Business Analytics (110-1)

Assignment 4

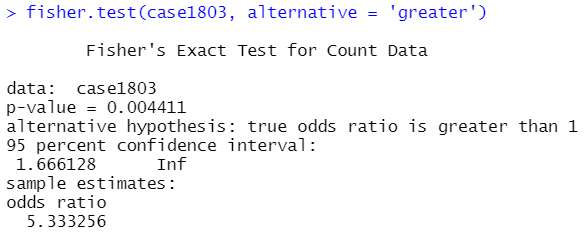
Due: 9:00 am, Tue 21-Dec-2021

1.

For the data in “Smoking & Lung Cancer” case, test whether the odds of lung cancer for smokers are equal to the odds of lung cancer for nonsmokers, using Fisher’s Exact Test.

H0: phi = 1

H1: phi > 1

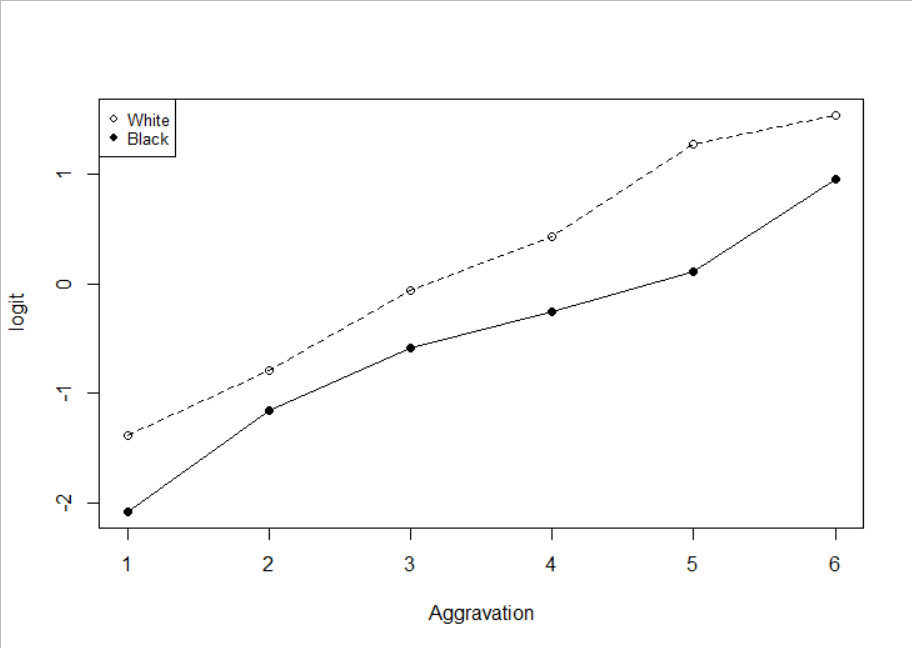


觀察圖中fisher’s exact test結果可知p-value = 0.004411 < 0.05，我們可以拒絕H0，得知odds ratio != 1，兩組的odds並不相等。

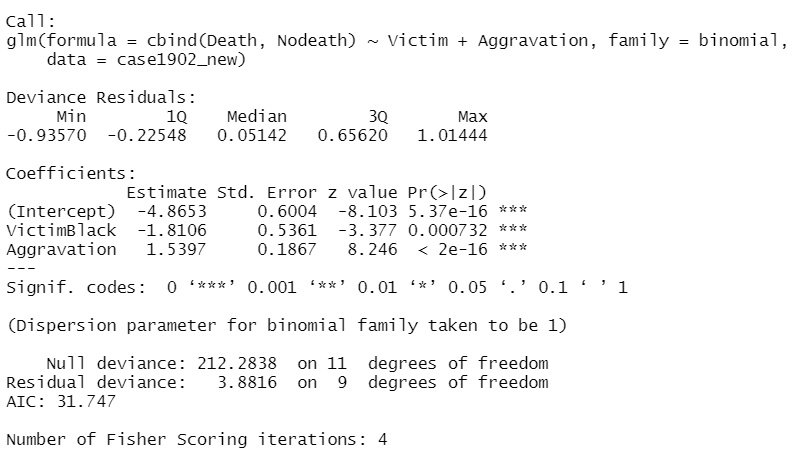
2.

Reconsider the case Death Penalty and Race of Murder Victim discussed in lecture 10 (case1902). Reanalyze the data **using logistic regression**. The response variable is the number of convicted murderers in each category who receive the death sentence, out of the **m convicted murderers** in that category.

1. Plot the logits of the observed proportions versus the level of aggravation. The logit, however, is undefined for the rows where the proportion is 0 or 1, so **compute the empirical logit = log[(y + 0.5) / (m – y + 0.5)] and plot this versus aggravation level**, using different plotting symbols to distinguish proportions based on white and black victims.

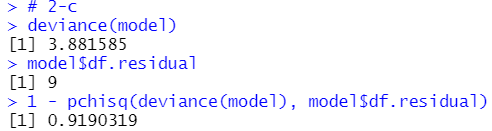


1. Fit **the logistic regression of death sentence proportions** on aggravation level and an indicator variable for race of victim.



1. Report the p-value from the deviance goodness-of-fit test for this fit.

H0: the chosen Logistic GLM holds

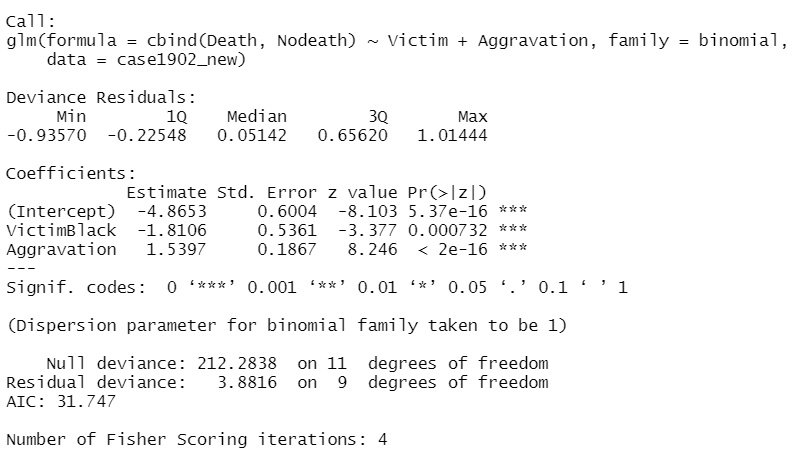


p-value = 0.919 > 0.05, there is no evidence that the model is inadequate

1. Test whether the coefficient of the indicator variable for race is equal to 0, using Wald’s test.

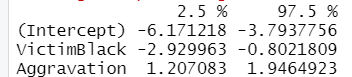
H0: Beta = 0

H1: Beta != 0



VictimBlack的p-value為0.000732 < 0.05，因此可以推翻H0，coefficient of the indicator variable for race is not equal to 0.

1. Construct a confidence interval for the same coefficient, and **interpret it in a sentence about the odds of death sentence for white-victim murderers relative to black-victim murderers**, accounting for aggravation level of crime.



1. Refit the model by treating the aggravation level as a factor. How would you interpret the results of this model?

