

Homework 5B

STAT 242: Intermediate Statistics

Medical School Admissions.

The data for Medical School Admissions is in `MedGPA.csv`, taken from undergraduates from a small liberal arts school over several years. We are interested in student attributes that are associated with higher acceptance rates.

- **Accept** = accepted (A) into medical school or denied (D)
- **Acceptance** = accepted (1) into medical school or denied (0)
- **Sex** = male (M) or female (F)
- **BCPM** = GPA in natural sciences and mathematics
- **GPA** = overall GPA
- **VR** = verbal reasoning subscale score of the MCAT
- **PS** = physical sciences subscale score of the MCAT
- **WS** = writing samples subscale score of the MCAT
- **BS** = biological sciences subscale score of the MCAT
- **MCAT** = MCAT total score
- **Apps** = number of schools applied to

For this assignment, we will use **Acceptance**, **GPA**, and **Sex**.

- (a) Fit an additive logistic regression model with Acceptance as the response variable, and GPA and dichotomized Sex as explanatory variables (no interactions needed). You may assume all assumptions are reasonably satisfied.
- (b) Find the 95% confidence interval for the effect of GPA on the log odds of Acceptance based on the model in (a).
- (c) Interpret the 95% confidence interval for the effect of GPA on the log odds of Acceptance based on the model in (a).
- (d) Find the 95% confidence interval for the effect of GPA on the odds of Acceptance based on the model in (a).
- (e) Interpret the 95% confidence interval for the effect of GPA on the odds of Acceptance based on the model in (a).
- (f) Fit a logistic regression model with Acceptance as the response variable, and GPA and Sex as explanatory variables, but this time include an interaction between GPA and Sex.
- (g) Conduct a test using the summary output from your model in (f) to determine whether the effect of GPA on the log odds of Acceptance differs for males and females, as defined in this data set. Write out your null and alternative hypotheses and state your results in the context of the problem.

References

- A. Cannon et al. 2018. Package Stat2Data. “<https://cran.r-project.org/web/packages/Stat2Data/Stat2Data.pdf>”
- P. Roback and J. Legler. 2020. *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*.