

P8110: Applied Regression II

Homework #7

School administrators study the attendance behavior of high school juniors at two schools in an academic year. Predictors of the school absence rate include gender of the student and standardized test scores in math and language arts. The data are saved in “HW7.csv”. The columns of variables from left to right are:

| | |
|----------|--|
| ID | = student ID |
| school | = 1 - school A, 2 - school B |
| male | = 1 - male, 0 - female |
| math | = standardized test score in math |
| langarts | = standardized test score in language arts |
| daysabs | = days absent |

1. Fit a Poisson regression model with days of school absence as the outcome and school, gender, and standardized test scores in math and language arts as covariates (model 1). Write the model. Is overdispersion a potential problem for this Poisson model? [2 points]
2. Refit model 1 with the scale parameter being equal to Deviance divided by DF. Estimate the absence rate ratio between male and female students. Provide the 95% confidence interval and interpret.[3 points]
3. Refit model 1 using negative binomial regression. Provide a formal test to decide whether a negative binomial model is needed for this data than a Poisson regression model. [2 points]
4. Use the negative binomial model to estimate the absence rate ratio between male and female students and provide 95% confidence interval. Is the conclusion different from the Poisson model in part (2)? [3 points]