

Matthew Jusino

Week 9 Reading Questions

I did not work on these questions with anyone else

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

Using the customized special-case procedures can be much faster and safer in computation because you're less likely to run into errors, so you don't have to worry as much about the validity of your modeling. Additionally, using the customized methods allows you to simply say we used the so-and-so method, whereas with nonstandard methods you have to explain in excruciating detail to allow for replication. Aside from that, it's also easier when working with R because a lot of the special-case modeling is already built into R, allowing for easy plug-and-play, so to speak.

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

Normality – This assumption means we should see a normal distribution of observations if the sampling is repeated many times.

Homogeneity – The assumption is made that the variance is the same for all X values, rather than estimating the variance for each X value individually.

Independence – Independence means each Y value should affect only itself, not other Y values. Each Y value should not be influenced by any other Y values in the model.

Fixed X – Fixed X means that we have to always know the exact value of X with almost no noise, or that any error in X is very small compared to the noise in Y values

- **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 with a response variable that is not normally-distributed because the assumption of normality applies only to the variation between the residuals, not the data set as a whole.