MATH 4044 – Statistics for Data Science

Practical Week 13

Question 1

The sinking of the Titanic is a famous event. Many well-known facts – from the proportions of first-class passengers to the 'women and children first' policy, and the fact that that policy was not entirely successful in saving the women and children in the third class – are reflected in the survival rates for various classes of passengers. These data were originally collected by the British Board of Trade in their investigation of the sinking. There is no complete agreement among primary sources as to the exact numbers on board, the number rescued, and the number lost.

The data for this question is stored in a SAS data file called `titanic.sas7bdat`. For each person on board the fatal maiden voyage of the ocean liner Titanic, this dataset records gender, age [adult/child], economic status [first/second/third class, or crew] and whether or not that person survived. Specifically, variables in the data file are as follows:

Variable	Description
Class	0 = 'crew', 1 = 'first', 2 = 'second', 3= 'third'
Age	1 = 'adult', 0 = 'child'
Gender	1 = 'male', 0 = 'female'
Survived	1 = 'Yes', 0 = 'No'

[Source: *Journal of Statistics Education* data archive.]

(a) Obtain mosaic plots of Survived versus Gender, Age and Class. Also consider tests of independence. Comment on chances of survival by gender, age and class. How well did the 'women and children first' policy work?

(b) Fit and interpret a logistic model for the probability of surviving the Titanic disaster with three main effects of Gender, Age and Class.

(c) Now fit and interpret a model with both main effects and interactions. Specifically, start with a model that includes the same three variables and all possible two-way interactions, and use a backwards elimination technique.

Note: A backwards elimination method will produce what is called a hierarchical model. In this kind of model, main effects cannot be removed from the model if these effects are involved in an interaction that remains in the model.