Outline

1. Setting: two categorical variables
   (a) A simple random sample that is then categorized by the two variables
   (b) A simple random sample from each of \( r \) populations, categorized by the second variable
   (c) A randomized comparative experiment with \( r \) treatments and a categorical response variable

2. The null hypothesis

   \( H_0: \) There is no relationship between the two categorical variables

3. The key idea is that if there is no relationship between the categorical variables, we know what the expected counts in each cell of the table should be.

   \[
   \text{expected count} = \frac{\text{row total} \times \text{column total}}{\text{table total}}
   \]

4. Some examples, tables on back.

5. We want a test statistic that is large if the expected counts are very different than the actual counts and small if the expected counts are similar to the actual counts.

6. The test statistic is called the chi-square statistic and is written \( X^2 \).
1. From the Baylor Religion Survey. [http://www.thearda.com](http://www.thearda.com) The questions are: in which region of the country do you live? and do you believe God exists?

<table>
<thead>
<tr>
<th>Region</th>
<th>Absolutely</th>
<th>Probably</th>
<th>Probably not</th>
<th>Absolutely not</th>
</tr>
</thead>
<tbody>
<tr>
<td>East</td>
<td>220</td>
<td>53</td>
<td>32</td>
<td>12</td>
</tr>
<tr>
<td>Midwest</td>
<td>409</td>
<td>55</td>
<td>25</td>
<td>14</td>
</tr>
<tr>
<td>South</td>
<td>373</td>
<td>45</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>West</td>
<td>266</td>
<td>78</td>
<td>43</td>
<td>29</td>
</tr>
</tbody>
</table>

2. The National Survey of Student Engagement. Random samples of first-year students and seniors. The question is: this institution has challenged me to critically evaluate and reconsider values that I have always held.

<table>
<thead>
<tr>
<th>Class</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR</td>
<td>58</td>
<td>121</td>
<td>221</td>
</tr>
<tr>
<td>SR</td>
<td>31</td>
<td>58</td>
<td>202</td>
</tr>
</tbody>
</table>

3. Testing whether a screening test for depression helps primary care physicians to diagnose depression. Will doctors recognize depression status within twelve months of first visit?

<table>
<thead>
<tr>
<th>Group</th>
<th>Recognized</th>
<th>Not recognized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressed, Doctor told</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>Depressed, Doctor not told</td>
<td>11</td>
<td>41</td>
</tr>
<tr>
<td>Not depressed</td>
<td>4</td>
<td>56</td>
</tr>
</tbody>
</table>