Tuesday PM

- Presentation of AM results
- What are nonparametric tests?
- Nonparametric tests for central tendency
  - Mann-Whitney U test (aka Wilcoxon rank-sum test)
  - Sign test, Wilcoxon signed-ranks test
  - Nonparametric ANOVA
- Chi-squared

Nonparametric tests

- As mentioned on Monday, t-tests and ANOVAs are *parametric*: they make assumptions about the distribution of populations (typically, normal distributions)
- *Nonparametric* tests don't require normality, but…
  - They are less powerful (require more subjects)
  - They test slightly different null hypotheses

Mann-Whitney U Test

- Goal: Determine whether two groups differ on a variable. "Nonparametric indepedent t-test"
- Equivalent to the *Wilcoxon rank-sum test*
- Works by ranking all scores across groups, and computing the sum of the ranks within each group. Those rank-sums should be similar if the distributions are similar in each group.
- U or W is reported, with significance.

Mann-Whitney U in SPSS

- Analyze…Nonparametric tests…2 independent samples
- Enter test (dependent) variable and grouping variable
- Do Asian Pacific countries have significant larger populations than Eastern European countries? (t-test might be too sensitive to skew in distribution):

	Test Statistics
Mann-Whitney U	51.000
Wilcoxon W	156.000
Z	-2.699
Asymp. Sig. (2-tailed)	.007
Exact Sig. [2*(1-tailed Sig.)]	.006

- AP countries have significantly larger populations than EE (Mann-Whitney U=51, p<.06)

Sign test

- Goal: Determine whether a variable, measured twice, differs between measurements. "Nonparametric paired t-test"
- Works by examining the difference between each pair of scores, and categorizing it as positive, negative, or zero.
- If the measurements differ, there should be significantly more positive or negative differences.

Sign test

- Analyze…Nonparametric tests…2 related samples
- Enter pairs of variables

Avg male LE - Avg female LE in 109 countries:

Negative Differences	107
Positive Differences	1
Ties	1
Total	109

Test Statistics

Z	-10.104
Asymp. Sig. (2-tailed)	.000

- Female life expectancy exceeds male life expectancy in nearly all countries (sign test, Z=-10.1, p < .05).

Wilcoxon signed-ranks test

- Goal: Determine whether a variable, measured twice, differs between measurements. "Nonparametric paired t-test"
- Works by ranking absolute differences between measurements, summing them up for positive and negative differences, and comparing the sums.
- Unlike sign test, gives more weight to pairs that show large differences than to pairs that show small differences.

Wilcoxon signed-ranks test in SPSS

- Analyze…Nonparametric tests…2 related samples
- Enter pairs of variables

Ranks: Avg male LE - Avg female LE

	N	Mean Rank	Sum of Ranks
Negative Ranks	107	54.98	5883.00
Positive Ranks	1	3.00	3.00
Ties	1		

Test Statistics

Z	-9.039
Asymp. Sig. (2-tailed)	.000

- Female LE exceeds male LE across countries (Wilcoxon signed-ranks test, Z=-9.0, p < .05).

Nonparametric ANOVA

- SPSS also offers nonparametric tests for:
  - 3+ independent groups (Kruskal-Wallis H) "Nonparametric one-way between-subject ANOVA"
  - 3+ repeated measures of same variable (Friedman's test) "Nonparametric one-way within-subject ANOVA"
  - 3+ measures by different raters (Kendall's W)

Chi-squared

- χ^2 is one of the most useful nonparametric statistics. It can be applied to many problems:
  - Is an observed distribution of responses different from an expected on?
  - Are there independent or interactive effects of two categorical variables on a distribution of responses?
  - Are there differences in two related proportions (e.g. proportion of students scoring >90% before and after an educational intervention)?

One-way χ^2

- Given:
  - a set of observed responses divided into categories
  - a set of expected responses divided into categories
    (often a null hypothesis of 'equal distribution')
- Goal: Determine if the observed distribution is significantly different than the expected distribution.

One-way χ^2: example

- Students are asked to choose if they prefer exams in the morning or afternoon. Is there a significant preference?

	Prefer AM	Prefer PM	Total
Observed	39	21	60
Expected	30	30	60

- $\chi^2 = \Sigma(O-E)^2/E = (39-30)^2/30 + (21-30)^2/30 = 5.4$
- Significantly more students prefer morning to afternoon exams ($\chi^2(1)=5.4$, $p<.05$)

One-way χ^2 in SPSS

- Nonparametric tests…Chi-square
- Enter test variable and set expected values if not equally distribute across categories
- Example: We are designing an evaluation in which residents are given a case and asked to make a yes or no decision about performing an LP. We don't expect the residents, on average, to know the right answer, so we expect equal numbers to say yes and no. Did that happen?

One-way χ^2 output

LP Decision

	Observed N	Expected N	Residual
No	28	20.0	8.0
Yes	12	20.0	-8.0
Total	40		

Test Statistics

Chi-Square	6.400
df	1
Asymp. Sig.	.011

- Significantly more residents believed they should not do the LP ($\chi^2(1)=6.4$, $p<.05$)

Two-way χ^2

- Given data in a contingency table (relating responses to two categorical variables)

	Prefer AM	Prefer PM	Total
M1	25	25	50
M2	15	35	50
Total	40	60	100

- Are the effects of the two categorical variables independent or related?
- Same algorithm as one-way (compute expected frequencies based on marginal totals)

Two-way χ^2 in SPSS

- A second case is developed about use of CT (and tested on different residents). Are the distribution of responses to the CT and LP cases the same?
- Analyze…Descriptive statistics…Crosstabs
- Enter a row and column variable to define the contingency table.
- Hit "Options" and check the box for chi-square

Two-way χ^2 output

Form * Prior Decision Crosstabulation

	Prior Decision No	Prior Decision Yes	Total
CT	25	20	45
LP	28	12	40
Total	53	32	85

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.882	1	.174
Continuity Correction	1.317	1	.251
Likelihood Ratio	1.897	1	.168

- The distributions of responses to the two items were not significantly different.

McNemar's test of correlated proportions

- Given two related proportions, is one significantly higher than the other?
- Example: 85 residents answered the LP case, and were then given a journal abstract that did not support doing LP in the case, and were asked to answer the case again. Did significant fewer do the LP after the evidence?

McNemar's test in SPSS

- Analyze…Nonparametric tests…2 related samples
- Enter variable pair and select McNemar checkbox

```
                       Prior Decision
Post Decision          0       1
                0      50      12
                1      3       20

                          Test Statistics
N                              85
Exact Sig. (2-tailed)  .035
```

- Residents were significantly less likely to order the LP after reading the evidence (McNemar's test, $p < 0.05$)

χ^2 data considerations

- Observations are assumed to be independent (except in McNemar's test)
- χ^2 is not reliable if the expected cell frequencies are smaller than about 5.
- A "correction for continuity" may be applied when expected frequencies are small, but there is argument about appropriateness (see Howell, p 146).

Tuesday PM assignment

- Using the clerksp data set, examine the i1/i1post items (self-rated differential diagnosis skills):
  - Are post-test scores higher than pre-test? Test this question using a paired t-test, a sign test, and the Wilcoxon signed-ranks test. How do the results differ?
  - Create a new variable, nastydoc, coded "1" for clerks whose pre-test i1 rating is higher than their pre-test i15 (expresses caring) rating, and "0" for others. Test whether more than half the clerks are nastydocs using one-way χ^2
  - Create a new variable, IM, coded "1" for clerks whose 1st choice residency before the clerkship was internal medicine, and "0" for all others. Is there a relationship between IM and nastydoc? Test using two-way χ^2 and interpret.