Today

- Example 1: Test for a Multinomial Distribution

- Example 2: Two Sample Tests for Location

Example 1: Consider a data set that could have come from a binomial distribution with $n=5$, but may also have come from a hypergeometric or some other distribution.

<table>
<thead>
<tr>
<th>X</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>#obs</td>
<td>9</td>
<td>21</td>
<td>16</td>
<td>10</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Test $H_0$: $X$ is binomial vs. $H_A$: it isn't
So in general we get

\[ G^2 = -2 \log \Lambda = 2 \sum_{i=1}^{n} O_i \log \left( \frac{O_i}{E_i} \right) \]

Page 342-343 demonstrates that this is asymptotically the same as the classic

\[ \sum_{i=1}^{n} \left( \frac{O_i - E_i}{E_i} \right)^2 \]

**Theorem A pg.310:** Under smoothness conditions on the pdf, the null distribution of \(-2\ln \Lambda\) has an approximate chi-square distribution with d.f. = \(\text{dim}\omega - \text{dim}\omega_0\) for large \(n\).
Note: You must be careful what estimates you use for the parameters!!!
What if the data is not normal?

What if we don't know the distribution of the data?

In this case a new paradigm must be used.

Many procedures in statistics are based on replacing the data with ranks.
In the Mann-Whitney-Wilcoxon test all of the data in both groups are replaced by their ranks from smallest to largest.
The population that tends to have larger values will have more large ranks.

Another way to look at it is by considering if the probability that a randomly chosen member of one group is less than a randomly chosen member of another group.
You can show that the number of pairs where \( X < Y \) is equal to the sum of the ranks of \( Y - \frac{m(m+1)}{2} \).