Skill Builder Series
Non parametric Methods
November 12, 2015

Patrick Dunn and Kim Palermo-Kielb
Welcome to the Session!

House Keeping in Adobe Connect

- Presenters can share a variety of content with participants in the share pod (screen, document, whiteboard)
- Other pods may also be visible, such as those for closed captioning and poll questions
- Participants can utilize the Q&A pod to ask questions or they can raise their hand to be unmuted
- Select the Help button in the upper right corner for technical support
Housekeeping

Live participants & archived viewers:
• Polls, files, and links are interactive. Different types of interactivity are used at each session through the day.

Live participants:
• Use the chat box to ask questions.

Archived viewers:
• Send questions to ASCTutoring@Waldenu.edu

Help & troubleshooting:
• Please type your question in the “chat” area
This session is being recorded and will be available for viewing on our website.

http://academicguides.waldenu.edu/ASC/skillbuilder
Skill Builder Series
Non parametric Methods
November 12, 2015

Patrick Dunn and Kim Palermo-Kielb
Agenda

• Review of previous skills builders
• Overview of non-parametric methods
• Demonstration of multivariate tests in SPSS and Excel
• Questions
• Additional Resources
Previous skill builders

• General concepts
  – Normal distribution
  – Variance
  – Null and alternative hypothesis

• Descriptive statistics
  – Means
  – Variation

• T tests and ANOVA

• Multivariate methods

• Power and sample size
What are non-parametric tests and how are they used?

- When you are analyzing categorical variables
- When the distribution is not normal

- Chi-Square
- Mann-Whitney
- Kruskal-Wallis
- Wilcoxon Signed Rank Test
- Sign Test
- Survival analysis
## Parameter or Statistic

### Parameter
- Based on the population
- An absolute number
- Example: Data in the form of birth certificates is collected on every person born in the U.S. Therefore, Natality Records are *population level data* and all information computed using that data are *parameters*.

### Statistic
- Based on a sample of the population
- An estimate of the parameter it represents
- Example: Data in National Surveys such as NHANES or BRFSS is collected from a *random sampling of the U.S. population*. All information computed using that data are *statistics*. 
Purpose of Statistical Test

Quantitative variables

Testing Differences

Description

Continuous: mean and standard deviation; Ordinal: median; Categorical: frequency

Number of groups?

One group mean compared to a test value

Two groups

More than two groups

One group: observed frequency compared to a proposed distribution

Testing Relationships

Categorical variables

Quantitative variables

Related samples

Independent samples

One-sample t-test

Related measures ANOVA

Independent samples t-test

One-way ANOVA

Paired samples t-test

Wilcoxon matched-pairs test

Mann-Whitney U test

Multiple Logistic Regression

Pearson's correlation coefficient

Pearson's rho

Multiple Linear Regression

Chi-square goodness-of-fit test

Chi-square test for association

Two or more independent variables, one dichotomous dependent variable

Two or more independent variables, one dependent variable
Distributions
Distributions

[Histogram showing frequency distribution of age categories 1 to 4.]

- Category 1
- Category 2
- Category 3
- Category 4

[Bar chart comparing categories 1 to 4 with different heights.]

- Category 1
- Category 2
- Category 3
- Category 4
Frequency comparisons

• Chi-Square
  – Odds ratio
  – Relative risk
The Chi-Square Statistic

- A very commonly used test statistic in Public Health
- It compares the results OBSERVED to those EXPECTED based on
  - Known theories
  - Hypotheses
  - Comparison Groups
- Question: Do the results we obtained differ significantly from those expected?
Chi-Square and 2x2 Tables

Here is a 2x2 table:

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Exposure</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Yes</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>No</td>
<td>n</td>
<td>N</td>
</tr>
</tbody>
</table>

• Suppose you want to determine if the two groups represented in the table are significantly different
• To do this you use the Chi Square Statistic:

\[
\chi^2 = \frac{(ad - bc)^2 \times N}{(a+c)(b+d)(a+b)(c+d)}
\]
Odds and Risk Ratios

• **Odds Ratio (OR):** The odds of exposure in the diseased group divided by the odds of exposure in the non-diseased group.
  – Odds ratio=OR

• **Risk Ratio (Relative Risk (RR)):** The ratio of the risk in the exposed group to the risk in the unexposed group.
  – Risk=RR

• **Both use 2x2 tables**
2 by 2 Contingency Table

- Comparison data between two groups
  - Used to show results of 2 groups
  - Column 1 is data from group 1
  - Column 2 is data from group 2
  - Row 1 is positive outcomes
  - Row 2 is negative outcomes
  - Each row and column is added
  - Used in many applications

<table>
<thead>
<tr>
<th>Exposure to variable of interest</th>
<th>Existence of Disease designated as case (yes disease) or control (no)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case</td>
<td>Control</td>
</tr>
<tr>
<td>Yes exposed</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>No not exposed</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>Total</td>
<td>All Cases (a + c)</td>
<td>All Controls (b + d)</td>
</tr>
</tbody>
</table>
Calculation of OR and RR

• OR = (odds for observing the outcome in the exposed group)/(odds for observing the outcome in the unexposed group) = (ad)/(bc)

• RR = (risk in the exposed group)/(risk in the unexposed group) = [a/(a + b)]/[c/(c + d)]

• OR used in retrospective studies; but can be used in prospective studies as well

• RR used in prospective studies only (RR uses incidence)
Independent Samples

• Mann-Whitney U – 2 samples
• Kruskal-Wallis – More than 2 samples
Mann-Whitney U test
Wilcoxon Rank-Sum test

• 2 independent samples
• Non parametric counterpart to the independent sample t test
• Comparing the medians
### Example

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Order

<table>
<thead>
<tr>
<th>Placebo</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

#### Rank

<table>
<thead>
<tr>
<th>Placebo</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>
$U = n_1 n_2 + n_1(n_1+1)/2 - R$

$U = n_1 n_2 + n_2(n_2+1)/2 - R$

$U = 5(6) + 5(5+1)/2 - 37 = 3$

$U = 5(6) + 5(5+1)/2 - 18 = 22$

Reject null if $U<2$
Kruskal-Wallis

• Comparing more than 2 independent samples
• Non parametric counterpart to ANOVA

\[ H = \frac{12}{N(N+1)} \times \text{Sum} \left( \frac{\text{Sum of ranks}^2}{N} \right) - 3(N+1) \]
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

Sum of ranks: 46, 62, 24, 78
\[ H = \left( \frac{12}{N(N+1)} \right) \times \text{Sum (Sum of ranks}^2/N) \right) - 3(N+1) \]

\[ 12/20 \times 21 \times \left( \frac{46^2}{5} + \frac{65^2}{5} + \frac{24^2}{5} + \frac{78^2}{5} \right) - 3(21) = 9.11 \]
Matched samples

- Wilcoxon Signed Rank Test
- Sign Test
## Example

<table>
<thead>
<tr>
<th>Pre</th>
<th>Post</th>
<th>Sign</th>
<th>Wilcoxon</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>75</td>
<td>10</td>
<td>-5</td>
</tr>
<tr>
<td>70</td>
<td>50</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>40</td>
<td>50</td>
<td>25</td>
<td>-10</td>
</tr>
<tr>
<td>65</td>
<td>40</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>80</td>
<td>20</td>
<td>60</td>
<td>-10</td>
</tr>
<tr>
<td>75</td>
<td>65</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>55</td>
<td>40</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-5</td>
<td>25</td>
</tr>
</tbody>
</table>

**Result:** Fail to reject

**Result:** Reject the Null
Time/Event analysis

- Comparing events over time.
- Non parametric
- Does not require a normal distribution
Questions
More questions? Stay Informed!

Current Walden students, e-mail our tutoring team at:
ASCTutoring@Waldenu.edu

Subscribe to our Facebook & Twitter channels to keep up-to-date on new information, ask questions and share your knowledge.

Subscribe to our YouTube channel (free!) to have instant access to new tutorials and webinar archives.

Visit our website for tutorials, event schedules, tutoring services, courses and workshops.