Chapter 26  
Chi-Square Tests

- Goodness of fit: \( df = \# \text{cells of table} - 1 \) (C-1 for cells arranged in a row).
- Homogeneity-Independence: \( df = (R-1)(C-1) \). Analyzed the same.
- Homogeneity is when "row counts are sampled separately."
- Chi-Square Statistic is always calculated \( \sum_{\text{cells of a table of counts}} \frac{(\text{obs} - \text{exp})^2}{\text{exp}} \geq 0 \).
- Significance level (P-value) = probability of getting chi-square statistic that is at least as large as your data gave if the null hypothesis is correct.
- Using a chi-square table:

  \[
  \begin{array}{c|c}
  \text{P-value} & 0.0145 \\
  \hline
  \text{df} & 0.0145 \\
  \end{array}
  \]

  \[
  \begin{array}{c|c|c}
  30 & 49.34 & \text{Prob(chi-sq with df 30 > 49.34) = 0.0145} \\
  \end{array}
  \]

- Require all expected counts \( \geq 5 \). Not required of observed counts!
- Can merge cells to achieve \( \geq 5 \) requirement.
- Can add independent chi-square statistics to combine experimental results. Add df to get the applicable df for the combined data.
- Remember: If you choose to "reject \( H_0 \) whenever \( P < 0.001 \)" then your type I error probability is 0.001. That is, if \( H_0 \) is true then you will "reject \( H_0 \)" with probability 0.001 (error of type I).
- Chance of error of type II \( \rightarrow 0 \) with lots of data. That is, if \( H_0 \) is false you are nearly certain to reject \( H_0 \) with enough data.
Goodness of fit example: Is the coin fair?

Suppose we find 63 heads in 100 tosses?

\[ H_0 : p = 0.5 \quad H_1 : p \neq 0.5 \]

Data: 63 heads in 100 tosses.

\[ \hat{p} = \frac{63}{100} = 0.63, \quad \hat{q} = 1 - \hat{p} = \frac{37}{100} = 0.37 \]

Test statistic \[ = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \sim Z \text{ if } H_0 \text{ is true (i.e. } p = 0.5) \]

Reject if test statistic is too far from 0 (2 - sided test).

Test statistic evaluates to \[ = \frac{0.63 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{100}}} = 2.60. \]

2-sided alternative hypothesis \[ \sqrt{\frac{p_0 q_0}{n}} \quad \sqrt{\frac{0.5 \times 0.5}{100}} \]

\[ P - value = 2 P (Z > 2.60) = 2 (1 - 0.9953) = 0.0094. \approx .01 \]

Conclusion: It is around 1% likely that a fair coin would produce either \( \leq 37 \) or \( \geq 63 \) heads. The data does exhibit a rarely seen departure from 0.5.
Is the coin fair? Apply chi-square instead of z-test.

Same data as above (\( P = 0.0094 \)) but analyzed by chi-square.

\[
\text{chi-square statistic} = \sum_{\text{cells}} \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = 6.76.
\]

\[
\frac{(63 - 50)^2}{50} + \frac{(37 - 50)^2}{50} = 3.38 + 3.38 = 6.76
\]

\( \text{DF} = C - 1 = 2 - 1 = 1 \)

\( P = 0.0093 \)

\[
\begin{array}{c|c|c}
\text{df} & 0.0093 & X^2 \\
1 & 6.76 & \text{STATISTIC}
\end{array}
\]

a. The P-value, using the z-test of chapter 19, is 0.0094.
b. This closely agrees with the P-value 0.0093 found using the chi-square test of Chapter 26.

Either the coin is fair and this data is “luck of the draw bad” or the coin is not fair. We may never know which.
Can students act like equal probability selectors?

Apply chi-square *goodness of fit* to their choice-data.

**H0:** choices 1, 2, 3, 4 are equally likely.

**H1:** not equally likely

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>total 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>expected</td>
<td>55/4</td>
<td>55/4</td>
<td>55/4</td>
<td>55/4</td>
<td>55</td>
</tr>
<tr>
<td>observed</td>
<td>8</td>
<td>27</td>
<td>12</td>
<td>8</td>
<td>55</td>
</tr>
</tbody>
</table>

\[
\text{chi-square statistic} = \sum \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = 17.8
\]

\[
\frac{(8 - 55/4)^2}{55/4} + \frac{(27 - 55/4)^2}{55/4} + \frac{(12 - 55/4)^2}{55/4} + \frac{(8 - 55/4)^2}{55/4} = 17.8
\]

DF = C - 1 = 4 - 1 = 3  
P = 0.0005

If students choose with equal probability, a chi-square at least as large as our 17.8 would only be seen with probability 0.0005. Which is it? We may never know for sure.
Is full moon statistically related to incidence of crime?

<table>
<thead>
<tr>
<th></th>
<th>FULL MOON</th>
<th>NOT FULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>violent</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>property</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>drugs</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td>abuse</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>other</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>62</td>
</tr>
</tbody>
</table>

\[
\text{expected} = \begin{pmatrix}
1.91304 & 1.7971 \\
18.1739 & 17.0725 \\
22.0000 & 20.6667 \\
11.9565 & 11.2319 \\
7.17391 & 6.73913
\end{pmatrix}
\]

\[
\chi^2 \text{ for } \text{INDEPENDENCE}\text{ } \chi^2_{\text{HOMOGENEITY}}
\]

\[
\text{df} = (R-1)(C-1) = 4
\]

a. Some expected counts are less than 5.
b. Possible "confounding factors."

\[\text{e.g. } \text{moon phases might coincide with holidays or "game nights," and thus with crimes.}\]
Merge cells to meet the “minimum of 5” requirement.

<table>
<thead>
<tr>
<th>FULL MOON</th>
<th>NOT FULL</th>
<th>Merge with abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>violent</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>property</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>drugs</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td>abuse</td>
<td>[11+2 = 13]</td>
<td>[14+2 = 16]</td>
</tr>
<tr>
<td>other</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>62</td>
</tr>
</tbody>
</table>

\[\frac{66}{138} \cdot \frac{29}{138}\]

\[
\begin{pmatrix}
18.1739 & 17.0725 \\
22.2 & 20.6667 \\
13.8696 & 13.029 \\
7.17391 & 6.73913
\end{pmatrix}
\]

\[
\text{df} = (4-1)(2-1) = 3
\]

\[
\chi^2 = \sum \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = 3.528 \text{ (merged)}
\]

\[
P = 0.317 \text{ (no evidence against the hypothesis of homogeneity)}
\]

Seems that we don’t have to worry over confounding factors.
Is there a statistical association between tattooing and hepatitis C?

<table>
<thead>
<tr>
<th></th>
<th>Hepatitis C</th>
<th>No Hepatitis C</th>
</tr>
</thead>
<tbody>
<tr>
<td>tattoo parlor</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>tattoo no parlor</td>
<td>8</td>
<td>53</td>
</tr>
<tr>
<td>no tattoo</td>
<td>22</td>
<td>491</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>579</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>626</td>
</tr>
</tbody>
</table>

**EXPECTED**

<table>
<thead>
<tr>
<th></th>
<th>Hepatitis C</th>
<th>No Hepatitis C</th>
</tr>
</thead>
<tbody>
<tr>
<td>tattoo parlor</td>
<td>52 47 / 626</td>
<td>52 579 / 626</td>
</tr>
<tr>
<td></td>
<td>(3.904)</td>
<td>(48.096)</td>
</tr>
<tr>
<td>tattoo no parlor</td>
<td>61 47 / 626</td>
<td>61 579 / 626</td>
</tr>
<tr>
<td></td>
<td>(4.580)</td>
<td>(56.420)</td>
</tr>
<tr>
<td>no tattoo</td>
<td>61 513 / 626</td>
<td>513 579 / 626</td>
</tr>
<tr>
<td></td>
<td>(38.516)</td>
<td>(474.484)</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>579</td>
</tr>
<tr>
<td></td>
<td></td>
<td>626</td>
</tr>
</tbody>
</table>

chi-square statistic $= \sum_{\text{cells}} \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = 57.91$

P << 0.0001

<table>
<thead>
<tr>
<th>table of chi - sq:</th>
<th>df</th>
<th>0.0001</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3-1)(2-1) = 2</td>
<td></td>
<td>18.42</td>
</tr>
</tbody>
</table>

**But wait! Are all of the expected counts at least 5?** | No-
Independence/Homogeneity

<table>
<thead>
<tr>
<th></th>
<th>Hepatitis C</th>
<th>No Hepatitis C</th>
</tr>
</thead>
<tbody>
<tr>
<td>tatoo parlor</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>tatoo no parlor</td>
<td>8</td>
<td>53</td>
</tr>
<tr>
<td>no tatoo</td>
<td>22</td>
<td>491</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>513</td>
</tr>
<tr>
<td></td>
<td>113</td>
<td>626</td>
</tr>
</tbody>
</table>

**OBSERVED**

- $\frac{47}{626} = 0.076$
- $\frac{579}{626} = 0.924$

**EXPECTED**

- $\frac{47}{626} = 0.076$
- $\frac{579}{626} = 0.924$

**Chi-square statistic**

- $\chi^2 = \sum \frac{(O - E)^2}{E} = 42.39$

- Degrees of freedom: $(R-1)(C-1) = 1$

- Are all of the expected counts at least 5? **Yes.**