PSYC 5143
Spring 2024
Problem Set #2

This is due on Monday, February 5 by 2pm via Blackboard, preferably as a single R file. Ask for help before frustration sets in. (An article is linked for the curious.)

1) Zaval et al. (2015) showed that priming so-called legacy motivation (i.e., the motivation to take action to benefit future generations) resulted in more pro-environmental behaviors and intentions. Imagine that a researcher attempted to replicate this effect, randomly assigning 20 participants to either a condition in which legacy motives were primed (i.e., participants wrote a short essay asking how they want to be remembered by future generations) or to a no-prime control. After, they were given the opportunity to donate as much of their $10 participation compensation as they liked to an environmental organization. Hypothetical results are in legacy.csv; higher numbers indicate a larger donation.

   a. Specify Models C and A (use words, like *donation = overall mean*, or something like that) and the null hypothesis.
   b. Find the means of the two conditions.
   c. Based on the means from part b, if you use contrast codes of ±½ for the two groups for the augmented model, what should the intercept and slope be?
   d. Estimate parameters for both Models A and C. Use contrast codes of ±½ for the groups for the augmented model. Do the estimates you find match those you predicted in part c?
   e. Show that Model A predicts the condition means.
   f. Provide a precise interpretation of the slope and intercept parameters in Model A. What do the numbers mean?
   g. What is your decision about the null hypothesis? Provide a substantive conclusion as well (i.e., which group scored higher/lower, was the difference significant, what statistics support your conclusion?).
   h. Find (as simply as you can, i.e., use software) a 95% CI for the slope parameter. Does the 95% CI for the slope parameter include 0? Does this match up with your conclusion in part g?

2) Using the data in ps3.csv, which has two variables, *Y* and *group*, answer part a; then in parts b-f fit two-parameter models with the specified numeric codes for the groups and answer the question that follows. (Assign the higher X value to the group with the higher mean.)

   a. What are the group means? What is the difference between the two means? What is the mean of the two means?
   b. Use X = ±1. How are the intercept and slope related to your answers to part a?
   c. Use X = ±½. How are the intercept and slope related to your answers to part a?
   d. Use X = 0 and +1. How are the intercept and slope related to your answers to part a?
   e. Use X = 0 and -1. How are the intercept and slope related to your answers to part a?
   f. Use X = +3 and +5. How are the intercept and slope related to your answers to part a?

3) Using the data in unequal.csv, which has three variables, Y, X, and group, answer the following questions.

   a. What are the group means? How many scores are in each group? What is the overall mean? What is the mean of the two group means?
   b. Model Y (the outcome) as a function of X, which has values of +½ and -½. Is the intercept equal to the overall mean or is it equal to the mean of the two group means?
   c. Using each of the formulas below, calculate what I've labeled as *SS1* and *SS2*; n_k is the number of scores in a group and $\bar{Y}_k$ is the mean of the corresponding group. Which of these is equal to *SSR* for the model you fit in part b?

$$SS1 = \sum_k n_k(\bar{Y}_k - \bar{Y}_{overall})^2 \qquad SS2 = \sum_k n_k\left(\bar{Y}_k - \bar{Y}_{mean\ of\ group\ means}\right)^2$$