**ESOC 302**

**Exam 2 Review Sheet**

**Fall 2019**

This review sheet lists key terms and concepts that you should know for Exam 2. **All** definitions and answers can be found from the lecture PowerPoint decks, the assigned readings, and from notes that you may have taken in class from our classroom discussion and activities.

Some notes about the exam (as discussed in class):

* Exam 2 is located under the “Quizzes” tab on our course homepage.
* Exam 2 will have 50 questions.
* You must start sometime between 9:30 a.m. and 9:45 a.m. on Tuesday, November 12th. The exam will not be available after 9:45 a.m.—think of this as a virtual lecture hall door locking people out.
* One attempt is allowed; you may not close your internet browser and start over once you have started the exam.
* You may take Exam 2 wherever you want to. Make sure you have a stable internet connection!
* You will have 75 minutes to complete Exam 2.
* There is a 2-minute grace period. Exams that exceed the 2-minute grace period will automatically receive 0 out of 100. No exceptions.

Some notes about the following review sheet:

* The items listed below are starting points that can be used to help you review for the exam (hence, a review sheet). **This is not necessarily an exhaustive list for all of the test questions.**
* You should know *all* relevant material related to a given issue. For example, the term “measures of central tendency” has some key components and different types. Know them.
* Most of the topics from the assigned readings and lecture PowerPoint decks correspond to the specific items listed below.

Work hard!

—Nate

**Experimental designs and internal validity**

* What notation do we use to describe experimental designs?

R = Random assignment if participants

X = Manipulation or observation of the IV

O = measurement of the DV

* What are the three broad categories of experimental designs?
  + True experimental
  + Quasi-experimental
    - Researchers don’t have as much control compared to true experiemnt
  + Pre-experimental
    - Random assignment is NOT present (no “R” Notation)
    - IV is observed, not manipulated
* What does random assignment mean?
  + Recall: Random selection is for sampling
  + Random assignment is after we have our sample
* Know the three true experimental designs
  + Posttest-only Control Group Design
    - Randomly assign participants to 2 different groups.
    - Group 1 is manipulated with IV; Group 2 is manipulated with DV
    - DV is measured in both groups afterwards
      * Group 1: treatment group (Nothing happens)
      * Group 2: control group
  + Pretest-Posttest Control Group Design
    - Randomly assign participants to 2 different groups
    - DV is measured in both groups at the start of the experiment and the end of experiment
    - Group 1 is manipulated with IV; Group 2 is not
  + Solomon 4-group
    - Combination of the first two true experimental designs
* Know the three quasi-experimental designs
  + Nonequivalent control group design
  + Interrupted time-series design
  + Multiple time-series design
* Know the three pre-experimental designs
  + One-shot case study
  + One group, pretest-posttest design
  + Static group comparison
* What is validity and, specifically, what is internal validity?
  + Validity: Are we measuring what we intend to measure
  + Internal Validity: Describes the extent to which a study identifies the “true” cause of variation in the DV
* How does internal validity relate to causality?
* Know the seven types of threats to internal validity
  + History
  + Testing
  + Maturation
  + Experimental Mortality/Attrition
  + Statistical Regression
  + Selection
  + Instrumentation

**Descriptive statistics and relative standing**

* Core assumptions of univariate statistics and normalcy
  + Univariate Normal: The data are not statistically significantly different from a normal distribution.
* The three measures of central tendency
  + Mean
    - Arithmetic average
    - Most sensitive to extreme scores
  + Median
    - Middle of all scores on one variable
    - Not sensitive to extreme scores
    - Indication compared to mean of non-normal distribution
  + Mode
    - Score that appear most often
    - Useful for continuous and categorical data
    - Some distributions have more than one mode (two modes = bimodal)
* Three ways to describe the curvature of the data
  + Normal Curve (mean = median = mode)
  + Positively skewed curve (Mean > Median > Mode)
  + Negatively Skewed Curve (Mean < Median Mode)
* The two measures of dispersion
  + Dispersion: How dispersed (stretched or squeezed) a distribution is
    - Skew
      * Skewness = symmetry
      * Positively skewed: The most frequent score and the middle score are less than the mean
      * Negatively skewed: The most frequent score and the middle score are greater than the mean
    - Kurtosis
* Standard deviation (p15)
  + In a nutshell: SD tells us the average distance of raw scores from the mean
  + Small SD = The data tend to be similar to the mean
  + Large SD = the data are spread apart from the mean
  + SD = zero = every single data point is exactly the same
* Variance (p21)
  + Variance is the standard deviation before taking the square root
* The theoretic normal distribution vs. the empirical normal distribution vs. the standard normal curve (p24)
* Know what z-scores are and how to calculate them (p32)
  + A Z-score states the number of standard deviations by which the original data point lies above or below the mean
  + Therefore, a Z-score means that the data point has a score that is the average
* Directional vs. nondirectional hypothesis testing and z-scores
* Standard error

**Hypothesis testing**

* Know the difference between a directional, a nondirectional, and a null hypothesis
  + Directional: A one-tailed directional hypo predicts the nature of the effect of the independent variable on the dependent variable. (Adults will correctly recall more words than children)
  + Nondirectional: A two-tailed non-directional hypo predicts that the independent variable will have an effect on the dependent variable, but the direction of the effect is not specified. (There will be a difference in how many numbers are correctly recalled by children and adults)
  + Null: The null hypo states that there is no relationship between the two variables being studied (One variable does not affect the other)
* Why do we use the null as our comparison value and what do we use when we make our decision to reject (or fail to reject) the null hypothesis? (6)
  + We test H0 because we can logically deduct what the H0 “value” is in the population
    - We compare our sample statistics to the null hypothesis value
    - We decide whether or not to reject H0 based on the likelihood that our sample statistics different from the H0 value based on various criteria we set
* What does it mean to reject the null?
  + When we reject H0
    - The probability of getting a sample statistic this large (or small) by random chance alone is so unlikely that there probably is some difference between or association between variables.
* What does it mean to fail to reject the null?
  + When we fail to reject H0
    - There isn’t enough evidence to reject the notion that some difference or relationship between our variables does not exist
* What types of errors can we make in hypothesis testing?
  + Type 1: We reject the null hypothesis
  + Type 2: We fail to reject the null hypothesis
* Are “alpha” and “p-value” synonymous and interchangeable terms?
  + Alpha
    - The likelihood of making a Type 1 error
    - Rejecting H0 when we really shouldn’t
    - Alpha is the criterion we use/set before doing statistics for our p-value
      * Set alpha to 0.05 -> p-value needs to be <0.05 to be Statistically significant
* How are alpha and p-value related?
  + Set alpha to 0.05 -> p-value needs to be <0.05 to be statistically significant
  + Set alpha to 0.01 -> p-value needs to be <0.01 to be statistically significant
  + Set alpha to 0.001 -> p-value needs to be <0.001 to be statistically significant
* What are the five steps used in hypothesis testing?
  + Propose a research hypothesis
  + Get a sample data
  + Compare sample data to the sampling distribution under conditions of the null hypothesis
  + Select a criterion (alpha) that will help you decide whether or not you reject H0
  + Calculate the CI and the decide whether or not to (1)reject or fail H0 (2) support or fail to support H1
* Know how to use the two confidence intervals we learned for Exam 1 and how to interpret them

**Categorical variables**

* What level of measurement is chi-square used for?
  + Big Picture: How are categorical variables and/or levels related to each other
  + Categorical Nominal
* What are the two different types of chi-square and how do they differ?
  + One-way chi-square
    - Only when we have one categorical variable
  + Tow-way chi-square
    - When we have two categorical variables of interest
    - Comparing how different levels of the first categorical variable vary with the levels of the second categorical variable
* Which type of chi-square is more useful, for our intents and purposes (i.e., which one are you allowed to use for the research paper)?
  + Two way
* Know the differences between expected and observed frequencies
  + Expected Frequencies (denoted by some form of “e” )
    - What we’d see under conditions of H0 (no association)
  + Observed Frequencies (denoted by some form of “o”)
    - What our sample data actually are
* Know how to calculate a one-way chi-square and a two-way chi-square
* Know how to interpret the expected cell totals, the chi-square statistic for each cell, the chi-square statistic, and the *p*-value
* What odds and odds ratios?
  + The odds and odds ratios allow us to interpret and talk about chi-square in meaningful ways
  + Odds: The ratio (think division) of the frequency of one event occurring to frequency of a second event occurring.
  + Odds ratios: The ratio (think division) of the odds of two levels of one variable in relation to two levels of a second variable
* Know how to use and interpret the odds ratio calculator
* Know the null hypothesis value for odds ratios
* Understand the difference behind two-way chi-square and the odds ratio