

## 14.2: Chi-squared Goodness of Fit

There are times when a hypothesis makes a statement about whether counts or proportions of subgroups in a sample will be approximately equal to specified counts or proportions. In these situations, a chi-squared goodness of fit test is appropriate. **The chi-squared goodness of fit** tests whether the counts of subgroups of a variable fit specified, expected proportions or counts. One thing that makes the chi-squared quite unusual is that the null hypothesis indicates the specified counts and, thus, sometimes a null result is desired. This is because the null states that the counts observed in the data will match those specified when using a goodness of fit test and the alternative hypothesis would, thus, be that the counts in the data do *not* match those specified.

Usually the chi-squared goodness of fit is used to test whether the distribution of data from a sample is similar to the distribution that is known or presumed true about a population. Chi squared goodness of fit, therefore, is assessing whether the data are a good fit to (i.e. are similar or dissimilar to) the counts or proportions specified in the null hypothesis. When a result is significant, it means that the counts are significantly different than those specified by the null hypothesis. When a result is not significant, it means that the counts are approximately equal to those specified by the null.

Let's take a look at this using an example. Suppose that among the population of college students, 25% were business majors, 25% were psychology majors, 20% were nursing majors, and 30% were statistics majors. Suppose that you wanted to test whether the proportion of students in each of these majors at a small college in Statistonia students, overall. In this case, you would expect that at the proportions of majors at Statistonia College would be approximately 25% business majors, 25% psychology majors, 20% nursing majors, and 30% statistics majors. In this case, these percentages would be used as the specified proportions and a non-significant result would mean that the counts observed at Statistonia College were *not significantly different* from those in the population. A significant result would mean that the counts observed at Statistonia College were *significantly different* from those in the population. This is because a significant result occurs when counts observed in the data are different from the expected counts. (Note: We will see how to test this specific example in the SPSS section).

There are two kinds of counts considered in a chi-squared goodness of fit: Observed counts and expected counts. Observed counts refer to the counts that exist in the sample data. The expected counts are based on either known population proportions or on a hypothesis about what counts are expected to be. We will review these in detail later in this chapter. For now, let's focus on being clear on the general idea of this test. Below shows a summary of how to interpret results from a chi-squared goodness of fit. Refer back to this as needed to remind yourself what the test is comparing and what a significant result means compared to what a non-significant result means.

Interpreting Significance in a Chi-squared Goodness of Fit

Significant Result	Counts observed in the sample are significantly different from those expected based on the population or a hypothesis.	$Observed\ Counts \neq Expected\ Counts$
Non Significant Result	Counts observed in the sample are <i>not</i> significantly different from those expected based on the population or a hypothesis. <i>Alternative wording:</i> Counts observed in the sample are approximately equal to those expected based on the population or a hypothesis.	$Observed\ Counts \approx Expected\ Counts$

### Variables

One or more grouping/categorical variables are used in chi-square. These variables can either include nominal or ordinal data which are separated into categories for the purposes of hypothesis testing. It is also possible to create categories with interval or ratio data using intervals, however, this is rarely used because other tests are generally a better fit to those data and their corresponding hypotheses.

In a univariate chi-square, one variable with at least two categories or levels is used. However, it is possible to use a version that looks at the counts of categories of two or more variables at once such as when testing whether the counts by gender and race/ethnicity in a sample are proportionally similar to those in the population. However, the focus of this section of the chapter will be univariate chi-square.

## Data and Assumptions

Each statistical test has some assumptions which must be met in order for the formula to function properly. In keeping, there are a few assumptions about the data which must be met before a chi-squared goodness of fit is used. The most important is that no category can have an expected count lower than 5. Thus, sample sizes should be large enough that each expected count is 5 or higher.

## Hypotheses

Hypotheses for chi-squared goodness of fit focus on whether observed counts are similar or significantly different from those expected based on populations or theories (and their corresponding hypotheses). We will focus on the non-directional hypothesis for this chapter because the goodness of fit test functions as an omnibus test. It can indicate whether counts are or are not as expected but cannot, on its own, indicate which counts are different than expected and which are not. When there are only two subgroups, only the chi-squared goodness of fit test is needed to know whether both groups were significantly different than expected. When there are three or more groups and a significant result, however, some ambiguity remains about the pairwise comparisons. Thus, what the goodness of fit test tells us is, overall, whether counts of subgroups are as expected or not. The non-directional research and corresponding null hypotheses for chi-squared goodness of fit can be summarized as follows:

Non-Directional Hypothesis for Chi-squared Goodness of Fit

Research hypothesis	Counts observed in the sample are significantly different from those expected based on the population or a theory.	$H_A : f_{\text{observed}} \neq f_{\text{expected}}$
Null hypothesis	Counts observed in the sample are <i>not</i> significantly different from those expected based on the population or a theory.	$H_0 : f_{\text{observed}} = f_{\text{expected}}$

In this example  $f_{\text{observed}}$  refers to the frequency of counts observed in the data for each subgroup or category of the test variable and  $f_{\text{expected}}$  refers to the frequency of counts expected in the data for each subgroup or category of the test variable based on what is known or theorized about the population. When the counts in the sample are similar or equal to those expected, the result will be non-significant and, thus, favor the null hypothesis. Conversely, when the counts in the sample are notably dissimilar from those expected, the result will be significant and, thus, support the alternative hypothesis.

## Limitations

A limitation of chi-squared goodness of fit is that it cannot always be used to determine cause effect relationships; however, under some conditions this may be appropriate using a different form of chi-squared known as a chi-squared test for independence (which we will cover later in this chapter). In addition, the chi-squared goodness of fit test can be hard to interpret when there are many categories and may not be an option when any of those groups has an expected count that is less than 5. However, it otherwise is a good option for comparing group counts when sufficient data are available.

## Reading Review 14.1

1. What kinds of variables are tested using chi-squared goodness of fit tests?
2. What does a significant result indicate when using chi-squared goodness of fit tests?
3. What does a non-significant result indicate when using chi-squared goodness of fit tests?

This page titled 14.2: Chi-squared Goodness of Fit is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by .