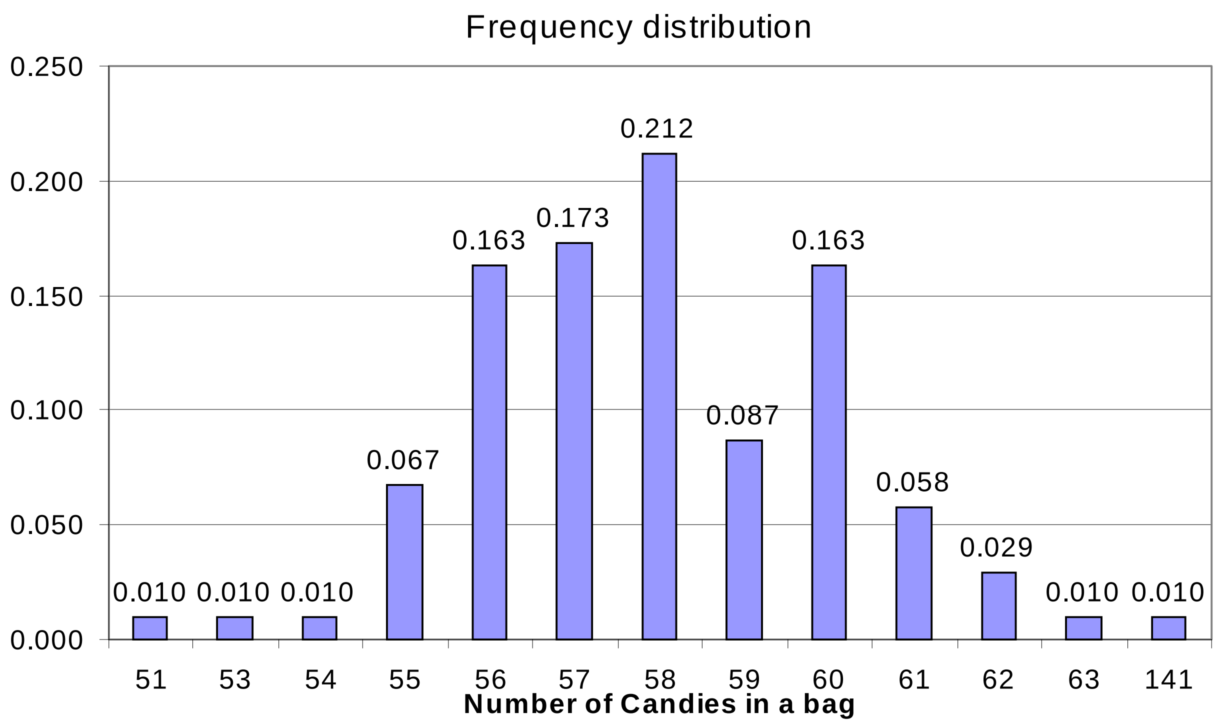

Figure 2. Frequency Distribution

The Single Sample – Comments on Normal-ness

The sample bag was fairly unique in its color distribution. Referring again to figure 1, the total number of candies matched the mean well within the standard deviation. However, the number of Blue candies in the sample package is outside the norm. Most of the packages had between 6.7 and 12.1 blue candies per package while the sample only had 6 (based on the mean +/- standard deviation). This sample was fairly unique also because it had a greater than the mean plus the standard deviation of brown candies.

Summary and Conclusions

There is a surprising variance in the total number of candies in a bag. As stated above, a company may not put less than the reported quantity (usually a weight) of a product in a package. In a case like this where there is a large variance, they must over fill the average package so that a large percentage of the packages have a minimum quantity of product. The variance of colors in each bag is not so surprising given the small quantity of each color. My sample was somewhat of an 'odd-ball' having the number of brown and blue candies outside of the standard deviation from the mean.

What I learned

I design food fillers for a living and this project raised an interesting question. Frequently when filling dry product like this, a filler will quickly put a certain quantity of product in the bag. This is called the gross fill. The bag is usually suspended on an electronic scale that tells how full the bag is. The information from the scale is used to top off the bag with what they call the fine fill in order to meet a close tolerance on the fill amount.

My question was, if they fill M&M bags this way, do they use a single color for the fine fill? To answer this I looked to see if any of the colors had a significant variation. Orange did, but not tremendously so. So I did another calculation. For each color, I took the deviation from the mean of the total number of candies in a bag and subtracted that number from the deviation from the mean of that color. Equation 1 describes this.

$$N = T - \bar{T} - (C - \bar{C}) \qquad (1)$$

T it the total count, and C is the color count. The idea being, if a bag were under filled in the gross fill then the fine fill would compensate for it. If this compensation happened with just one color, the mean value of N would be much more negative than for other colors. This turned out to be untrue. See Table 1

Table 1. Mean Values of N.

Color	Red	Orange	Yellow	Green	Blue	Brown
Mean	-0.32349	-0.1615	-0.82743	-0.16559	-0.61158	-0.03971
Std Dev of N	8.56129	3.028793	8.776667	8.648768	8.336004	8.553346