Review

Sample Comparisons

What are we testing?

- Single samples
  - Single sample mean to a population mean posited by the null hypothesis
- Paired sample t
  - Single sample of difference scores → mean difference score compared to a population mean difference score
- Independent samples
  - Compare the difference between sample means to the population mean difference

Going further: Effect size d

- Effect size
- Standardized mean difference
  - Cohen's d, Glass's delta and variants
- Applicable to both independent and dependent though caution must be applied to the latter
  - There is choice in standardizer

Going further: Equivalence testing

- What if your question regards equivalence rather than a difference between scores?
  - Example: want to know if two counseling approaches are similarly efficacious
- Just do a t-test, and if non-significant conclude they are the same
- Wrong! It would be logically incorrect.
  - Can you think of why?

- Equivalence testing has effect size as a starting point
- Choose a meaningful effect (Δ) such that any difference that size or greater could not be construed as indicating equivalent groups
- However, to show equivalence you must show that the small difference seen did not occur by chance

- Equivalence testing actually tests both difference and equivalence at the same time, places focus on effect size rather than statistical significance, and allow ambiguous outcomes to not be overstated in terms of difference or equivalence if the evidence just simply isn't there

Confidence Intervals

- The mean or difference between means is a point estimate
- We should also desire an interval estimate reflective of the variability and uncertainty of our measurement seen in the data

$$Limits = \bar{X} \pm t_{.95}(s_{\bar{X}})$$

Where

$\bar{X}$ = sample mean

$t_{.95}$ = t critical value

$s_{\bar{X}}$ = standard error of the mean

What confidence intervals really tell you

- A 95% confidence interval means that:
- 95% of the confidence intervals calculated on repeated sampling of the same population will contain μ, and this is one of them
- Note that the population value does not vary in this scenario i.e. it's not a 95% chance that it falls in that specific interval
- There are an infinite number of 95% CIs that are consistent with a fixed population parameter
- In other words, the CI attempts to capture the mean, and the one you have is just one of many that are correct

- http://www.ruf.rice.edu/~lane/stat_sim/conf_interval/index.html

Confidence Intervals

- As suggested by many leading quantitative psych guys, using confidence intervals in lieu of emphasizing p-values may make for a better approach for evaluating group differences etc.
- Confidence intervals can provide all the relevant NHST info as well as a means of graphically displaying statically significant differences
- One approach is to simply see if our interval for the statistic of interest contains the H_0 value.
- If not, reject H_0

Initial Summary

- T-tests provide minimal evidence that the difference seen is unlikely assuming a nil or other proposed size of difference
- Effect sizes provide more information regarding the actual effect, and interval estimates of them give a sense of the true effect in the population
- Confidence intervals are an important component statistical analysis and should always be reported, and can be used in lieu of observed p-values[1]
- Non-significance on a test of difference does not allow us to assume equivalence
- Methods exist to test the group equivalency which force one to think about the meaningfulness of the effect first and foremost, and should be implemented whenever that is the true goal of the research question but is always viable since it provides the regular t-test information also

One-way ANOVA

- One-way ANOVA extends the dependent and independent samples t-tests to include more than two groups
  - Between-groups design
  - Repeated measures design
- As we will see later it can have more than one factor
  - More than
  - Mixed
- The test tells you that one of the groups is statistically distinct form the other, and the primary effect size is eta-squared (η^2) and it's bias-corrected counterpart omega-squared (ω^2) which tells you the amount of variance accounted for in the DV by the grouping factor
  - Yes it's R^2
- We will review more later, but suffice it to say that a one-way analysis serves as a starting point, it is never a sufficient analysis by itself

Categorical Data

- If you have *only* categorical data for your predictor and outcome[1], one could perform a chi-square analysis
- It, like the one way design can only tell you there is some relationship, careful inspection will be required for more detail

What do undergrad stat students do with their free time?	Updating their Myspace/Facebook or whatever blog thing whose contents will get them not hired/ fired from some job in the future	Talking on cell phone about their drama loudly enough that now total strangers know how the 'tests' turned out	Texting instead of just calling the person and actually talking to them	Staring at the ceiling	Total
Males (E)	30 (25)	40 (35)	20 (30)	10 (10)	100
Females (E)	20 (25)	30 (35)	40 (30)	10 (10)	100
	50	70	60	20	200

$$X^2 = \sum \frac{(O-E)^2}{E}$$

df = (R-1)(C-1)

Summary

- Understanding group comparisons on a this basic level will help you understand Analysis of Variance techniques later on
  - They allow for more than two independent or dependent samples
  - They allow for the testing of interactions
  - They allow for mixing independent and dependent sample factors
- Even when using those, the specifics typically only involve specialized t-tests, and in that sense will not provide a lot that is new conceptually
- The concerns about effect size etc. will all be maintained