**Homework 13**

**Due Wednesday, 13 December 2017, 5:00pm**

**Design-your-own homework**

One of the goals of this class is ensuring you can apply the methods you learn to your own research (what would the point of the class be if it wasn’t practically useful?). However, we frankly don’t do much to very actively work toward this goal in class, usually forcing you to use examples that are rarely or never directly applicable to the things you might be doing.

To this end, we’ve created this ~\*~*new*~\*~ homework format. If it goes well, we might use it more frequently in the future. In order for it to work, you need to follow our instructions precisely (or else it will be impossible to grade).

I’ve created a generic data file (codebook below) in long format using a 2 (between) x 3 (within) design. The outcome measure, Y, is on a theoretical 100-point scale. Your task is to create an example *from your own research / academic area* to fit to these data, explain the example, and conduct the analysis as described. Specifically, you should hypothesize an interaction between the two independent variables. Be creative! Consider ways you can think about outcome variables that are most relevant to you in terms of the provided scale (e.g., using a 100-point scale instead of a 7-point Likert scale, some proportion of things). If you’re not sure whether your example makes sense, email the TAs to check. Remember that your subjects do not need to be humans, and your groups don’t need to be manipulated.

Here’s an example from my own research: People are randomly assigned to either receive a prejudice-reduction intervention or not (G). The intervention is a perceived heterogeneity manipulation: those in the intervention group read an article about how Muslims are actually a quite heterogeneous group, while control participants read an article about how birds are a heterogeneous class of organisms. Months later, participants complete a supposedly unrelated climate survey. One part of the survey is feeling thermometers, which measure warmth toward a given social group on a 100-point scale (Y). For this study, we focused on the thermometers for Muslims, Blacks, and gay men (X). I have two hypotheses.

* My hypothesis that can be tested using a contrast is that the intervention will improve warmth toward Muslims more so than warmth toward the other groups.
* My hypothesis that must be tested with dummy codes is that warmth toward Muslims will increase more than warmth toward both Blacks and gay men, but also that warmth toward Blacks will increase more than warmth toward gay men (visible versus concealed identity).

Some questions have very specific instructions about how you name and report things. Please respect these, or else it will be impossible for me to grade the homework. I recommend you read the whole assignment at least twice before starting (just like Mary Berry suggests for her technical challenges on amazing small screen wonder *The Great British Bake Off*).

Codebook for ‘HW13\_Data.dat”

|  |  |  |
| --- | --- | --- |
| **VarName** | **Description** | **Values** |
| X | 3-level within-subjects independent variable | 1 - 3 |
| G | 2-level between subjects independent variable | 1 - 2 |
| Y | Outcome measure | 25.39 – 99.08 |

1. Read in the data. *While they are still in long format,* run one command that shows how Y varies by each combination of X and G. Using R (we haven’t taught you how to do this, but a quick Google will tell you how) *or* good ol’ pen and paper, make a simple bar plot that displays these means (this is for your own reference).
2. Use this bar plot to help you think of an example from your own research (if I were you, I’d label my pen and paper plot with the variable names). Your example doesn’t need to fit the appearance of the data perfectly (i.e., you can hypothesize something different from what the data appear at this stage to be saying), but this is to prevent you from making predictions that are totally absurd in the face of the data you’ve been provided. I’m fine with you not confirming hypotheses, but I’d rather not have you find the opposite of your predictions.  
     
   One of your hypotheses should involve a pattern of means (i.e., use a contrast) while the other should involve 2+ pairwise comparisons.  
     
   In a Word doc, write out your design and hypotheses in the same general manner I did in my example above. Give me just enough information to make it clear what you’re proposing and a rationale for your prediction. Also know I might be very unfamiliar with your research area, so assume you’re explaining the design to a general educated audience. I also have to grade it, so it needs to be a clear enough story that I can follow it through the analysis.
3. Convert the data from long to wide format, and rename variables to make sense for your example. Keep the original name (i.e., Y or G) at the beginning of the new name for it’s easier for me to follow. Use the “factor” command to change the names of the levels of the G variable.  
   So using my example, my new names would be “Gintervention,” “Ymuslim,” “Yblack,” and “Ygmen.” I’d rename the levels of G from 1 and 2 to “control” and “intervention.”
4. Test whether there is an overall effect of group on your outcome of interest. Summarize your result briefly.
5. Test your hypothesis based on a contrast. Restate the hypothesis in your R script so I understand why you’re doing what you’re doing. Summarize your result briefly.
6. Test your hypothesis based on 2+ pairwise comparisons. Restate the hypothesis in your R script so I understand why you’re doing what you’re doing. Summarize your result briefly.
7. Make a graph of these results (like we did in lab, i.e. using the models from your pairwise comparisons). Paste it in your Word doc.
8. In your Word doc, summarize your results (in a separate paragraph from where you wrote the design). Make sure you report the stats that speak directly to your hypotheses. Conclude briefly at the end.

Comment frequently (but concisely) in your R script to make sure I’m on your page. This is also good practice, as open science standards will often have you sending R scripts with this level of detail to people who are curious about your research and analysis.