1. **Log transformations and interactions.** Some of you may have encountered the idea of using a log-transformation on response variables such as reaction times to make them look more normal and hence be more faithful to the assumptions of linear models. Now suppose you are analyzing a $2 \times 2$ design and are interested in the possibility of a super-additive interaction between the two factors. Your data are reaction times and look more normal when log-transformed. What are the potential consequences of log-transforming your response variable for investigating whether there is an interaction between your two predictors of interest? **Hint:** try constructing a set of four condition means for a two-by-two that reflect a super-additive pattern, and then look at the pattern when you take the log of each cell.

2. **Exercises in mixed-effects regression formula specification, keeping it maximal.** For each of the below informal descriptions of a planned statistical data analysis, state a pair of mixed-effects R regression formulae that will form the basis of a likelihood-ratio test to assess the strength of evidence for the particular scientific hypothesis in question. Predictor variable names are in small capitals. Justify the pair of formulae in each case — in particular, explaining what random-effects structure specification is required to avoid anti-conservative inference due to potential cluster-level idiosyncrasies. (Note that there will not always be clear-cut answers as to what the “right” model formula is; what is most important is that your formula specification matches your assumptions and justification.)

   (a) Rohde et al. (2011) looked at the effect of main-cause VERBTYPE (implicit causality or IC, such as *detests*, versus “ordinary” verb, such as *babysits*) on ATTACHMENT preferences for relative clauses modifying a complex NP object:

   (1)   a. John *detests* the children of the musician who are generally arrogant and rude. (VERBTYPE=IC, ATTACHMENT=low)
b. John detests the children of the musician who is generally arrogant and rude. (VERBTYPE=IC, ATTACHMENT=high)
c. John babysits the children of the musician who are generally arrogant and rude. (VERBTYPE=ordinary, ATTACHMENT=low)
d. John babysits the children of the musician who is generally arrogant and rude. (VERBTYPE=ordinary, ATTACHMENT=high)

They predicted a theoretically critical significant interaction between VERBTYPE and ATTACHMENT on reading times of the relative clauses, with reading times super-additively low for high attachment with IC verbs. They constructed 20 sentence frames (items) with 4 instantiations each along the lines of I above, and had each of 58 subjects read one condition of each sentence frame while recording relative-clause reading times (RT; each participant read 5 sentences of each of the four conditions). They focused on the reading times of the third word of the relative clause (generally above) for their analysis, and included the LENGTH in characters of that word for each item as a covariate to soak up residual noise. What model formula would you use to test their theoretically critical interaction?

(b) In the do–be construction (Wasow et al., 2015), the word to immediately before the copula is often optional:

(2) a. the thing that I tried to do was (to) keep the score close
b. all they can do is (to) circumvent themselves

You are interested in the effects of the following properties of the POST-COPULAR VERB—keep and circumvent in the examples above—on the tendency to include or exclude to (TO-USE): (i) whether it possesses word-initial lexical stress (STRESS) and (ii) the log-word-frequency (FREQ). You collect corpus data and annotate each example for TO-USE, STRESS and FREQUENCY. The examples come from such a diverse collection of speakers and authors that speaker-specific idiosyncrasies are not a major concern. However, you also realize that it may be the case that each VERB of English may have idiosyncratic preferences for use of to in this construction that are not captured by STRESS and FREQUENCY. What mixed-effects model formulae would you use to test whether word-initial lexical stress and verb frequency are associated with differential rates of to-use?

(c) You are interested in studying whether ADJECTIVE–NOUN pairs that are collocations (co-occur particularly often), such as heavy traffic are rated as sounding “more natural” than pairs that do not co-occur particularly often, such as thick traffic or heavy mist. You operationalize collocational strength as point-wise mutual information (pMI), defined as:

\[
pMI(\text{adj}, \text{noun}) = \frac{P(\text{adj}, \text{noun})}{P(\text{adj})P(\text{noun})}
\]
and obtain pMI estimates for a large number of adjective–noun pairs (the estimation procedure is irrelevant for this problem). You then conduct a large rating study where each subject rates the naturalness of 25 adjective–noun combinations varying in pMI on a scale of 1 to 7. Because you have so many adjective–noun pairs, you collect only one rating for each specific pair, but each adjective in your materials set appears with many nouns; likewise, each noun in your materials set appears with many adjectives. Thus, each adjective and each noun were part of stimuli for many collected ratings. What mixed-effects model formulae would you use to test whether pMI has a reliable effect on naturalness ratings?

**Bonus:** how would you change your model specification after collecting lots more data with the same materials, so that now you have several different ratings for each adjective–noun pair?

**References**
