Lecture 22: Non-Parametric Tests

Parametric vs. Non-parametric Statistics

**Parametric statistics**
- Most common type of inferential statistics
  - \( r, t, F \)
- Make strong assumptions about the population
  - Mathematically fully described, except for a few unknown parameters
  - Powerful, but limited to situations consistent with assumptions

When parametric statistics fail
- Assumptions not met
- Ordinal data: Assumptions not meaningful

**Non-parametric statistics**
- Alternatives to parametric statistics
- "Naive" approach: Far fewer assumptions about data
- Work in wider variety of situations
- Not as powerful as parametric statistics (when applicable)

Assumption Violations

**Parametric statistics** work only if data obey certain properties

**Normality**
- Shape of population distribution
- Determines shape of sampling distributions
- Tells how likely extreme results should be; critical for correct p-values
- More important with small sample sizes (Central Limit Theorem)

**Homogeneity of variance**
- Variance of groups is equal (t-test or ANOVA)
- Variance from regression line does not depend on values of predictors

**Linear relationships**
- Pearson correlation cannot recognize nonlinear relationships

If these assumptions are true:
- Population is almost fully described *in advance*
  - Goal is simply to estimate a few unknown parameters

If assumptions violated:
- Parametric statistics will not give correct answer
  - Need more conservative and flexible approach
Ordinal Data
Some variables have ordered values but are not as well defined as interval/ratio variables
Preferences
Rankings
Nonlinear measures, e.g. money as indicator of value
Can’t do statistics based on differences of scores
Mean, variance, \( r, t, F \)
More structure than nominal data
Scores are ordered
Chi-square goodness of fit ignores this structure
Want to answer same types of questions as with interval data, but without parametric statistics
Are variables correlated?
Do central tendencies differ?

Non-parametric Tests
Can use without parametric assumptions and with ordinal data
Basic idea
Convert raw scores to ranks
Do statistics on the ranks
Answer similar questions as parametric tests

<table>
<thead>
<tr>
<th>Parametric Test</th>
<th>Non-parametric Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson correlation</td>
<td>Spearman correlation</td>
</tr>
<tr>
<td>Independent-samples t-test</td>
<td>Mann-Whitney</td>
</tr>
<tr>
<td>Single- or paired-samples t-test</td>
<td>Wilcoxon</td>
</tr>
<tr>
<td>Simple ANOVA</td>
<td>Kruskal-Wallis</td>
</tr>
<tr>
<td>Repeated-measures ANOVA</td>
<td>Friedman</td>
</tr>
</tbody>
</table>

Your job: Understand what each is used for and in what situations

Spearman Correlation
Alternative to Pearson correlation
Produces correlation between -1 and 1
Convert data on each variable to ranks
For each subject, find rank on \( X \) and rank on \( Y \) within sample
Compute Pearson correlation from ranks
Works for
Ordinal data
Monotonic nonlinear relationships (consistently increasing or decreasing)
Mann-Whitney Test
Alternative to independent-samples t-test
Do two groups differ?
Combine groups and rank-order all scores
If groups differ, high ranks should be mostly in one group and low ranks in the other
Test statistic ($U$) measures how well the groups’ ranks are separated
Compare $U$ to its sampling distribution
  Is it smaller than expected by chance?
  Compute p-value in usual way
Works for
  Ordinal data
  Non-normal populations and small sample sizes

Wilcoxon Test
Alternative to single- or paired-samples t-test
Does median differ from $\mu_0$?
  Does median difference score differ from 0?
Subtract $\mu_0$ from all scores
  Can skip this step for paired samples or if $\mu_0 = 0$
Rank-order the absolute values
Sum the ranks separately for positive and negative difference scores
  If $\mu > \mu_0$, positive scores should be larger
  If $\mu < \mu_0$, negative scores should be larger
If sums of ranks are more different than likely by chance, reject $H_0$
Works for
  Ordinal data
  Non-normal populations and small sample sizes

Kruskal-Wallis Test
Alternative to simple ANOVA
Do groups differ?
Extends Mann-Whitney Test
  Combine groups and rank-order all scores
  Sum ranks in each group
If groups differ, then their sums of ranks should differ
Test statistic ($H$) essentially measures variance of sums of ranks
  If $H$ is larger than likely by chance, reject null hypothesis that populations are equal
Works for
  Ordinal data
  Non-normal populations and small sample sizes
**Friedman Test**
Alternative to repeated-measures ANOVA
Do measurements differ?
Look at each subject separately and rank-order his/her scores
Best to worst for that subject, or favorite to least favorite
For each measurement, sum ranks from all subjects
If measurements differ, then their sums of ranks should differ
Test statistic ($\chi^2$) essentially measures variance of sums of ranks
If larger than likely by chance, reject $H_0$
Works for
Ordinal data
Non-normal populations and small sample sizes

**Summary**

<table>
<thead>
<tr>
<th>Test</th>
<th>Replaces</th>
<th>When Useful</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman correlation</td>
<td>Pearson correlation</td>
<td>• Ordinal data</td>
<td>Rank each variable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Nonlinear relationships</td>
<td></td>
</tr>
<tr>
<td>Mann-Whitney t-test</td>
<td>Independent-samples t-test</td>
<td>• Ordinal Data</td>
<td>Rank all scores</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Non-normal + small samples</td>
<td></td>
</tr>
<tr>
<td>Wilcoxon t-test</td>
<td>Single- or paired-samples</td>
<td>• Ordinal Data</td>
<td>Rank absolute values</td>
</tr>
<tr>
<td></td>
<td>t-test</td>
<td>• Non-normal + small samples</td>
<td></td>
</tr>
<tr>
<td>Kruskal-Wallis t-test</td>
<td>Simple ANOVA</td>
<td>• Ordinal Data</td>
<td>Rank all scores</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Non-normal + small samples</td>
<td></td>
</tr>
<tr>
<td>Friedman ANOVA</td>
<td>Repeated-measures ANOVA</td>
<td>• Ordinal Data</td>
<td>Rank all scores</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Non-normal + small samples</td>
<td>for each subject</td>
</tr>
</tbody>
</table>