W203: Exploring and Analyzing Data
Summer 2015
Lab 3

Instructions

Please answer **ALL** parts on the bcourses **and also** construct an R markdown report that addresses Part 3 showing your code, and what R reports back to you.

Part 1: Multiple Choice (32 points)

For the following questions, please choose the best answer and provide the correct letter in your response.

1) A software bug affects all Mac laptops that were manufactured in April of 2012, permanently replacing the user's desktop with photos of cats. You measure productivity levels among users that own these "buggy" laptops and compare them to users that own "non-buggy" laptops made in the months immediately after April 2012. This is an example of a…
   a. Pretest-postest experimental design
   b. Natural experiment
   c. Associational non-experiment
   d. Confounding covariate
   e. Solomon 4-group design

2) A researcher obtains a convenience (non-random) sample of 100 undergraduates to come to a lab. Each participant is given an anxiety questionnaire before being randomly assigned to only one of four conditions, "music by Rebecca Black," "music by Nickleback," "music by Yanni," and "quiet room." The participants are then given the anxiety questionnaire again after their treatment. This is an example of a…
   a. Pretest-postest experimental design
   b. Quasi-experiment
   c. Associational non-experiment
   d. Confounding covariate
   e. Solomon 4-group design

3) Which of the following is true about parametric tests of mean difference compared to rank-sum tests?
   a. Non-parametric tests are more appropriate than parametric tests when the assumptions of the central limit theorem are met.

b. Parametric tests typically have higher statistical power than non-parametric test when parametric assumptions are met.
c. Parametric tests capture the meaningful information in Likert-type scales without assuming regular intervals between levels.
d. Parametric tests are always more robust than rank-sum tests.
e. Both a and c.
f. None of the above.

4) We follow a sample of 100 neurosurgeons for a week and record how many MRI and CAT scans they request as their first diagnostic test for new patients suspected of having a brain tumor. What is the most appropriate way to tell if there is a significant difference between the mean number of these two types of scans?
   a. Independent-sample t-test
   b. Chi-square test
   c. Dependent-sample t-test
   d. Effect size calculation
   e. ANOVA
   f. None of the above

5) Suppose that 60% of San Franciscans drink beer, and 70% wear plaid. According to the assumption that drinking beer and wearing plaid are independent, if you survey a random sample of 100 San Franciscans, how many do you expect to drink beer and wear plaid?
   a. It depends on the standard deviation of each variable.
   b. It depends on which alpha value (.05, .01, etc) we choose.
   c. It depends on the assumption of linearity.
   d. 42
   e. 60
   f. 50

6) You have a data set of 100 people, recording how many hours of TV they watch during each month of a calendar year. It turns out that December has a higher mean number of TV-hours than any other month, and an uncorrected t-test of December against the other months is significant. Should you run a correction (e.g. Bonferroni) on your p-value for the t-test involving December?
   a. Yes, because the odds of getting a type 1 error for at least one month are inflated above 5%.
   b. No, because December is a unique month due to the holiday shopping season.
   c. It depends on your theory and how you came to observe this difference in means.
   d. Run an ANOVA and see what happens.

7) You randomly assign participants to treatment and control in a study of self-confidence, but you learn that your lab assistant only invited tall and attractive people walking on Sproul plaza to be participants. This could potentially threaten which of the following?
   a. Internal validity
   b. External validity
   c. Your ability to draw causal conclusions
   d. Ecological validity
   e. Your measurement precision
   f. Construct validity

8) In a study of Pacific spinner dolphins, you measure the ratio of dorsal fin length to playfulness (operationalized as flips in the air per minute) and you find a statistically significant difference between means for el Niño and la Niña years. Which of the following is the best effect size measure for this phenomenon?
   a. The correlation between the values measured in El Niño year and the values measured in La Niña years.
   b. Difference of means, divided by the pooled standard deviation of the variable.
   c. Difference of means, divided by the standard error of the statistic.
   d. Difference of means, divided by the sample size

Part 2: Test Selection (20 points)

Every other year, the General Social Survey collects responses to thousands of questions, covering a wide variety of topics. You will be using a subset of data from 1993, including a small number of variables.

While some variables may be self-explanatory, others may not make sense until you look at the GSS codebook. An easy way to investigate a variable is to look it up in the GSS mnemonic index, located at:

http://www3.norc.org/GSS+Website/Browse+GSS+Variables/Mnemonic+Index/

Before you run a statistical test on a variable, you should always read its description in the codebook, in order to understand what different values mean. For example, the codebook may explain if certain values stand for missing data. If this occurs, you should make sure those values are recorded as NA in R before proceeding.

Like any survey, GSS data creates additional concerns that would normally go into a statistical analysis. Surveys are usually weighted in order to compensate for over- or under-representation

of subgroups. For this lab, however, you will be using unweighted data, which limits how well your findings generalize to the U.S. population.

For the following problems, select the most appropriate statistical test to answer the question from the given choices.

Note: You do **NOT** need to execute the test.

9. Is income (income91) higher for men who visit art museums (visitart) or men who don't?
   a. t-test
   b. Pearson correlation
   c. Wilcoxon signed-rank test
   d. Wilcoxon rank-sum test
   e. Chi-square test
   f. Anova

10. Is age related to opinion of country music (country)?
    a. t-test
    b. Wilcoxon rank-sum test
    c. Chi-square test
    d. Anova
    e. Fisher's exact test

11. Do Catholic (relig) men or Catholic women have a greater number of siblings (sibs)?
    a. t-test
    b. Pearson correlation
    c. Wilcoxon rank-sum test
    d. Chi-square test
    e. Anova
    f. Shapiro-Wilk test

12. Is years of education (educ) related to hours spent watching television (tvhours)?
    a. t-test
    b. Pearson correlation
    c. Wilcoxon rank-sum test
    d. Chi-square test
    e. Anova
    f. Levene's test

13. Is political orientation (politics) related to how often someone reads the news (news)?
    a. t-test
    b. Pearson correlation

     c. Wilcoxon signed-rank test
     d. Chi-square test
     e. Anova

Part 3: Data Analysis and Short Answer (48 points)

Write a well-commented R script to perform each of the following tasks, following the best practices described in class. For each new variable, look for obvious errors and make sure that appropriate values are coded as NA.

Include all important output, variable recoding, and discussion in your main report. I should be able to understand what you did and what your answer is for each item without opening your R script.

14. Task 1: Conduct a chi-square test to determine if there is an association between marital status (marital) and political orientation (politics).

    A. What are the null and alternative hypothesis for your test?

    B. What test statistic and p-value do you get?

    C. Conduct an effect size calculation for your relationship.

    D. Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?

15. Task 2: Conduct a Pearson correlation analysis to examine the association between age when married (agewed) and hours of tv watched (tvhours).

    A. What are the null and alternative hypotheses for your test?

    B. What test statistic and p-value do you get?

    C. Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?

16. Task 3: Create a new binary/dummy variable, "married", that denotes whether an individual is currently married or not currently married. Next, we want to consider just the subpopulation of 23-year olds in this sample. Conduct a Wilcox rank-sum test to

determine whether your new "married" variable is associated with the number of children (childs) *for respondents who are 23 years old*.

A. **What is the mean of your new "married" variable among 23-year-olds (e.g., the proportion of cases in the category coded "1")?**

B. **What is the null and alternative hypotheses for your test?**

C. **What test statistic and p-value do you get?**

D. **Conduct an effect size calculation for your relationship.**

E. **Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?**

17. Task 4: Conduct an analysis of variance to determine if there is an association between religious affiliation (relig) and age when married (agewed).

A. **What is the null and alternative hypotheses for your test?**

B. **What test statistic and p-value do you get?**

C. **Are there statistically significant differences between individual pairs of groups, and if so, how do you know?**

D. **Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?**