Michelle Newcomer

Stat 242

Section Discussion

1. Suggest some alternatives to how the authors designed their study

To set up a simulation study, instead of choosing parameters for which to simulate the data, it would be value to generate data that is randomly mixed and try to fit a model to the data to derive the model parameters. Those parameters could then be used to generate the data which would provide a good dataset for error analysis. The mixture variation parameters (α) are very limited and do not allow the reader to assess how the null model will perform given largely different coefficients. They follow this method in the example cases when they fit the model to the wine adulteration dataset and I think their simulation should follow the same direction of the example (i.e. have data and generate parameters)

1. Interpret their tables on power-do the results make sense?

In terms of how the power varies as a function of the data generating mechanism, the results do make sense. Generally, as the number of samples increases, the probability that the test will reject the null hypothesis when the alternative hypothesis is true (a correct prediction) increases. However, this is only observed for the null models that have no mixture variation. Generally, there is no significant improvement when alpha varies slightly. The standard deviation also introduces a significant improvement in the power suggesting that a dataset with more variation would better conform to a mixture model analysis.

1. Discuss the extent to which they follow JASA’s guidelines

In this paper, it was really difficult to determine how the authors obtained the values for the parameters α, θ, and σ. Those values are very important, and a value such as α, the mixture coefficient can dramatically change the outcome. The reader does not know if α can only vary slightly from the other value or if a mixture model with coefficients (0.90, 0.05, 0.05) is acceptable. Results would not be reproducible without this knowledge. Also, a reader using this model for application purposes would not know if the model is appropriate for the given dataset. The authors do follow the JASA guidelines regarding a discussion of the use of code, software, accuracy, and algorithms.