Chi-Square Goodness of Fit Test

This lesson explains how to conduct a **chi-square goodness of fit test**. The test is applied when you have one [categorical variable](https://stattrek.com/Help/Glossary.aspx?Target=Categorical%20variable) from a single population. It is used to determine whether sample data are consistent with a hypothesized distribution.

For example, suppose a company printed baseball cards. It claimed that 30% of its cards were rookies; 60% were veterans but not All-Stars; and 10% were veteran All-Stars. We could gather a random sample of baseball cards and use a chi-square goodness of fit test to see whether our sample distribution differed significantly from the distribution claimed by the company. The [sample problem](https://stattrek.com/chi-square-test/goodness-of-fit.aspx#example1) at the end of the lesson considers this example.

When to Use the Chi-Square Goodness of Fit Test

The chi-square goodness of fit test is appropriate when the following conditions are met:

* The sampling method is [simple random sampling](https://stattrek.com/Help/Glossary.aspx?Target=Simple%20random%20sampling).
* The variable under study is [categorical](https://stattrek.com/Help/Glossary.aspx?Target=Categorical%20variable).
* The expected value of the number of sample observations in each [level](https://stattrek.com/Help/Glossary.aspx?Target=Level) of the variable is at least 5.

This approach consists of four steps: (1) state the hypotheses, (2) formulate an analysis plan, (3) analyze sample data, and (4) interpret results.

State the Hypotheses

Every hypothesis test requires the analyst to state a [null hypothesis](https://stattrek.com/Help/Glossary.aspx?Target=Null%20hypothesis) (Ho) and an [alternative hypothesis](https://stattrek.com/Help/Glossary.aspx?Target=Alternative%20hypothesis) (Ha). The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false; and vice versa.

For a chi-square goodness of fit test, the hypotheses take the following form.

* Ho: The data are consistent with a specified distribution.
* Ha: The data are *not* consistent with a specified distribution.

Typically, the null hypothesis (Ho) specifies the proportion of observations at each level of the categorical variable. The alternative hypothesis (Ha) is that *at least* one of the specified proportions is not true.

Formulate an Analysis Plan

The analysis plan describes how to use sample data to accept or reject the null hypothesis. The plan should specify the following elements.

* Significance level. Often, researchers choose [significance levels](https://stattrek.com/Help/Glossary.aspx?Target=Significance%20level) equal to 0.01, 0.05, or 0.10; but any value between 0 and 1 can be used.
* Test method. Use the [chi-square goodness of fit test](https://stattrek.com/Help/Glossary.aspx?Target=Chi-square%20goodness%20of%20fit%20test) to determine whether observed sample frequencies differ significantly from expected frequencies specified in the null hypothesis. The chi-square goodness of fit test is described in the next section, and demonstrated in the sample problem at the end of this lesson.

Analyze Sample Data

Using sample data, find the degrees of freedom, expected frequency counts, test statistic, and the P-value associated with the test statistic.

* Degrees of freedom. The [degrees of freedom](https://stattrek.com/Help/Glossary.aspx?Target=Degrees%20of%20freedom) (DF) is equal to the number of levels (k) of the categorical variable minus 1.

DF = k - 1

* Expected frequency counts. The expected frequency counts at each level of the categorical variable are equal to the sample size times the hypothesized proportion from the null hypothesis

Ei = npi

where Ei is the expected frequency count for the *i*th level of the categorical variable, n is the total sample size, and pi is the hypothesized proportion of observations in level *i*.

* Test statistic. The test statistic is a chi-square random variable (Χ2) defined by the following equation.

Χ2 = Σ [ (Oi - Ei)2 / Ei ]

where Oi is the observed frequency count for the *i*th level of the categorical variable, and Ei is the expected frequency count for the *i*th level of the categorical variable.

* P-value. The P-value is the probability of observing a sample statistic as extreme as the test statistic. Since the test statistic is a chi-square, use the [Chi-Square Distribution Calculator](https://stattrek.com/Tables/ChiSquare.aspx) to assess the probability associated with the test statistic. Use the degrees of freedom computed above.

Interpret Results

If the sample findings are unlikely, given the null hypothesis, the researcher rejects the null hypothesis. Typically, this involves comparing the P-value to the [significance level](https://stattrek.com/Help/Glossary.aspx?Target=Significance%20level), and rejecting the null hypothesis when the P-value is less than the significance level.

Test Your Understanding

**Problem**

Acme Toy Company prints baseball cards. The company claims that 30% of the cards are rookies, 60% veterans but not All-Stars, and 10% are veteran All-Stars.

Suppose a random sample of 100 cards has 50 rookies, 45 veterans, and 5 All-Stars. Is this consistent with Acme's claim? Use a 0.05 level of significance.

**Solution**

The solution to this problem takes four steps: (1) state the hypotheses, (2) formulate an analysis plan, (3) analyze sample data, and (4) interpret results. We work through those steps below:

* **State the hypotheses.** The first step is to state the null hypothesis and an alternative hypothesis.
  + Null hypothesis: The proportion of rookies, veterans, and All-Stars is 30%, 60% and 10%, respectively.
  + Alternative hypothesis: At least one of the proportions in the null hypothesis is false.
* **Formulate an analysis plan**. For this analysis, the significance level is 0.05. Using sample data, we will conduct a [chi-square goodness of fit test](https://stattrek.com/Help/Glossary.aspx?Target=Chi-square%20goodness%20of%20fit%20test) of the null hypothesis.
* **Analyze sample data**. Applying the chi-square goodness of fit test to sample data, we compute the degrees of freedom, the expected frequency counts, and the chi-square test statistic. Based on the chi-square statistic and the [degrees of freedom](https://stattrek.com/Help/Glossary.aspx?Target=Degrees%20of%20freedom), we determine the [P-value](https://stattrek.com/Help/Glossary.aspx?Target=P-value).

DF = k - 1 = 3 - 1 = 2 (Ei) = n \* pi  
(E1) = 100 \* 0.30 = 30  
(E2) = 100 \* 0.60 = 60  
(E3) = 100 \* 0.10 = 10  
Χ2 = Σ [ (Oi - Ei)2 / Ei ]   
Χ2 = [ (50 - 30)2 / 30 ] + [ (45 - 60)2 / 60 ] + [ (5 - 10)2 / 10 ]  
Χ2 = (400 / 30) + (225 / 60) + (25 / 10) = 13.33 + 3.75 + 2.50 = 19.58

where DF is the degrees of freedom, k is the number of levels of the categorical variable, n is the number of observations in the sample, Ei is the expected frequency count for level i, Oi is the observed frequency count for level i, and Χ2 is the chi-square test statistic.

The P-value is the probability that a chi-square statistic having 2 degrees of freedom is more extreme than 19.58.

We use the [Chi-Square Distribution Calculator](https://stattrek.com/Tables/ChiSquare.aspx) to find P(Χ2 > 19.58) = 0.0001.

* **Interpret results**. Since the P-value (0.0001) is less than the significance level (0.05), we cannot accept the null hypothesis.

**Note:** If you use this approach on an exam, you may also want to mention why this approach is appropriate. Specifically, the approach is appropriate because the sampling method was simple random sampling, the variable under study was categorical, and each level of the categorical variable had an expected frequency count of at least 5.