PSYC 5143
Spring '24
Problem Set #5

This is due on Monday, February 26, at the beginning of class, submitted on Blackboard, ideally as an R file.

1) A researcher is interested in the effectiveness of different ways of introducing oneself to attractive others. Two confederates (one hot, one not; this is factor A, attractiveness, with two levels: (1) attractive and (2) not attractive) are instructed to use one of two ways to get to know someone (factor B, approach, with two levels: (1) casual conversation and (2) humor). Each confederate finds and engages five different people using each approach, and the dependent measure is how long the confederate is able to engage an individual in conversation. The data are in pick-up.csv. In this file, for the attract variable, 1 = the attractive confederate & 2 = the not-so-attractive one; for the approach variable, 1 = casual conversation & 2 = humor. Find the cell and marginal means. Dummy code each of the factors, with not attractive & casual conversation as the reference groups. Model the outcome based on the dummy codes and their interaction. (If you think modeling the data once more w/different dummy-coding will be informative, do it!) Write a few sentences summarizing and interpreting the results of the modeling as if you were writing to someone seeking advice about meeting others. Don't worry about FWER or the FDR.

2) (This question uses similar to the data from a problem set in the fall, so the sense of déjà vu is real.) The file exercise2.csv includes simulated data for *n* = 244 subjects on physical endurance (*yendu*), age (*xage*), and years of vigorous physical exercise (*zexer*) in which the subjects have engaged. Two new variables have been added to the data to dichotomize the participants into old and young (*ageGroup*) and high and low exercisers (*exerGroup*); the cut-points for high and low values on these variables were chosen so that there are 61 scores in each combination of high and low in this research if it is conceptualized as a 2 × 2 design. This dichotomization is overwhelmingly a bad idea as an analytic strategy (MacCallum et al., 2002), but it serves a useful purpose here. Please make sure that the grouping factors are treated as factors by R. **Statistical significance is of no interest in any part of this question. If you write something about a variable being significant or not, you will get a score of 0 on this assignment.**

   a. Fit a model with endurance (yendu) as the outcome and continuous age (xage) and continuous exercise (zexer) as predictors; do NOT include an interaction term. Report the slopes of the two predictors.

   b. Mean-center xage and zexer. Fit a model with endurance (yendu) as the outcome and mean-centered continuous age and mean-centered continuous exercise (zexer) as predictors; do NOT include an interaction. Report the slopes of the two predictors. Did anything change from your answers to part a? Why or why not?

   c. Fit a model with endurance (yendu) as the outcome and mean-centered continuous age and mean-centered continuous exercise (zexer) as predictors, *along with their interaction*. Report the slopes of the three predictors. Also, interpret each of the slopes. That is, what does the value of each slope tell you precisely.

   d. Center continuous age and continuous exercise at a value that is 1 SD below the mean for each respective variable. Fit a model with these new variables and their interaction as predictors for endurance. Report the slopes and interpret each precisely, as you did in part c.

   e. Center continuous age and continuous exercise at a value that is 1 SD above the mean for each respective variable. Fit a model with these new variables and their interaction as predictors for endurance. Report the slopes and interpret each precisely, as you did in part c.

   f. Find the cell means for the four groups created by crossing the age and exercise grouping variables. (Ask for help if you need it.)

   g. Find the marginal means for the age grouping variable.

   h. Find the marginal means for the exercise grouping variable.

   i. Contrast code each of the grouping variables with values of +½ and -½ for the high and low groups for each variable, respectively. Fit a model with these new contrast codes ***without*** their interaction as predictors for endurance. Report the slopes and interpret each precisely.

   j. Contrast code each of the grouping variables with values of +½ and -½ for the high and low groups for each variable, respectively. Fit a model with these new contrast codes ***and*** their interaction as predictors for endurance. Report the slopes and interpret each precisely.

   k. Dummy code each of the grouping variables with values of 1 and 0 for the high and low groups for each variable, respectively. Fit a model with these new contrast codes ***and*** their interaction as predictors for endurance. Report the slopes and interpret each precisely.