P8110: Applied Regression II Spring 2024

Homework #7 [12 points]

Due on Apr 1, 11:59**AM**

The “hwdata3.csv” contains data on 1959 patients hospitalized subsequent to undergoing either the CABG or PTCA cardiovascular procedure. The variables included are:

|  |  |
| --- | --- |
| procedure | 1=CABG, 0=PTCA |
| gender | 1=male, 0=female |
| age | age of patient |
| los | hospital length of stay(days) |
| type | 1=emergency/urgent, 0=elective |

1. Fit a Poisson regression model with hospital length of stay as the outcome and procedure, gender, age, and type as covariates (model 1). Write down the model. Is overdispersion a potential problem for this Poisson model? [2 points]

where

1. Refit model 1 with the scale parameter being equal to Pearson chi-square divided by residual DF. Estimate the length of stay rate ratio between patients undergoing CABG and PTCA procedures. Provide the 95% confidence interval and interpret. [3 points]
2. Use the fitted model in part (2), calculate the expected days of hospital stay for male patients aged 68 who underwent CABG procedure and stayed in an elective type. [2 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 length of stay rate ratio between patients undergoing CABG and PTCA procedures and provide 95% confidence interval. Is the conclusion different from the Poisson model in part (2)? [3 points]

1