Lecture 6: Analysis of Independent Groups Design

III ANALYSIS
A. Descriptive Statistics

- Analysis of Independent Groups Design involves two types of statistical concepts and procedures.
  - Descriptive statistics are used to summarize a data set, to estimate population parameters, and to reduce a large body of raw information (observations) to a smaller body of summarized information
  - Inferential statistics are used to make some judgments about the population of interest based upon the sample statistics.

III ANALYSIS
A. Descriptive Statistics

- Let's discuss a particular study
  - College students are told of Jane and Bob who were resting by a tree on campus.
    - They were told how Bob climbed the tree and played on one of the branches while Jane was watching. Also how Bob sidled up to Jane and she scampered away.
  - Students were then asked to judge whether "Bob was romantically interested in Jane" on a 7-point Likert scale
    - "Very Strong Agree" to "Strongly Disagree"

III ANALYSIS
A. Descriptive Statistics

- There was one IV and one DV.
  - Fifteen students were randomly assigned to the "student" condition, where Bob and Jane were described as students who lived on campus.
  - Another fifteen students were randomly assigned to the "squirrel" condition, where Bob and Jane were described as squirrels who lived on campus.
  - It was hypothesized that the explanation would be judged more acceptable in the Human than the Squirrel condition.

III ANALYSIS
A. Descriptive Statistics

LIKERT SCALE

LIKERT SCALE	Frequency
1. I very strongly agree with the explanation.	1
2. I strongly agree with the explanation.	3
3. I agree with the explanation.	5
4. I neither agree nor disagree with the explanation.	12
5. I disagree with the explanation.	5
6. I strongly disagree with the explanation.	3
7. I very strongly disagree with the explanation.	1

III ANALYSIS
A. Descriptive Statistics

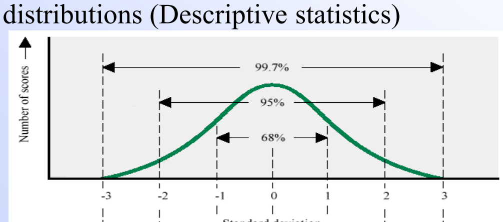
- Compute the mean of the sample
  - A measure of central tendency found by computing the average observation.
  - $M = \Sigma X / n$
  - $120/30 = 4$

III ANALYSIS
A. Descriptive Statistics

- Compute the Standard Deviation of the sample:
  - A measure of variability found by computing the average distance from the mean of the observations.
  - $\sqrt{\Sigma(X-M)^2/(n-1)}$
  - $\sqrt{52/29} = 1.34$
  - $sd^2 = s$ (variance) $= 1.34^2 = 1.80$

III ANALYSIS
A. Descriptive Statistics

- Having a mean (M) and a standard deviation (sd) of a sample is a powerful combination of numbers with which you can figure out a lot!
- You can figure out properties of the sampled distributions (Descriptive statistics)



III ANALYSIS
A. Descriptive Statistics

- But you can do more with these values than that!
- You can use the mean and standard deviation to infer characteristics of the population.
  - **Sample:** a subset of the subjects chosen from the population by a specified procedure
  - **Population**: the subjects, items, elements or units in a defined group.
- Moving from a sample to a population is the goal of **Inferential Statistics**

III ANALYSIS
B. Inferential Statistics

- The descriptive statistics have a corresponding inferential one.

SAMPLE	POPULATION
Mean (M)	Mu (μ)
Standard Deviation (SD)	Sigma (σ)

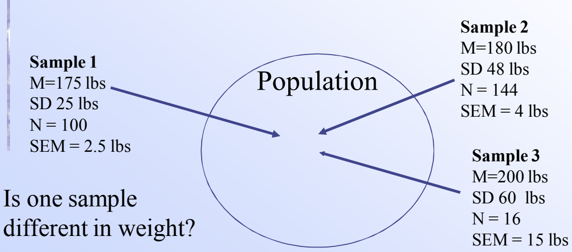
  - Knowing our sample's M and sd, we can make inferences about the population of all students.
    - We can estimate the quality μ of as an estimate of M.
  - This may be important not just as an end in itself, but for us to figure out what to expect by way of other sample means.

III ANALYSIS
B. Inferential Statistics

- The way to go from sample M to population μ is to compute the **Standard Error of the Mean.**
  - SEM is a measure of the extent to which different samples may have different means.
  - The SEM is computed by dividing the sample SD by the square root of the number of estimates.
  - $SD/\sqrt{n}$
  - 1.34/5.48 = .24

III ANALYSIS
B. Inferential Statistics

- Imagine randomly selecting classes in the university and weighing the students in them.



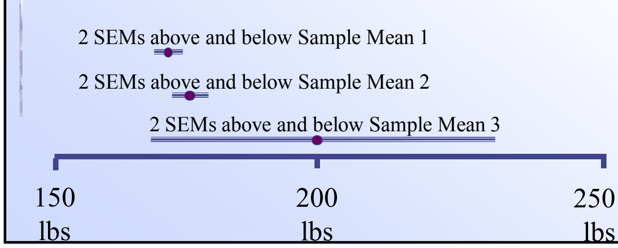
Is one sample different in weight?

III ANALYSIS
B. Inferential Statistics

- Sample 3 looks different from the other samples
  - From a Descriptive Statistics point of view, it looks a lot different!
    - Its mean is much higher (200 vs. 175 & 180 lbs)
    - Its standard deviation is much higher (60 vs. 25 & 48 lbs)
    - Its SEM is also much higher (15 vs. 2.5 & 4 lbs)
  - From an Inferential Statistics point of view, Sample 3 may not be so different!
    - Its SEM suggests that it is a weak estimate of the population mean, ranging from 170 lbs (2 SEMs below its mean) to 230 lbs (2 SEM above its mean)
    - This range includes the ranges of the other means

III ANALYSIS
B. Inferential Statistics

It is very reasonable to assume that sample 3 is consistent with samples 1 and 2, in terms of the population each sample is estimating!



III ANALYSIS
B. Inferential Statistics

- **B. Null Hypothesis Testing**
  - This exercise in judging whether separate samples make the same estimates of populations is at the heart of *Null Hypothesis Testing*.
  - The logic of null hypothesis testing is at the heart of analyzing IG designs:
    - Assume that the independent groups are the same. (Null Hypothesis)
    - Find evidence that they are different. (Significance Testing)
    - Conclude with a degree of certainty that your initial assumption was incorrect. (Statistical Conclusion)

III ANALYSIS
B. Inferential Statistics

- **B.1.i Null Hypothesis**
- Why do we form the null hypothesis (M_1=M_2)?
  - Trying to fool ourselves? Pretend that we don't want to find what we want to find?
- Statistically, we are assuming that they are sampled from the same population.
  - The population estimates of each sample will be compared to see whether they ***triangulate*** – that is, whether they agree on the same estimates.
  - While the samples may be quite different, they may agree on the same population estimates.

III ANALYSIS

B. Inferential Statistics

- **B.2.ii Significance Testing**
- To test whether samples agree on population estimates, we have to do significance testing
  - Significance testing asks the question, Do group means differ from each other more than we would expect from chance?
  - If so, then the difference between groups is not just a difference that may be expected by sampling two random samples from the same population.
  - Rather, the two groups differ because they can not be said to come from the same population!

III ANALYSIS

B. Inferential Statistics

- **B.2.ii Significance Testing**
- Growing Plants:
  - You want to find out whether or not the fertilizer you use is cost-effective in growing tomatoes.
  - Randomly assign plots of land to be treated or untreated by the fertilizer
  - Grow tomato plants on both plots and find that the fertilized plots have 1.65 more tomatoes per plant.
  - Is it worth using the fertilizer?
  - Yes, if you expect little variability in the number of tomatoes per plant. No, if you expect much more.

III ANALYSIS

B. Inferential Statistics

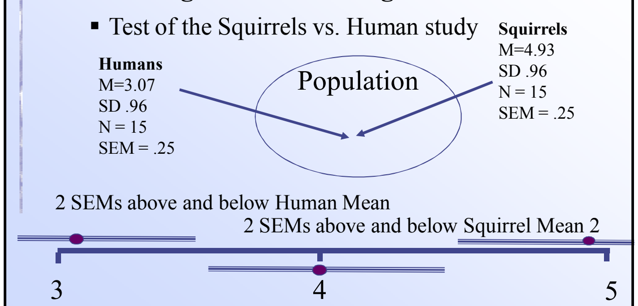
- **B.2.ii Significance Testing**
- We just outlined the basic procedure of *t*-test.
- *t*-Statistic: Compares the difference between the means to an estimate of the extent to which randomly selected sample means will vary.
  - *t* = Difference between means / SEMs

$$\frac{M_1 - M_2}{\sqrt{(SD_1/\sqrt{n_1})^2 + (SD_2/\sqrt{n_2})^2}}$$

III ANALYSIS

B. Inferential Statistics

- **B.2.ii Significance Testing**
  - Test of the Squirrels vs. Human study



III ANALYSIS
B. Inferential Statistics

- **B.2.ii Significance Testing**
  - If the *t*-value ratio is above roughly 2, then we say that the difference between groups is real, not due to chance.
  - Why "2" is a long story, but basically because 2 standard deviations from the mean represents a very unlikely event ($p < .05$), so a critical ratio of 2 is also considered very unlikely to occur by chance alone.
  - Actually, the critical t value is determined as a function of the DEGREES OF FREEDOM (n-2) and checked on a "critical values of *t*" table.

III ANALYSIS
B. Inferential Statistics

- **B.2.ii Statistical Conclusion**
- The critical value of *t* further depends on whether your test is one-tailed or two-tailed.
  - A two-tailed test assumes no directionality to the hypothesis. The prediction is that one group is different than the other, without specifying which one.
  - A one-tailed test assumes directionality to the hypothesis. The prediction specifies that one particular group (Students) scores higher than the other (Squirrels).

III ANALYSIS
B. Inferential Statistics

- **B.3. Other Concepts**
- **1. Type 1 and Type II Error**
  - When the observed value of *t* is greater than the critical value of *t*, we conclude that the difference is significant!
  - It doesn't mean it's an important or valuable difference, only one which is greater than what we would expect by chance.
  - Statistically significant doesn't even mean not unlikely, only that the difference would happen 5 times or less out of 100.

III ANALYSIS
B. Inferential Statistics

- We may be wrong in our inferential conclusion!
- **Type I Inference Error**: Reject null hypothesis when it's true
  - CONSEQUENCE: You earn a bad reputation because you will publish data which looks significant but can't be replicated.
- **Type II Inference Error**: Fail to reject null hypothesis when it's false
  - CONSEQUENCE: Lost chance at finding significance. It was there, but you missed it!

III ANALYSIS

B. Inferential Statistics

- **B.3. Other Concepts**
- **2. Parametric and Non-Parametric Statistics**
- A *t*-test is a parametric statistic because it requires making estimates of populations.
  - Such estimates are central in null hypothesis testing.
  - But sometime such estimates make no sense.
  - Consider a distribution of 10 boys and 10 girls, what is the population estimate of gender? 1.5?
  - Only Interval and Ratio scaled variables can be assumed to offer meaningful population estimates.

III ANALYSIS

B. Inferential Statistics

- **B.3. Other Concepts**
- **2. Parametric and Non-Parametric Statistics**
- Non-parametric statistics (e.g., chi-square) do not require making population estimates.
  - Non-parametric statistical methods can be used to perform statistical significance testing on Nominal or Ordinal variables.

III ANALYSIS

C. Chance and the *t* distribution

- 2. Experiments as the production of variance.
- Science as the production and understanding of variance can now understood not only at the level of design (IV ➔ DV), but also at the level of statistical analysis.
  - A *t*-test is an examination of two types of variability (difference between means and variability of groups) and computing a ratio between them.
- Think variability.

III ANALYSIS

C. Chance and the *t* distribution

$$\frac{M_1 - M_2}{\sqrt{(SD_1/\sqrt{n_1})^2 + (SD_2/\sqrt{n_2})^2}}$$

- **1. Research with Statistics in mind.**

What is the consequence on the significance of the t-value of…

1. Lowering alpha level ($p<.05$) to ($p.<01$)?
2. Increasing sample size?
3. 2-tail vs.1-tail testing?
4. Increasing effect size (M_1-M_2)/sd (pooled)

Significance is …

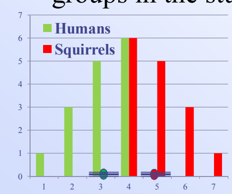
1. harder to find
2. easier to find
3. depends
4. easer to find

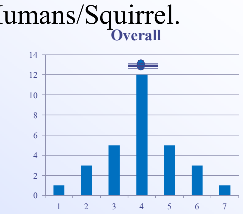
III ANALYSIS
D. F and X^2 Tests

- 1. Analysis of Variance (ANOVA) tests
  - Premised by the idea of a triangulation of *variance estimates* of a population mean, which can be estimated by
    - averaging the deviations of **each observation** in the sample from the overall mean.
    - averaging the deviations of **each subgroup mean** (IV level) in the sample from the overall mean.
  - If estimates converge, then the subgroups are assumed to be sampled from the same overall population, but if not, then the subgroups are not assumed to be sampled from different populations.

III ANALYSIS
D. F and X^2 Tests

- Think of the estimates within and between groups in the study of Humans/Squirrel.





Do estimates of the variance of the overall mean based on variability of the two samples converge with estimates based on variability within the sample of 30 participants ?

III ANALYSIS
D. F and X^2 Tests

- ANOVAs analyze the variance in the data.
  - In a simple ANOVA, variance is devided or partitioned into.
    - **Between Group Variation**: Variability in scores associated with IV but also with individual differences and error. Assumed to contain both *Error* and *Systematic* Variance.
      - What is the variance of the population mean, based on the sample means?
    - **Within Group Variation**: Variability in scores associated with individual differences and measurement error. Assumed to contain only *Error* Variance.
      - What is the variance of the population mean, based on the overall sample?

III ANALYSIS
D. F and X^2 Tests

- An *F*-ratio is computed to test for significance in an ANOVA
  - The F ratio is relates variability between Between Groups / Within Groups.
    - Error + Systematic Variance / Error Variance
    - These sources of variability is computed as the **sum of squares** (squared differences from the mean)
      - Just like a *t*-ratio, The *F*-ratio provides information about the ratio of IV-related variability relative to the variability expected by chance.
    - If that ratio is higher than a critical value (on an ANOVA table) it is significant.

III ANALYSIS

D. F and X^2 Tests

- **2. X^2 (chi square) Test**
  - The chi-square test is a non-parametric test as there is no estimating of population statistics.
    - For use with nominal variables.
  - In a simple chi-square test, observed distribution of scores is compared to an expected chance distribution.
  - Consider a study of whether blondes do indeed have more fun.
    - Blonde and non-blonde participants took a "Fun" questionnaire and divided into those having and not having fun

III ANALYSIS

D. F and X^2 Tests

	Blonde	Non-blonde	
Fun	10 (7.5)	5 (7.5)	15
No Fun	5 (7.5)	10 (7.5)	15
	15	15	30

III ANALYSIS

D. F and X^2 Tests

- 2. X^2 (chi square) Test

To compute the Chi-square test, the formula is:

$$\Sigma(O-E)^2/E$$

$4\ (\Sigma(2.5)^2/7.5) = 4\ (.83) = 3.33$

Chi Square Critical = 3.84 (with df= 1 and p = .05)

So the there is no significance

IV PUTTING IT ALL TOGETHER

A. Statistics for Researchers

- The moral of the statistics story for researchers is that you have design research with statistics in mind! Two particularly important rules:
  - **1. Think about error variance and try to minimize it because all error variance will lower your chances of finding significance.**
    - Use reliable instruments!
    - Hold extraneous variables constant
    - Match or control variables
    - Randomize error variance across groups through random assignment.
      - May not reduce error variance but it will minimize Type I error.

IV PUTTING IT ALL TOGETHER
A. Statistics for Researchers

- Floor and Ceiling effects in measurement can now be understood as a dangerous source of Type 1 (Bad Reputation) error.
  - If a measuring instrument has (unnaturally) little variability, what happens to sd, SEM, t- and F-ratios?
  - The sd and SEM lower, making the t- and F-ratios much higher, resulting in unwarranted statistical significance (Type 1 error).

IV PUTTING IT ALL TOGETHER
A. Statistics for Researchers

- **2. Increase the Effectiveness of your IV**
  - Use extreme groups, weak IVs produce only Type II error.
  - Use sensitive DVs which will detect differences in IVs.