Hypothesis or Hypotheses: that is the question
Elizabeth Hipp, PhD

Conflicts of Interest
None to disclose

Outline
• Null and alternative hypothesis
• Comparing two groups
• Parametric data
• Non-parametric data
• Multiple comparisons
• Conclusion
The null and alternative hypothesis

- A clinical research hypothesis is not the same as a statistical hypothesis
- Formulate a hypothesis
  - H₀ null
  - H₁ alternative
- Compute the appropriate test
- Determine if you reject the null hypothesis
- Assess what this statistical hypothesis means for your clinical research inquiry

The null and alternative hypothesis

- H₀ – the null hypothesis: there is no association
- This will be rejected with varying degrees of confidence
- One sided – testing for the direction of the association
- Two sided – testing for association without direction

What is a p value?

- P value tells us the probability that our conclusion is wrong
- Low p value = unlikely that we wrongly rejected H₀
- Why α=.05 as the standard?
Comparing two groups

• One sample
  – Dr Awesome has seen 12 of the 315 patients she treated using radioactive seeds develop urinary strictures
  – The published average is 5%
  – How does Dr. Awesome compare?

• Two sample
  – Dr Awesome hypothesized that primary tumors and metastatic tumors would be visible on diffusion weighted MRI imaging

Z test and t test

• Continuous, parametric data (normal distribution)
• Observed difference between means

T test

– N<30
– Variance of the population (not just your sample) is unknown
– Paired (e.g before and after)
– Unpaired (independent, unrelated samples)

Unequal variance: non-parametric

• Chi squared
  – Fisher Exact
  – McNemar
• Mann Whitney
• Wilcoxon signed rank
Chi-squared

- Used with categorical data
- Determines whether two (or more) independent populations are homogenous
- whether two (or more) characteristics are independent
- Tests for independence

<table>
<thead>
<tr>
<th></th>
<th>Malignant</th>
<th>Benign</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-enhancing</td>
<td>15</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Enhancing</td>
<td>7</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>47</td>
<td>69</td>
</tr>
</tbody>
</table>

Unequal variance: non-parametric

- Mann Whitney
  - Compares two independent population distributions
  - Does not assume a distribution
- Wilcoxon Signed rank
  - An alternative to the paired t test
  - Assumes distribution is symmetric with respect to its median

Multiple comparisons: ANOVA

- Compares the means of two or more independent groups
- Assumes a normal distribution
Multiple comparisons: Bonferroni

- Adjusts the threshold for significance based on the number of variables being considered
- Reducing the level of significance also reduces the power of the test
- Most conservative approach
- Alternatives: Holm, Sidak

Erroneous relationships due to data mining

- Follow up imaging for two arms of a study (control and treated)
- Initially you only look at diffusion coefficient, then want to compare other image metrics

<table>
<thead>
<tr>
<th>Number of independent variables</th>
<th>Probability of erroneously declaring one significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>13</td>
<td>50%</td>
</tr>
<tr>
<td>30</td>
<td>80%</td>
</tr>
</tbody>
</table>

Conclusion

- Formally write out your null hypothesis before you run any tests
- How many hypotheses do you have? Are they related?
- Now choose your tests
- If you start analyzing your data and running tests in excel, take a moment and go back and write it out