Week #9 Assignment

Poisson Regression

Due: 04/04/2021, 11:59 PM

Table of Contents

* **Data Descriptions**

Daily total of bike counts conducted monthly on the Brooklyn Bridge, Manhattan Bridge, Williamsburg Bridge, and Queensboro Bridge. To keep count of cyclists entering and leaving Queens, Manhattan and Brooklyn via the East River Bridges. The Traffic Information Management System (TIMS) collects the count data. Each record represents the total number of cyclists per 24 hours at Brooklyn Bridge, Manhattan Bridge, Williamsburg Bridge, and Queensboro Bridge.

* **Data Formats and Loading**

To save your time for finding a data set for Poisson regression, I created subsets that contain the relevant information based on the combinations of bridges and months so every will have a distinct data set for this week’s assignment. The data was saved in the Excel format multiple tabs. Please find the tab with your last name and then copy-and-paste your data file to a new Excel sheet and save it as a CSV format file or simple copy-and-paste to the Notepad to create a TXT format file so you can read the file to R. You can also read the Excel file directly to R using appropriate R functions in relevant R libraries.

* **Assignment Instructions**

Your analysis and report should be similar to section 3 of my class note. PLEASE TELL STORIES BEHIND ALL R OUTPUTS (tables and figures) YOU GENERATED IN THE ANALYSIS. You can earn the half of the credit if you only generate relevant outputs correctly but with no good storytelling. *The model diagnostic is not required in this assignment but will be required in the next assignment*.

The following components must be included in you analysis report.

* Your description of the data and the variables
  + data collection
  + variable names and definition. Keep in mind that the variable **Date** is the observation ID.
* Build a Poison regression model on the counts only.
  + Use the p-values to perform variable selection - keep only significant variables in the final model
  + Interpret the regression coefficients of the Poison regression model
  + Explain/describe the steps of your analysis, motivation, and findings.
* Build a Poison regression on proportions (rates) of cyclists enter and leave the bridge in your data.
  + Use the p-values to select the significant variables
  + Interpret the regression coefficients of the Poison rate model
  + Explain/describe the steps of your analysis, motivation, and findings.
* Summarize the findings of the above two models.