STAT 477/STAT 577
HW 8 - Module 3: Section 2

Every year, graduating seniors submit applications for medical school. There are many factors medical schools use to select applicants. The data file **Med.csv** contains data on 55 seniors who applied to medical school, all from the same liberal arts college in the Midwest. For each applicant, the variables GPA, MCAT, Gender, Apps, and Acceptance were collected. Here is some information on these five variables:

- GPA - Cumulative undergraduate grade point average
- MCAT - Medical School Admission Test, a standardized test that measures aptitude and achievement for medical school
- Sex.1.0 - 1 = Female, 0 = Male
- Apps - Number of medical schools to which the student applied
- Acceptance - 1 = Accepted, 0 = Denied

1. Fit a logistic regression model for predicting the log odds of being accepted into medical school from the variables MCAT and GPA. Use this logistic regression to answer the following questions.

   (a) Give the equation for predicting the log odds of acceptance from the variables MCAT and GPA.

   (b) Give the equation for predicting the probability of acceptance from the variables MCAT and GPA. Use this equation to predict the probability of acceptance for a student with a GPA of 3.54 and a MCAT score of 38.

   (c) Find and interpret a 95% confidence interval for the probability of acceptance for a student with a GPA of 3.54 and a MCAT score of 38.

   (d) Test for the significance of the overall model using the likelihood ratio test.

   (e) Test for the significance of GPA and MCAT scores separately using the Wald test.

2. Now fit a logistic regression model for predicting the log odds of being accepted into medical school from the variables MCAT, GPA and Sex. Use this logistic regression to answer the following questions.

   (a) Give the equation for predicting the log odds of acceptance from the variables MCAT and GPA for Females and the equation for predicting the log odds of acceptance from the variables MCAT and GPA for Males.

   (b) Interpret the coefficient for Sex in the model. Calculate and interpret a 95% confidence interval for the associated odds ratio.

   (c) Test for the significance of the variable Sex in the model with GPA and MCAT using the Wald test.

3. Add an interaction term between GPA and MCAT to the logistic regression model from Problem 1 above. Use this logistic regression to answer the following questions.

   (a) Give the equation for predicting the log odds of acceptance from the variables MCAT and GPA and their interaction.

   (b) Give the equation for predicting the probability of acceptance from the variables MCAT and GPA and their interaction. Use this equation to predict the probability of acceptance for a student with a GPA of 3.54 and a MCAT score of 38. How does this value compare to the one you calculated in Problem 1, part (b)?

   (c) Find and interpret a 95% confidence interval for the probability of acceptance for a student with a GPA of 3.54 and a MCAT score of 38. How does this confidence interval compare to the one you calculated in Problem 1, part (c)?

   (d) Test for the significance of the interaction term in this model using the Wald test.

4. Use the **step** function in R and the criteria AIC and BIC to find a good model to predict the probability of acceptance from the possible explanatory variables MCAT, GPA, Sex and Apps. Explain your selection process. Once you have found a good model, summarize the fit of the model using pseudo R^2, the Hosmer-Lemeshow goodness of fit test, a confusion table and associated statistics, and the ROC Curve and area under this curve.