PSYC 5143
Spring '24
Problem Set #6

This is due on Monday, March 4, at the beginning of class, submitted on Blackboard, ideally as an R file.

1) In a study of memory consolidation processes, animals were tested in a one-trial avoidance-learning task (the data file). The animals were presented with a fear-producing stimulus on a learning trial as soon as they stepped across a line in a test chamber. The dependent variable was the time it took them to step across the line on a later test trial (thus, ***longer times indicate better learning***). Two groups of animals differed in terms of the area in which they had electrodes implanted in their cortex (a Neutral site known not to be involved in memory, or Area A, which is the area of interest in this study—these correspond to areas 1 and 2 in the data file, respectively). Each group was further divided and given electrical stimulation either 50, 100, or 150 msec (these correspond to delays 1, 2, and 3 in the data file, respectively) after crossing the line and being presented with the fear-inducing stimulus. If the brain area that was stimulated is involved in memory, stimulation would be expected to interfere with memory consolidation, therefore impeding learning of the avoidance response, and the animal should be faster (relative to the Neutral condition) in recrossing the line. (The manipulation of the timing of stimulation is intended to reveal the timing of consolidation processes. If these processes happen at a particular time, delay should have a simple effect, which may lead to main effect.)

   a. Find cell and marginal means, as well as the mean of all the cell means. Find variance and sample sizes for each cell, too. If the cell sizes are the same, find the mean of the cell variances.

   b. Add to the data contrast codes for the area variable and for the delay variable. (Despite that the delay is inherently numeric, treat it as a factor. If you don't, you're assuming linearity between its numeric characteristics and the outcome, an assumption that may not be warranted.) For the delay variable, choose whatever orthogonal contrasts you want. Finally, add variables that are the products of the contrasts.

   c. Model the outcome on all of the variables you created in part b. For each parameter estimate, including the intercept, say how it is related to the means from part a.

   d. Without worrying about FWER/FDR (i.e., use .05 for all significance tests), write a brief interpretation of each slope's hypothesis test. Be informative (i.e., don't simply say something is "significant").

   e. Use `anova(model)` or `modelEffectSizes(model)` to find *SS* (really it's *SSR*) for each of the predictors in the model in part c.

   f. Create factor versions of the area and delay factors called areaF and delayF and conduct an ANOVA on the data using the R command below (where d is whatever you named your data).

```
summary(aov(latency ~ areaF * delayF, d)
```

   g. In the summary table from part f are the values of *SS* (Sum Sq) related to the *SSR* values from part e? If so, try to say how/why they're related. And is SS_{residual} equal to anything from above?

   h. Using *SSR* values from part e and SS_{residual}, show how these are related to `pEta-sqr` and `dR-sqr` produced for reach predictor by `modelEffectSizes` for the model in part c. (Remember that this function rounds things off, which will make answers slightly imprecise.)

   i. Add to the data dummy codes for the area variable (use the Neutral site as the reference group) and for the delay variable (use 50 msec as the reference group). Also, add variables that are the products of the dummy codes.

   j. Model the outcome on all of the variables you created in part h, and for each parameter estimate, including the intercept, say how it is related to the means from part a.