Lab 4: Alpha and Kappa

Today's Activities
- Consider Alpha
- Consider Kappa
- Homework and Media Write-Up

Reliability
- Reliability refers to consistency
- Types of reliability estimates
  - Test-retest reliability
  - Internal consistency
  - Interobserver (or interrater) reliability
- Alpha measures the consistency of responses across items
Calculating Alpha by Hand

\[ \alpha = \frac{k \cdot \bar{r}_{ij}}{1 + (k - 1) \bar{r}_{ij}} \]

Calculating Alpha of the Extraversion Scale by Hand

- \( k \): number of items
  - Extraversion scale: \( k = 8 \)
- \( \bar{r}_{ij} \): average inter-item correlation
  - Extraversion scale: \( \bar{r}_{ij} = .46 \)

\[ \alpha = \frac{k \cdot \bar{r}_{ij}}{1 + (k - 1) \bar{r}_{ij}} \]

- How large would alpha be if the scale had 16 items?

Calculating Alpha of the Extraversion Scale for the Time 2 BFI

- SPSS: Analyze \( \rightarrow \) Scale \( \rightarrow \) Reliability Analysis
- Put the following items in the items box: 1, 11, 16, 26, 36, 6R, 21R, 31R
- Click on Statistics
  - Inter-Item: Correlations
  - Summaries: Correlations
    - Then hit Continue
- Make sure you use the Time 2 reverse scored items!
- Then click OK
Questions to Answer from Output

1. What was the average inter-item correlation?
2. How many items are on the scale?
3. What was the standardized alpha?
   - Reliability rules of thumb reminder:
     - above .8 indicates high to moderate reliability
     - .7 to .8 indicates moderate reliability
     - under .7 indicates low reliability

Calculating Alphas for the other four scales (A, C, N, and O)

- Conscientiousness: 3, 13, 28, 33, 38, 8R, 18R, 23R, 43R
- Neuroticism: 4, 14, 19, 29, 39, 9R, 24R, 34R
- Openness: 5, 10, 15, 20, 25, 30, 40, 44, 35R, 41R

What happens if you forget to reverse score the items?

- Calculate alpha for the Extraversion scale again, but now use the following items:
  1, 6, 11, 16, 21, 26, 31, 36
- What is the new alpha?
- How can you tell that you did something wrong?
What happens if you calculate alpha of two scales at the same time?

- Calculate alpha for the following eight items:
  - 6R, 16, 26, 36 (Extraversion items)
  - 3, 13, 28, 33 (Conscientiousness items)

- Can you interpret your result? What might clue you into your mistake?

- Remember: Alpha does not indicate unidimensionality!

Agreement Between Observers

Study of parallel play (Bakeman & Brownlee, 1980)

- 3-year olds were observed during indoor free play
- Observers coded children’s play as either Unoccupied (Un.), Solitary (Sol.), Together (Tog.), Parallel (Par.), Group (Gr.)

<table>
<thead>
<tr>
<th>First Observer</th>
<th>Second Observer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Un.</td>
<td>7</td>
</tr>
<tr>
<td>Sol.</td>
<td>1</td>
</tr>
<tr>
<td>Tog.</td>
<td>1</td>
</tr>
<tr>
<td>Par.</td>
<td>0</td>
</tr>
<tr>
<td>Gr.</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
</tr>
</tbody>
</table>
Percentage of Agreement

\[ P_A = \frac{N_A}{N_A + N_D} \times 100 \]

\( P_A \) = percentage of agreement  
\( N_A \) = number of agreements  
\( N_D \) = number of disagreements

Percentage of Agreement

- \( N_A \) (number of agreements) = 87  
  \( (7 + 24 + 17 + 25 + 14) \)
- \( N_D \) (number of disagreements) = 13  
  \( (1 + 1 + 1 + 2 + 2 + 3 + 1 + 1) \)
- \( P_A \) (percentage of agreement) = 87%  
  \( \frac{87}{(87 + 13)} \times 100 \)

Percentage of Agreement: Problems

- By chance alone, we expect some agreement between observers
- We need an agreement statistic that corrects for chance agreement
- Enter: Cohen’s kappa (Cohen, 1960)
Cohen’s Kappa

\[ \kappa = \frac{P_{\text{obs}} - P_{\text{chance}}}{1 - P_{\text{chance}}} \]

\( \kappa \) = Kappa
\( P_{\text{obs}} \) = Proportion of agreement actually observed
\( P_{\text{chance}} \) = Proportion of agreement expected by chance

Computing Kappa I

\( P_{\text{obs}} \) = Proportion of agreement observed
- Sum up the numbers representing agreement (in the diagonal) and divide by the total number of observations
- In our example:
  \( \frac{7 + 24 + 17 + 25 + 14}{100} = .87 \)

Computing Kappa II

\( P_{\text{chance}} \) = Proportion of agreement expected by chance

\[ P_{\text{chance}} = \frac{\sum_{i=1}^{k} x_{+i} \cdot x_{i+}}{N^2} \]

So what does this mean???
Computing Kappa III

\( P_{\text{chance}} = \) Proportion of agreement expected by chance

- Multiply the first column total by the first row total. In our example: \( 9 \times 8 = 72 \)
- Multiply the second column by the second row total. In our example: \( 25 \times 25 = 625 \)
- Third column by third row total: \( 21 \times 23 = 483 \)
- Fourth column by fourth row total: \( 28 \times 29 = 812 \)
- Fifth column by fifth row total: \( 17 \times 15 = 255 \)

Computing Kappa IV

\( P_{\text{chance}} = \) Proportion of agreement expected by chance

- Add the resulting numbers. In our example: \( 72 + 625 + 483 + 812 + 255 = 2,247 \)
- Divide this sum by the total number of observations squared. In our example: \( 100^2 = 10,000 \)
- \( 2,247 / 10,000 = .2247 \)
- \( P_{\text{chance}} = .2247 \)
(We expect an agreement of 22.5% by chance)

Computing Kappa V

\[ \kappa = \frac{P_{\text{obs}} - P_{\text{chance}}}{1 - P_{\text{chance}}} \]
\[ \kappa = \frac{.87 - .2247}{1 - .2247} = .8323 \]

How does this compare to the percentage of agreement?
Interpreting Kappa (Fleiss, 1981)

- .40 - .60: fair
- .60 - .75: good
- Over .75: excellent

Homework Questions 1 & 2

1. Using SPSS, report the standardized alphas for all five BFI scales at Time 2. Also report the average inter-item correlation and number of items. (1 point)

2. Imagine that the average inter-item correlation for a scale is .35. Compute (by hand) the standardized alphas for tests with the following values of k: 5, 10, 15, 20, 25. (1 point)

Question 3: Kappa Calculation (3 pts)

3. Compute kappa for the following confusion matrix. In your answer, show each step of the calculation. How does kappa differ from the percentage of agreement? Why does it differ from the percentage of agreement?
Confusion Matrix

Two researchers observed a child on a playground. They coded 100 30-second intervals for fighting or no fighting.

<table>
<thead>
<tr>
<th></th>
<th>1. Observer</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Fight</td>
<td>Fight</td>
</tr>
<tr>
<td>2. Obs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Fight</td>
<td>80</td>
<td>5</td>
</tr>
<tr>
<td>Fight</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>82</td>
<td>18</td>
</tr>
</tbody>
</table>

New Popular Press Article (5 points)

- You have now read a piece of original research and a description of this research in the popular media. Last week you compared a piece of original research and a description of that research in the popular media. Hopefully you applied a critical eye and detected errors of omission or questionable interpretations in the popular media description. Now the task is to take those criticisms to heart and rewrite a short (i.e. 500 or 600 word) article for the popular press about the original research.
- Make sure you explain the importance of the topic or why readers should find it interesting (1 point)
- Provide basic details of the study in everyday language so that a wide variety of readers will understand the research method (2 points)
- Describe the main findings and the implications (1 point)
- Explain any limitations in clear English and provide a clue as to future research (1 point)
- The goal here is to explain the research in an engaging fashion that is clear, accessible, and accurate.