Jared Yu

Module 14 Discussion

I thought that your use of a table was a great choice for this discussion. Even though I wrote more, I think that your table perhaps is a better and more concise summary that more directly answers the discussion prompt. I see you are comparing traditional linear regression on the left and a ranked approach on the right. I agree that the advantage of the original approach is that it is basically following linear regression which we have a decent understanding of at this point. You mention also that a ranked approach can better handle outliers, along with not requiring some normality assumption. These are all great points and didn't think of them before. In the disadvantages, you mention issues for the original approach when outliers exist or when there is non-normality. I think it's good you mention this, since we learned some techniques to try and adjust by using something like outlier removal or transformation. However, using these requires us to be careful and considerate of the interpretation following such results. You also mention biasedness for large results in the ranking, I suppose this means that it could perhaps overfit some ranked data by ignoring the fact that certain points are particularly distant from the rest. This is something that I hadn't considered either and I find it to be useful knowledge. In the assumptions, you mention normality for the traditional approach and i.i.d. errors for the ranked approach. I had trouble finding the ranked information in the book, so I tried to just associate it with logistic regression.