STAT 2132:
Applied Statistical Methods II
HW #5
Due on Wednesday, February 23

1. We are interested in exploring the effect of socioeconomic status (SES) and writing skills on the selection of tracks in high school:

   - Response is three types of high school programs: 1=general, 2=academic, and 3=vocational.
   - Predictors are SES (1,2,3, larger for higher SES) and writing scores (continuous).

   Use the 'hsbdemo.txt' data posed on canvas and answer the following questions.

   (a) Examine the frequency table of SES and the selection of tracks in high school. What do you notice?

   (b) Examine the mean values of writing scores separated by the three tracks and report your findings.

   (c) Fit a logistic regression model only using the data for those people who chose the general or academic programs. The outcome is the choice of tracks (setting academic as 0 and general as 1), and the predictors are SES (as a categorical variable) and writing scores. Interpret your results.

   (d) Next fit a logistic regression model only using the data for those who chose the vocational or academic programs. This time set academic as 0 and vocational as 1. Discuss your results.

   (e) Now fit a multinomial regression model to the entire data set (setting academic as the reference level), and again with SES and writing scores as predictors. You may use the R function 'multinom' in the package 'nnet.' Are your results from this multinomial regression model consistent with your results in c and d?

   (f) Now, we want to see if SES can be included as an ordinal linear predictor in your multinomial regression model. Do you think the Deviance test you've learned for logistic regression can be used here?