

11.3: Goodness-of-Fit (1 of 2)

Learning Objectives

- Conduct a chi-square goodness-of-fit test. Interpret the conclusion in context.

In this section, we learn a new hypothesis test called the *chi-square goodness-of-fit test*. A goodness-of-fit test determines whether or not the distribution of a categorical variable in a sample fits a claimed distribution in the population.

We can answer the following research questions with a chi-square goodness-of-fit test:

- According to the manufacturer of M&M candy, the color distribution for plain chocolate M&Ms is 13% brown, 13% red, 14% yellow, 24% blue, 20% orange, and 16% green. Do the M&Ms in our sample suggest that the color distribution is different?
- During the presidential election of 2008, the Pew Research Center collected survey data that suggested that 24% of registered voters were liberal, 38% were moderate, and 38% were conservative. Is the distribution of political views different this year?
- The distribution of blood types for whites in the United States is 45% type O, 41% type A, 10% type B, and 4% type AB. Is the distribution of blood types different for Asian Americans?

The null hypothesis states a specific distribution of proportions for each category of the variable in the population. The alternative hypothesis says that the distribution is different from that stated in the null hypothesis. To test our hypotheses, we select a random sample from the population and determine the distribution of the categorical variable in the data. Of course, we need a method for comparing the observed distribution in the sample to the expected distribution stated in the null hypothesis.

Example

Distribution of Color in Plain M&M Candies

According to the manufacturer of M&M candy, the color distribution for plain chocolate M&Ms is 13% brown, 13% red, 14% yellow, 24% blue, 20% orange, and 16% green. This statement about the distribution of color in plain M&Ms is the null hypothesis. The alternative hypothesis says that this is not the distribution.

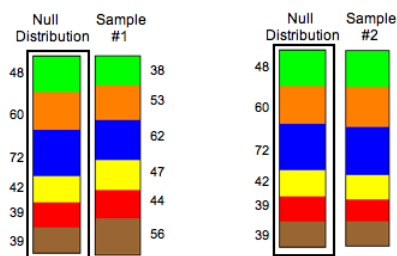
H_0 : The color distribution for plain M&Ms is 13% brown, 13% red, 14% yellow, 24% blue, 20% orange, 16% green.

H_a : The color distribution for plain M&Ms is different from the distribution stated in the null hypothesis.

We select a random sample of 300 plain M&M candies to test these hypotheses. If the sample has the distribution of color stated in the null hypothesis, then we expect 13% of the 300 to be brown, 13% of 300 to be red, 14% of 300 to be yellow, 24% of 300 to be blue, and so on. Here are the expected counts of each color for a sample of 300 candies:

| Color | Brown | Red | Yellow | Blue | Orange | Green |
|----------|----------------|----------------|----------------|----------------|----------------|----------------|
| Expected | $0.13(300)=39$ | $0.13(300)=39$ | $0.14(300)=42$ | $0.24(300)=72$ | $0.20(300)=60$ | $0.16(300)=48$ |

Of course, the distribution of color will vary in different samples, so we need to develop a way to measure how far a sample distribution is from the null distribution, something analogous to a z-score or T-score. Before we discuss this new measure, let's look at two random samples selected from the null distribution to practice recognizing different amounts of variability. We can compare the distributions visually using ribbon charts.



Which random sample deviates the most from the null distribution? We address this question in the next activity.

Learn By Doing

Observed Counts for Two Random Samples

Here are the observed counts for the two random samples shown above. This is the same information shown in the ribbon charts.

| Random Sample #1 (n=300) | | | | | | |
|--------------------------|-------|-----|--------|------|--------|-------|
| Color | Brown | Red | Yellow | Blue | Orange | Green |
| Observed | 56 | 44 | 47 | 62 | 53 | 38 |
| Expected | 39 | 39 | 42 | 72 | 60 | 48 |

| Random Sample #2 (n=300) | | | | | | |
|--------------------------|-------|-----|--------|------|--------|-------|
| Color | Brown | Red | Yellow | Blue | Orange | Green |
| Observed | 39 | 34 | 38 | 76 | 64 | 49 |
| Expected | 39 | 39 | 42 | 72 | 60 | 48 |

<https://assessments.lumenlearning.co...sessments/3794>

<https://assessments.lumenlearning.co...sessments/3795>

Learn By Doing

<https://assessments.lumenlearning.co...sessments/3796>

Statisticians use the following formula to measure how far the observed data are from the null distribution. It is called the *chi-square test statistic*. The Greek letter chi is written χ .

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

Notes about this formula:

- Recall that the symbol \sum means sum. Each category contributes a term to the sum, so the chi-square test statistic is based on the entire distribution. If the categorical variable has six categories, then the chi-square test statistic has six terms. If the categorical variable has three categories, then the chi-square test statistic has three terms, and so on.
- Notice that the difference “observed minus expected” for each category is part of the formula, but each difference is squared. This is necessary because the differences will add to 0, as we saw in the previous activity.
- Notice also that each squared difference is divided by the expected count for that category. The chi-square test statistic looks at the difference between the observed and expected counts *relative* to the size of the expected count.

Example

Calculating χ^2

For Sample 1, the chi-square test statistic is approximately 12.94. For Sample 2, the chi-square test statistic is approximately 1.53. Usually, we use technology to calculate χ^2 , but here we show two calculations in detail to illustrate how the formula works. Notice that we are adding six terms. Each term represents the deviation for one color category.

| Random Sample #1 (n=300) | | | | | | |
|--------------------------|-------|-----|--------|------|--------|-------|
| Color | Brown | Red | Yellow | Blue | Orange | Green |
| Observed | 56 | 44 | 47 | 62 | 53 | 38 |
| Expected | 39 | 39 | 42 | 72 | 60 | 48 |

$$\chi^2 = \frac{(56-39)^2}{39} + \frac{(44-39)^2}{39} + \frac{(47-42)^2}{42} + \frac{(62-72)^2}{72} + \frac{(53-60)^2}{60} + \frac{(38-48)^2}{48} = 7.41 + 0.64 + 1.39 + 1.39 + 0.82 + 2.08 = 12.94$$

Comment: In Sample 1, notice that both blue and green observed counts deviate from the expected counts by 10 candies. But green contributes more to the chi-square test statistic. This makes sense because the chi-square test statistic measures relative difference. Relative to the expected count of 48 green candies, an absolute error of 10 is large. It is almost 20% of the expected count. (10/48 is about 0.20). The squared difference relative to the expected count is 100/48, about 2.08. Relative to the expected count of 72 blue candies, an error of 10 candies is smaller. It is only about 14% of the expected count (10/72 is about 0.14). The squared difference relative to the expected count is 100/72, about 1.39.

Here is the chi-square calculation for Sample 2.

| | | | |
|-------|-------|-----|---|
| Color | Brown | Red | Y |
|-------|-------|-----|---|

| Color | Brown | Red | Yellow | Blue | Orange | Green |
|----------|-------|-----|--------|------|--------|-------|
| Observed | 39 | 34 | 38 | 76 | 64 | 49 |
| Expected | 39 | 39 | 42 | 72 | 60 | 48 |

[illegible]

Calculating χ^2

Learn By Doing

<https://assessments.lumenlearning.co...sessments/3717>

<https://assessments.lumenlearning.co...sessments/3718>

<https://assessments.lumenlearning.co...sessments/3719>

<https://assessments.lumenlearning.co...sessments/3716>

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** CC BY: Attribution

This page titled [11.3: Goodness-of-Fit \(1 of 2\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Bill Pelz](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.