STAT 796: Homework 2

Spring 2019

Due Sunday, February 3 at 11:59pm on Canvas.

For this assignment, you will analyze data from the ICU data you summarized in Homework 1. A pre-processed version of the dataset is available on Canvas in the file `icu.csv`.

- Provide R code and R output where, and only where, it is explicitly requested.
- Tables and figures should be labelled and formatted in a manner suitable for sharing with colleagues.

1. Consider the relationship between vital status (as coded by the variable `died`, with 1 indicating death during ICU stay and 0 indicating survival during ICU stay) and systolic blood pressure (`sbp`) at time of admission.
   a. Provide a plot that jointly summarizes these two variables. (You can choose boxplot, scatterplot, density plot, histogram, etc.). Explain what trends in the data you can identify from the plot.
   b. Fit a logistic regression model using vital status as the outcome and systolic blood pressure (SBP) as the predictor of interest. Provide the estimated regression coefficients.
   c. What is the estimated odds ratio for dying during an ICU visit, comparing those who differ in SBP at admission by 1 mmHg?
   d. What is the estimated odds ratio for dying during an ICU visit, comparing those who differ in SBP at admission by 15 mmHg?
   e. Provide the code you used for parts (b), (c), and (d) as an Appendix after all questions.
2. Consider the relationship between vital status (`died`) and visit type (`type`).
   a. Fill in the following 2x2 table with the number of subjects having each combination of these two variables:

Table 1: Counts of Vital Status by visit type.

	Elective Visit	Emergency Visit
Survived		
Died		

   b. Of those whose visit was elective, what proportion died?
   c. Of those whose visit was an emergency, what proportion died?
   d. Fit a logistic regression model using vital status as the outcome and visit type as the predictor of interest. Provide the coefficient estimates from the fit model.
   e. What is the estimated odds ratio for dying during an ICU visit, comparing those whose visit was an emergency to those whose was elective.
   f. What is the predicted risk of dying during an ICU visit for people whose procedure is elective?
   g. What is the predicted risk of dying during an ICU visit for people whose procedure is an emergency?
   h. How do your answers for (f) and (g) compare to (b) and (c)?
   i. Provide the code you used for parts (d)-(g) as an Appendix after all questions.
3. Again, consider the relationship between vital status and visit type.
   a. Create a new variable `emergency` that takes the value of 1 when `type` is `emergency` and takes the value of 0 when `type` is `elective`. (In other words, create an indicator variable for `type` being `emergency`.)
   b. Fit a logistic regression model using `emergency` as the outcome and vital status as the predictor of interest. Provide the coefficient estimates from the fit model.
   c. What is the estimated odds ratio for the procedure being emergency, comparing those who died during their ICU visit to those who did not.
   d. How does your response to (c) compare to part (e) of Question 2?
   e. Provide the code you used for parts (a)-(c) as an Appendix after all questions.