

## 16.2: Introduction to Goodness-of-Fit Chi-Square

The first of our two  $\chi^2$  tests, the Goodness of Fit test, assesses the distribution of frequencies into different categories of one quantitative variable against any specific distribution. Usually this is equal frequency distributions because that's what we would expect to get if categorization was completely random, but it can also be a specific distribution. For example, if Dr. MO wanted to compare a specific class's frequency of each ethnicity to a specific distribution, it would make more sense to compare her class to the ethnic demographics of the college rather than an assumption that all of the ethnic groups would have the same number of students in the target class.

### Hypotheses for Chi-Square

All  $\chi^2$  tests, including the goodness-of-fit test, are non-parametric. This means that there is no population parameter we are estimating or testing against; we are working only with our sample data. This makes it more difficult to have mathematical statements for  $\chi^2$  hypotheses (symbols showing which group is bigger or whatever). The next section will walk through the mathematical hypotheses. For now, we will learn how to still state our hypotheses verbally.

### Research Hypothesis

The research hypothesis is that we expect a pattern of difference, and then we explain that pattern of difference.

Using Dr. MO's sample class, she works at a college that is designated as a Hispanic-Serving Institution (HSI), so we would expect a pattern of difference such that there will be more students who are Hispanic in her class than students from any other ethnic group.

### Null Hypotheses

For goodness-of-fit  $\chi^2$  tests, our null hypothesis is often that there is an equal number of observations in each category. That is, there is no pattern of difference between the frequencies in each category. Unless we're looking at the situation above in which we have a distribution of frequencies that we are comparing our sample to, the null hypothesis is that each group will be the same size.

### Degrees of Freedom and the $\chi^2$ table

Our degrees of freedom for the  $\chi^2$  test are based on the number of categories we have in our variable, not on the number of people or observations like it was for our other tests. Luckily, they are still as simple to calculate:

$$df = k - 1$$

Do you remember what "k" stood for when we discussed ANOVAs?

#### ? Exercise 16.2.1

What does "k" stand for?

#### Answer

The letter "k" usually stands for the number of groups. In Chi-Square, this would be the number of different categories.

So for our pet preference example, we have 3 categories, so we have 2 degrees of freedom. Our degrees of freedom, along with our significance level (still defaulted to  $\alpha = 0.05$ ) are used to find our critical values in the  $\chi^2$  table, which is next, or can be found through the [Common Critical Value Tables](#) at the end of this book.

This page titled [16.2: Introduction to Goodness-of-Fit Chi-Square](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [14.2: Goodness-of-Fit](#) by [Foster et al.](#) is licensed [CC BY-NC-SA 4.0](#). Original source: <https://irl.umsl.edu/oer/4>.