**Feedback on Categorical Response Analysis memo**

**Emily Swinth for Ben Bui**

**9/23/24**

Really wonderful memo, Ben! This work is very thorough and conveys a ton of difficult concepts clearly. My one overall recommendation is to read through the memo to catch a handful of typos and sentence fragments. I also have a few specific comments to improve clarity below.

3.2:

* I would delete the final sentence that states to skip following section if not interested in internal workings of factors in R. I think the last paragraph of 3.2.1 is likely to be relevant to all readers and I wouldn’t want people to miss it.

4:

* I would mention that using ordinal categorical outcome variables in models earlier — maybe even in section 3 where you initially introduce the types of categoricals.
* You lost me a little bit with the second paragraph of section 4. I see where you’re trying to go with the explanation, but I think the wording is confusing from “First consider the non-rigorous…” to the end of the paragraph. Maybe you could replace with “For example, consider a data set with a categorical variable with the groups “small”, “medium”, and “large”, each of which has been dummied to be used as predictors in a model. What would it mean to find a significant effect only on the “large” dummy variable? By definition, the “small” and “medium” dummies are collinear with the “large” dummy, which would seem to indicate that they should have corresponding significant effects. For this reason, a single, shared statistical test, used to produce a single measure of significant would be preferred when evaluating a categorical variable.”

5:

* Incomplete sentence at the end of the first paragraph
* In your recommendation to use the continuous variable rather than categorical where possible, can you provide an example? I’m struggling to think of a time in my work where this was possible.

5.1

* The first sentence in the final paragraph of 5.1 does not read correctly.

5.3

* I’m not sure I understand the intention of showing the first example model in this section (predicting pre\_income\_raw from treatment and side random effects). It seems like this model code is essentially replicated in the next example, so I don’t think this needs to be included.
* The code to create the mixed\_contrasts object is a little hard to read in the memo. Maybe add more spacing between the steps to increase readability.