Ednita Tavarez-Jimenez

Nicole Yunes-Perez

November 6, 2022

Week 9 Reading Questions

Q1: Briefly (1 - 2 short paragraphs) describe at least two tradeoffs between the customized ML methods and the canned methods.

Two trade-offs between the customized ML methods and the canned methods are connection and varying models and comparison hypotheses. The book states that “Convention: if you use a standard method, you can just say (for example) “we used linear regression” in your Methods section and no one will think twice. If you use a non-standard method, you need to explain the method carefully and overcome readers’ distrust of “fancy” statistics — even if your model is actually simpler and more appropriate than any standard model. Similarly, it may minimize confusion to use the same models, and the same parameterizations, as previous studies of your system.” In addition, the book also states that “Varying models and comparing hypotheses: the machinery built into R and other packages make it easy to compare a variety of models. For example, when analyzing a factorial growth experiment that manipulates nitrogen (N) and phosphorus (P), you can easily switch between models incorporating the effects of nitrogen only (growth~N), phosphorus only (growth~P), additive effects of N and P (growth~N+P), or the main effects plus interactions between nitrogen and phosphorus (growth~N*P). You can carry out all of these comparisons by hand with your own models, and mle2’s formula interface is helpful, but R’s built-in functions make the process easy for classical models.”

Q2 (1 pt.): Briefly (1 - 2 sentences) describe each of the four key assumptions of the general linear modeling approach.

The four key assumptions of the general linear model observed values are 1. independent and 2. normally distributed with a 3. constant variance (homoscedastic), and 4. any continuous predictor variables (covariates) are measured without error.

Q3 (1 pt.): Explain how the normality assumption can be met in a general linear model, even if the response variable is not normally-distributed. (1 - 2 paragraphs)

The normality assumption can be met in a general linear model, even if the response variable is not normally distributed by the assumption of normality being applied to the variations around the expected value – the residuals – not to the whole data set. The Bolker book states, “linear” part of the “general linear model” means that the models are linear functions of the parameters, not necessarily of the independent variables.”