11-2 Goodness of Fit Test

In this section we consider sample data consisting of observed frequency counts arranged in a single row or column (called a one-way frequency table). We will use a hypothesis test for the claim that the observed frequency counts agree with some claimed distribution, so that there is a *good fit* of the observed data with the claimed distribution.

A *goodness-of-fit test* is used to test the hypothesis that an observed frequency distribution fits (or conforms to) some claimed distribution.

**H₀**: The random variable follows a particular distribution.

**H₁**: The random variable does not follow the distribution specified in **H₀**.

Ex 1) Consider the observed frequencies and relative frequencies of browser preference from a survey of 200 Internet users.

<table>
<thead>
<tr>
<th>Browser</th>
<th>Observed frequency</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft IE</td>
<td>140</td>
<td>0.785</td>
</tr>
<tr>
<td>Firefox</td>
<td>40</td>
<td>0.15</td>
</tr>
<tr>
<td>Safari/other</td>
<td>20</td>
<td>0.065</td>
</tr>
</tbody>
</table>

The following model shows how the market shares are distributed in the null hypothesis:

\[ H₀: P_{\text{Ms IE}} = 0.785, \quad P_{\text{Firefox}} = 0.15, \quad P_{\text{Safari/other}} = 0.065 \]

**H₁**: The random variable does not follow the distribution specified in **H₀**.
How a Goodness-of-Fit Test Works

The goodness-of-fit test is based on a comparison of the observed frequencies (actual data from the field) with the expected frequencies when \( H_0 \) is true. That is, we compare what we actually see with what would expect to see if \( H_0 \) were true. If the difference between the observed and expected frequencies is large, we reject \( H_0 \).

As usual, it comes down to how large a difference is large. The hypothesis we conduct to answer this question relies on \( \chi^2 \) distribution.

Performing the \( \chi^2 \) Goodness of Fit Test

The following conditions must be met:

- The data have been randomly selected.
- The sample data consist of frequency counts for each of the different categories.
- None of the expected frequencies is less than 1.
- For each category, the expected frequency is at least 5.

Finding Expected Frequencies

The expected frequency for a category is the frequency that would occur if the data actually have the distribution that is being claimed. For the \( i^{th} \) category, the expected frequency is \( E_i = n \cdot p_i \), where \( n \) represent the number of trials and \( p_i \) represents the population proportion for the \( i^{th} \) category.

If we assume that all expected frequencies are equal, then each expected frequency is \( E = n/k \), where \( n \) is the total number of observations and \( k \) is the number of categories.

The \( \chi^2 \) goodness of fit test may be performed using (a) the critical value, and (b) the p-value method.
(a) $\chi^2$ goodness of fit test. (Critical value method)

Step 1: State the hypotheses and check the conditions.

The null hypothesis is states that the qualitative random variable follows a particular distribution. The alternative hypothesis states that the random variable does not follow that distribution.

Step 2: Find the $\chi^2$ critical value, $\chi^2_{critical}$, from table A-4 by using $k-1$ degrees of freedom, where $k$ is the number of categories.

Note, Goodness-of-fit hypothesis are always right tailed.

And state the rejection rule.

Reject if $\chi^2_{data} > \chi^2_{critical}$.

Step 3: Find the test statistic $\chi^2_{data}$.

$$\chi^2_{data} = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where $O_i$ = observed frequency, and $E_i$ = expected frequency.

Step 4: State the conclusion and the interpretation.
Ex 2) Perform the hypothesis test shown in example 1, use 0.05 as a significance level.

\[ H_0: P_{\text{MS IE}} = 0.785, \ P_{\text{firefox}}=0.15, \ P_{\text{Safari/other}} = 0.065 \]

\[ H_1: \text{The random variable does not follow the distribution specified in } H_0. \]

<table>
<thead>
<tr>
<th>Browser</th>
<th>Observed freq</th>
<th>Relative freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>( O_i )</td>
<td>( p_i )</td>
</tr>
<tr>
<td>MS IE</td>
<td>140</td>
<td>0.785</td>
</tr>
<tr>
<td>Firefox</td>
<td>40</td>
<td>0.15</td>
</tr>
<tr>
<td>Safari/other</td>
<td>20</td>
<td>0.065</td>
</tr>
</tbody>
</table>

Interpretation: There is evidence that the random variable browser does not follow the distribution in \( H_0 \). In other words, there is evidence that the market shares for internet browsers have changed.
Note carefully what this conclusion says and what it doesn’t say. The $\chi^2$ goodness of fit test shows that there is evidence that the random variable does not follow the distribution specified in $H_0$. In particular, the conclusion does not state, for example, that Firefox’s proportion is significantly greater.

(b) $\chi^2$ goodness of fit test: (p-value Method)

Step 1: State the hypotheses and check the conditions.

Step 2: Find the test statistic $\chi^2_{data}$.

$$\chi^2_{data} = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where $O_i$ = observed frequency, and $E_i$ = expected frequency

Step 3: Find the p-value.

$$p\text{-value} = P(\chi^2 > \chi^2_{data})$$

Step 4: State the conclusion and the interpretation.

Ex (3) The following tables show figures on the market share of cable modem, DSL, and wireless broadband from a 2002 survey and a 2006 survey which was based on a random sample of 1000 home broadband users. Test whether the population proportions have changed since 2002, using the p-value method, and level of significance is 0.05.

<table>
<thead>
<tr>
<th>2002 broadband adoption survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cable modem</td>
</tr>
<tr>
<td>67%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2006 broadband adoption survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cable modem</td>
</tr>
<tr>
<td>410</td>
</tr>
</tbody>
</table>