The most commonly employed non-parametric hypothesis test is the Chi-square goodness of fit test. Goodness of fit refers to the degree to which observed events “fit with” our expectations.

If we have a system with observed and expected events, where $O_i$ and $E_i$ are the numbers of observed and expected events in category $i$ and there are $k$ categories, then

$$u = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

is a Chi-square random variable with $k-1$ degrees of freedom. We use the random variable $u$ to test the hypothesis that generated the expectations.

EXAMPLE
Suppose we have a die that we suspect is unfair. Then we wish to test the null hypothesis that the die is fair against the alternative hypothesis that the die is not fair, and, based on our suspicions, we anticipate rejecting the null hypothesis.

Thus, we have

$$H_0: \text{the die is fair}$$
$$H_A: \text{the die is not fair}$$

Note that neither the null nor the alternative hypothesis mentions a parameter; hence the term “non-parametric test”.

Suppose we roll the die 36 times. Under the null hypothesis, i.e., if the die were fair, we would expect to see six 1’s, six 2’s, six 3’s, six 4’s, six 5’s, and six 6’s. Suppose we observe four 1’s, eight 2’s, eight 3’s, five 4’s, four 5’s, and seven 6’s. This data can be put into a table as follows;

<table>
<thead>
<tr>
<th>rolls</th>
<th>1’s</th>
<th>2’s</th>
<th>3’s</th>
<th>4’s</th>
<th>5’s</th>
<th>6’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>expected</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>observed</td>
<td>4</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>
The value of the test statistics is

\[ u = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(4-6)^2}{6} + \frac{(8-6)^2}{6} + \frac{(8-6)^2}{6} + \frac{(5-6)^2}{6} + \frac{(4-6)^2}{6} + \frac{(7-6)^2}{6} \]

which reduces to

\[ \frac{4}{6} + \frac{4}{6} + \frac{4}{6} + \frac{1}{6} + \frac{4}{6} + \frac{1}{6} = \frac{18}{6} = 3. \]

The critical region is the interval on the horizontal axis to the right of \( \chi^2_{5 \, \text{dof.}} \). As before, the critical region is determined by the alternative hypothesis. If the alternative hypothesis is true, then the numerators of the elements of \( u \) should be large, where upon \( u \) should be large. Hence, Chi-square goodness of fit tests are always right-tailed tests.

Suppose we test at the .05 level of significance. Then \( \chi^2_{5 \, \text{dof.}} = 11.071 \). Since

Graphically, we have the following:

Since \( u* \) is not in the critical region, we maintain the null hypothesis. Simply put, we have failed to establish that the die is unfair.

Now suppose we get a little more sophisticated and suspect that the die is unfair in the proportions 1:2:2:1:1:1. Then the hypotheses are

- \( H_0 \): the proportions are 1:2:2:1:1:1
- \( H_A \): the proportions are not 1:2:2:1:1:1

These proportions add to 1+2+2+1+1+1 = 8, so that the null hypothesis says

1/8 of the rolls will be 1’s, 2/8 will be 2’s, 2/8 will be 3’s, 1/8 will be 4’s, 1/8 will be 5’s, and 1/8 will be 6’s.

Suppose we roll the die 48 times and observe four 1’s, sixteen 2’s, fifteen 3’s, three 4’s, eight 5’s, and two 6’s. If the null hypothesis is true, then we would expect to see, in 48
rolls of the die, six 1’s, twelve 2’s, twelve 3’s, six 4’s, six 5’s, and six 6’s. In tabular form, we have the following:

<table>
<thead>
<tr>
<th>rolls</th>
<th>1’s</th>
<th>2’s</th>
<th>3’s</th>
<th>4’s</th>
<th>5’s</th>
<th>6’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>expected</td>
<td>6</td>
<td>12</td>
<td>12</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>observed</td>
<td>4</td>
<td>16</td>
<td>15</td>
<td>3</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

The value of the test statistics is

\[ u = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(4 - 6)^2}{6} + \frac{(16 - 12)^2}{12} + \frac{(15 - 12)^2}{12} + \frac{(3 - 6)^2}{6} + \frac{(8 - 6)^2}{6} + \frac{(2 - 6)^2}{6} \]

which reduces to

\[ \frac{4}{6} + \frac{16}{12} + \frac{9}{12} + \frac{9}{6} + \frac{4}{6} + \frac{16}{6} = 6.33. \]

As before, the critical region is the interval on the horizontal axis to the right of \( \chi^2_{5 \text{ dof}} \). If we conduct an .05 test, then \( \chi^2_{5 \text{ dof}} \) is again 11.071.

Graphically, we have the following:

Since \( u^* \) is not in the critical region, we maintain the null hypothesis. Simply put, we have maintained the hypothesis that the die is unfair in proportions 1:2:2:1:1:1.

The text uses the goodness-of-fit test, in frequency format, as a test of independence. (See pages 163-166.) The key point is that the Chi-square test can be employed in many different ways.

PROBLEM:
Ch. 4, #22