Goodness-of-Fit Test

In the following examples, identify the null hypothesis, alternative hypothesis, test statistic, critical value and state the conclusion of the test.

Example #1

Absenteeism of college students from math classes is a major concern to math instructors because missing class appears to increase the drop rate. Suppose that a study was done to determine if the actual student absenteeism rate follows faculty perception. The faculty expected that a group of 100 students would miss class according to the following table. A random survey across all mathematics courses was then done to determine the actual number of absences in a course. Assume $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>Number of Absences</th>
<th>Expected Number of Absences</th>
<th>Actual Number of Absences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 2</td>
<td>50</td>
<td>35</td>
</tr>
<tr>
<td>3 - 5</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>6 - 8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>9 - 12</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>12+</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

$H_0$: Student absenteeism fits faculty perception

$H_a$: Student absenteeism does not fit faculty perception

Test statistic

$$
\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(35-50)^2}{50} + \frac{(40-30)^2}{30} + \frac{(20-12)^2}{12} + \frac{(1-6)^2}{6} + \frac{(4-2)^2}{2}$$

$$= 19.333$$

$\alpha = 0.05$

$df = k - 1 = 5 - 1 = 4$

$\chi^2 = 9.488 \text{ Critical value}$

Since $19.33 > 9.488$ we reject the critical region

Reject $H_0$ and conclude that the student absenteeism does not fit faculty perception
Example #2

A factory manager needs to understand how many products are defective versus how many are produced. The number of expected defects is listed in the following table. A random sample was taken to determine the actual number of defects. Assume $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>Number Produced</th>
<th>Expected Number of Defective</th>
<th>Actual Number of Defective</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 100</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>101 - 200</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>201 - 300</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>301 - 400</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>401 - 500</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

$H_0$: The actual number of defective items fits the expected number of defective items.

$H_a$: The actual number of defective items does not fit the expected number of defective items.

Test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{(5-5)^2}{5} + \frac{(7-6)^2}{6} + \frac{(8-7)^2}{7} + \frac{(9-8)^2}{8} + \frac{(11-10)^2}{10}$

$= 0.5345$

$\alpha = 0.05$

$df = k - 1 = 5 - 1 = 4$

$\chi^2_{0.05} = 9.488$

Since $0.5345 < 9.488$ our test statistic is not in the critical region.

Fail to reject $H_0$ and conclude that the actual mean distribution of defective items fits the expected distribution of defective items.
Test of Independence

In the following examples, identify the null hypothesis, alternative hypothesis, test statistic, critical value and state the conclusion of the test.

Example #1

When Paul Meier of the University of Chicago wrote about the 1954 clinical trial of the Salk poliomyelitis vaccine, he stated that it was "the biggest public health experiment ever." In that experiment, some children were given the Salk poliomyelitis vaccine and others were given a placebo. The experiments results are summarized in the following table. Determine whether a person develops paralytic polio is dependent on whether that person was given the Salk vaccine or a placebo. Assume \( \alpha = 0.05 \).

<table>
<thead>
<tr>
<th></th>
<th>Polio</th>
<th>No Polio</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salk Vaccine Treatment Group</td>
<td>33</td>
<td>200712</td>
<td>200745</td>
</tr>
<tr>
<td>Placebo Group</td>
<td>115</td>
<td>201114</td>
<td>201229</td>
</tr>
<tr>
<td>Column Total</td>
<td>148</td>
<td>401836</td>
<td>401974</td>
</tr>
</tbody>
</table>

\[
H_0: \text{Paralytic Polio is independent of the Salk vaccine}\\
H_1: \text{Paralytic Polio is not independent of the Salk vaccine}
\]

Test Statistic \( \chi^2 \) : \[
\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{200745}{401974} = 73.911
\]

\[
E_1 = \frac{(200745)(148)}{401974} = 73.911
\]

\[
E_2 = \frac{(200745)(401826)}{401974} = 200671.089
\]

\[
E_3 = \frac{(201114)(148)}{401974} = 74.089
\]

\[
E_4 = \frac{(201114)(401826)}{401974} = 201154.911
\]

\[
\chi^2 = \frac{(33 - 73.911)^2}{73.911} + \frac{(200712 - 200671.089)^2}{200671.089} + \frac{(115 - 74.089)^2}{74.089} + \frac{(201114 - 201154.911)^2}{201154.911}
\]

\[
= 45.252 + 0.005 + 22.591 + 0.008 = 45.252 < \text{Test Statistic}
\]

\[
\alpha = 0.05\\
df = (r-1)(c-1) = (2-1)(2-1) = 1
\]

Reject \( H_0 \) since 45.252 > 3.481

Conclusion: Paralytic polio is not independent of the Salk vaccine.
Test of Homogeneity

In the following examples, identify the null hypothesis, alternative hypothesis, test statistic, critical value and state the conclusion of the test.

Example #1

Scientists wanted to study whether the color of your eyes made you more predisposed to smoking. They collected data from both smokers and non-smokers from those with blue eyes, green eyes, and brown eyes which is summarized in the following table. Determine whether the proportion of eye colors for smokers and non-smokers are the same or not. Assume $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>Brown Eye</th>
<th>Blue Eye</th>
<th>Green Eye</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>70</td>
<td>20</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Non-Smoker</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Column Total</td>
<td>75</td>
<td>25</td>
<td>15</td>
<td>115</td>
</tr>
</tbody>
</table>

$E_i = \frac{(\text{Row Total}) \times (\text{Column Total})}{\text{Grand Total}}$

$E_1 = \frac{(100) \times (75)}{115} = 66.22$

$E_2 = \frac{(100) \times (25)}{115} = 21.74$

$E_3 = \frac{(100) \times (15)}{115} = 13.29$

$E_4 = \frac{(15) \times (75)}{115} = 9.78$

$E_5 = \frac{(15) \times (25)}{115} = 3.26$

$E_6 = \frac{(15) \times (15)}{115} = 7.17$

$
\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{(70 - 66.22)^2}{66.22} + \frac{(20 - 21.74)^2}{21.74} + \frac{(10 - 13.24)^2}{13.24} + \frac{(5 - 9.78)^2}{9.78} + \frac{(5 - 3.26)^2}{3.26} + \frac{(5 - 7.17)^2}{7.17}
$

$\chi^2 = 5.12 \leftarrow \text{Test Statistic}$

$\alpha = 0.05$

$\chi^2_{df} = (4-1) = 3$,

$\chi^2_{df} = (4-1) \times (3-1) = (3) \times (2) = 6$

$\chi^2 = 5.12 < 5.991$

Fail to reject $H_0$ since $5.12 < 5.991$

Conclusion: The proportion of eye colors is the same in ---