Sweetening STATISTICS

with Minitab® 16

Statistical Concepts: What M&M's Can Teach Us.

M&M's can provide an ideal way to make statistics immediately understandable to students who are timid about the subject. By simply weighing a few bags of the peanut variety and counting their contents, students gather useful data. Through applying statistics to this data on the weight, color, quality, and other factors of the M&M's—students take an everyday item such as candy and transform it into a powerful learning tool.

Are the M&M's weights accurate?

According to the package, each small bag of peanut M&M's should weigh 49.3 grams. To determine if the bag weights are on target, ask the students to measure the weight of say, 30 bags of M&M's and conduct a 1-sample t test.

In our analysis, the t-test results include a very small p-value of 0.000. Since the p-value is much smaller than an α-level of 0.05 or even 0.01, we can reject the null hypothesis and conclude that the average bag weight is not equal to 49.3 grams.

In this case, however, it's good news for consumers—we can see from the average weight of 52.040 grams that in fact, the average bag weight is significantly more than the 49.3 grams target.

One-Sample T: Weight(g)

```
Test of mu = 49.3 vs not = 49.3

Variable     N    Mean  StDev  SE Mean
Weight(g)   30  52.040  2.807    0.512

Variable        95% CI          T      P
Weight(g)  (50.992, 53.088)  5.35  0.000
```

Results generated using Stat > Basic Statistics > 1-Sample t.

Can two people measure the same weights?

If one person measures 30 bags of M&M's and their friend measures the same 30 bags using the same scale, will they get similar measurements? Students can test this hypothesis using a paired t test.

In our analysis, the p-value of 0.394 is larger than our chosen α-level of 0.05. Therefore, we fail to reject the null hypothesis and can conclude that there is not a significant difference between the average measurements taken by the two friends.

Paired T-Test and CI: Person1, Person2

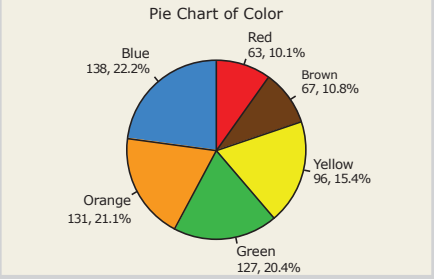
```
Paired T for Person1 - Person2

            N      Mean    StDev  SE Mean
Person1    30    52.040    2.807    0.512
Person2    30    52.047    2.802    0.512
Difference 30  -0.00700  0.04435  0.00810

95% CI for mean difference: (-0.02356,
   0.00956)
T-Test of mean difference = 0 (vs not = 0):
   T-Value = -0.86  P-Value = 0.394
```

Results generated using Stat > Basic Statistics > Paired t.

Are there equal amounts of each color in a bag?



Graph created using Graph > Pie Chart.

A pie chart lets us easily visualize the counts of each color.

For example, there were 138 blue M&M's and only 63 red M&M's in our sample. But is the difference between these counts statistically significant? A Chi-square test can tell us.

The p-value of 0.000 suggests that the observed counts are significantly different than what we would expect to see if there were an equal number of red, orange, yellow, green, blue and brown M&M's.

Chi-Square Goodness-of-Fit Test: CountPerColor

```
Using category names in Color

                          Test
Category  Observed  Proportion  Expected
Red             63    0.166667   103.667
Yellow          96    0.166667   103.667
Orange         131    0.166667   103.667
Blue           138    0.166667   103.667
Green          127    0.166667   103.667
Brown           67    0.166667   103.667

          Contribution
Category     to Chi-Sq
Red            15.9528
Yellow          0.5670
Orange          7.2069
Blue           11.3708
Green           5.2519
Brown          12.9689

  N  DF   Chi-Sq  P-Value
622   5  53.3183    0.000
```

Results generated using Stat > Tables > Chi-Square Goodness-of-Fit Test (One Variable).

Test and CI for One Proportion

```
Test of p = 0.15 vs p > 0.15


                                  95% Lower
Sample    X    N  Sample p          Bound
1        87  622  0.139871       0.117497


           Exact
Sample  P-Value
1         0.776
```

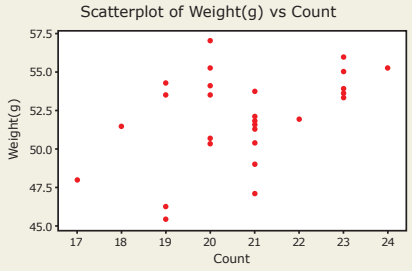
Results generated using Stat > Basic Statistics > 1 Proportion.

Do enough M&M's have the "m"?

M&M's are easily identified by the signature "m" printed on each piece of candy. It must pose a challenge to stamp the familiar symbol on a surface as uneven as a peanut M&M. It's not surprising, therefore, that sometimes this "m" is not perfectly printed.

Suppose there is a requirement that no more than 15% of M&M's have a misprinted "m." If we count the total number of M&M's and the number with misprints, we can conduct a 1 proportion test.

Of the 622 M&M's we evaluated, 87 had misprints. Using a 1 proportion test and an alternative hypothesis of greater than 15%, we get a p-value of 0.776. Because the p-value is greater than an α equal to 0.05, we can conclude that the proportion of misprinted M&M's is 15% or less.



Graph created using Graph > Scatterplot.

Is there a correlation between the number of M&M's in each bag and the bag weights?

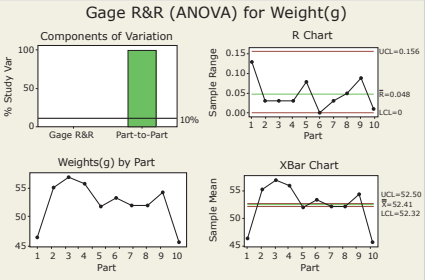
You may suspect that as the number of M&M's in each bag increases, so does the weight of the bags. We can use a scatterplot to examine this relationship and a correlation test to see if this theory is true.

The scatterplot of our data shows that, in general, as the number of M&M's in a bag increases, so does the weight. In addition, the correlation p-value of 0.011 indicates that we can reject the null hypothesis and conclude that there is a significant positive, linear relationship between the bag weights and the number of M&M's inside.

Correlations: Weight(g), Count

```
Pearson correlation of Weight(g) and
     Count = 0.458
P-Value = 0.011
```

Results generated using Stat > Basic Statistics > Correlation.



Results generated using Stat > Quality Tools > Gage Study > Gage R&R Study (Crossed).

Additional Considerations

Although measuring Gage repeatability and reproducibility is well beyond the scope of a typical university statistics course, you can also use M&M's to teach more advanced statistical techniques. For instance, we used Minitab to create a Gage R&R measurement plan, followed the plan to measure 10 bags of M&M's twice each in a random order, and then analyzed the measurement results.

The resulting total Gage R&R percent study variation of 1.14% is well below the 10% ideal, indicating that very little variation was due to the measurement system. In addition, the R Chart is in control, while the XBar Chart is out-of-control. Therefore, our measurement system is acceptable.

Additional Teaching Resources

Enhance your instruction with articles that describe how to use some of the most important statistical tools in Minitab.

Visit www.minitab.com/academic

MBREM16.STAT
10/11