**STAT 477/STAT 577**

**HW 8**

**Neha Maddali**

1. Model with GPA and MCAT

a. Equation for predicting log odds of acceptance = β0 + β1(MCAT) + β2(GPA)

b. Equation for predicting probability of acceptance = (eβ0 + β1(MCAT) + β2(GPA))/(1 + eβ0 + β1(MCAT) + β2(GPA))

Prediction = (e-22.3727 + 0.1645(38) + 4.6765(3.54))/(1 + e-22.3727 + 0.1645(38) + 4.6765(3.54)) = 0.6069

c. 95% confidence interval for probability of acceptance = (0.4159114, 0.7695156)

Interpretation: We are 95% confident that the true probability of acceptance for a student with a GPA of 3.54 and an MCAT score of 38 lies between 41.59% and 76.95%

d. Test for Overall Significance

* + Null Hypothesis: The intercept only model is sufficient to explain the variation in the response variable.
  + Alternative Hypothesis: The model with MCAT and GPA predictors provides a better fit to the data compared to the intercept only model.
  + Test statistic: 21.777
  + P-value: 1.867e-05
  + Conclusion: There is significant evidence that indicates that the model with MCAT and GPA predictors provides a significantly better fit to the data compared to the intercept-only model.

e. Test for GPA variable (use Wald Test)

* + Null Hypothesis: the coefficient of GPA is equal to zero
  + Alternative Hypothesis: the coefficient of GPA is not equal to zero
  + Test statistic: 2.849
  + P-value: 0.004389
  + Conclusion: There is significant evidence that the GPA is a significant predictor of Acceptance.

Test for MCAT variable (use Wald Test)

* + Null Hypothesis: the coefficient of MCAT is equal to zero
  + Alternative Hypothesis: the coefficient of MCAT is not equal to zero
  + Test statistic: 1.595
  + P-value: 0.110786
  + Conclusion: There is weak evidence to suggest that the MCAT variable is statistically significant in predicting acceptance into medical school.

2. Model with GPA, MCAT, and Sex

a. Equation for predicting log odds of acceptance for Females = −23.9851 + 0.1809(MCAT) + 5.1392(GPA)

Equation for predicting log odds of acceptance for Males = (−23.9851−1.2580) + 0.1809(MCAT) + 5.1392(GPA)

b. Slope coefficient for Sex variable = -1.2580

95% confidence interval for Slope for Sex variable = (0.0605968,1.1173393)

Interpretation: With 95% confidence, the odds of acceptance for males compared to females could be between approximately 0.061 times to 1.117 times higher, holding all other variables constant. Since the confidence interval includes 1, the coefficient is not statistically significant at the significance level of 0.05. This suggests that there may not be a significant difference in the odds of acceptance between males and females when controlling for MCAT scores and GPAs.

c. Test for Sex variable (Use Wald Test)

* + Null Hypothesis: the coefficient of the Sex variable is equal to zero
  + Alternative Hypothesis: the coefficient of the Sex variable is not equal to zero
  + Test statistic: -1.723
  + P-value: 0.084965
  + Conclusion: There is weak evidence to conclude that the sex variable is a significant predictor in acceptance.

3. Model with GPA, MCAT, and interaction term between GPA and MCAT

a. Equation for predicting log odds of acceptance = β0 + β1(MCAT) + β2(GPA) + β3(MCAT x GPA)

b. Equation for predicting probability of acceptance = 1 / (1 + e-(17.0632 – 0.9359(MCAT) – 6.6350(GPA) + 0.3154(MCAT\*GPA)))

Prediction = 1 / (1 + e-(17.0632 – 0.9359(38) – 6.6350(3.54) + 0.3154(38\*3.54))) = 0.60795

Compare prediction to one from Problem 1 part b: Both predictions are very close to each other. The inclusion of the interaction term in the model seems to have had minimal impact on the predicted probability of acceptance for this student. So, in this case, the interaction between GPA and MCAT doesn’t significantly alter the predicted outcome compared to the model without the interaction term.

c. 95% confidence interval for probability of acceptance = (0.405431, 0.777963)

Interpretation: The 95% confidence interval for the probability of acceptance for a student with a GPA of 3.54 and an MCAT score of 38, as calculated in this model is between 40.54% to 77.80% for the probability of acceptance.

Compare CI to one from Problem 1 part c: Both intervals show a high likelihood of acceptance of acceptance with slight variations in the exact range.

d. Test for Interaction Term (Use Wald Test)

* + Null Hypothesis: the coefficient of the interaction term is equal to zero
  + Alternative Hypothesis: the coefficient of the interaction term is not equal to zero
  + Test statistic: 1.101
  + P-value: 0.271
  + Conclusion: There is weak evidence to conclude that there is a statistically significant interaction effect between MCAT and GPA on the probability of acceptance.

4. Describe Selection Process: To select the best model, utilize the step function, considering the criteria of AIC and BIC. The stepwise selection method evaluates models by iteratively adding or removing variables based on their impact on model fit, measured by AIC and BIC. Start with the full model including all possible explanatory variables, the step function sequentially adds or removes variables to minimize AIC or BIC, resulting in a final model with the most parsimonious yet informative set of predictors.

Final Model = Acceptance ~ GPA + Sex

Pseudo R^2 = 0.3299149

Hosmer-Lemeshow Goodness of Fit Test

* Null Hypothesis: the model fits the data well
* Alternative Hypothesis: the model doesn’t fit the data well
* Test Statistic: 7.7273
* P-value: 0.052
* Conclusion: There is moderate evidence to suggest that the model fits the data well.

Confusion Table

A black background with white text

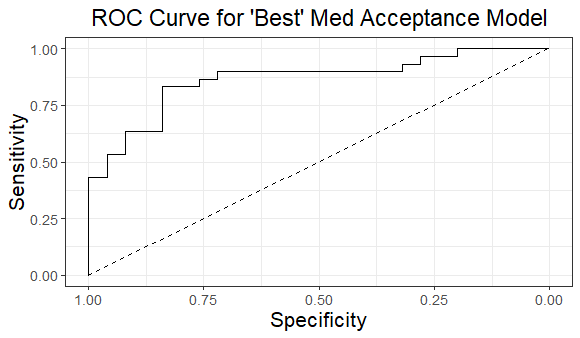
Description automatically generated

Agreement: 0.8363636

Sensitivity: 0.8333333

Specificity: 0.84

ROC Curve



Area under ROC Curve: 0.8653

Interpretation: The AUC of 0.8653 suggests that the model is effective at correctly ranking the probabilities of acceptance for different individuals. So, the model demonstrates strong predictive performance in distinguishing between accepted and rejected applicants.