

Quantitative Methods II

Assignment 10

October 30, 2011

Usually, assignments in this course give you practice in the methods introduced in the course the previous week. This assignment does that, but it also has you read a published paper and analyze it. I chose a subject away from FEMA so that you could focus on the methods used and not the theory tested.

The second problem has you determine error rates for a model you create. As such, this problem melds last week with this week. You will be creating a model to predict on a binary dependent variable. You will also compare predictions from your model to reality, as provided by the data available.

When you turn in this assignment, you will email to me the usual two files: your write-up and your R script.

As always, the email needs to include the subject line:

POLS6123: Assignment 10

Also as always, if you have problems, let me know. I am happy helping.

PROBLEM: FAIR SEAFARERS?

[[10]]

This week has you read an article that examines a binary choice model (for seafarers, in this case). While the topic has little to do with fire or with emergency management, it does show a use for binary dependent variable regression we did this week and last. Furthermore, while the article concerns itself with ships, the structure of the paper and the choice model used in the paper are very applicable to everyone in this course.

When I go to conferences to share my research findings, I make it a point to go to paper sessions that have nothing to do with my substantive area. I do this so I can focus on the theories and the methods they use. Several times I sat in a section that dealt with who-knows-what when I realized that the theory the researcher was discussing applied to my current research. As such, reading and thinking about this article will give you practice in ignoring the discipline-specific concepts and focus on the important parts.

The assignment is for you to outline the statistical method used by the authors. Explain why they had to use that method. State the model equation they formed. State which of the independent variables were statistically significant. Write a short paragraph explaining whether or not you think this method used was appropriate.

- Reference: Ding, Ji-Feng and Gin-Shuh Liang (2005). “The choices of employing seafarers for the national shipowners in Taiwan: an empirical study.” *Maritime Policy Management* 32(2): 123 – 37.

PROBLEM: THAT FLIPPIN' COIN!

[[15]]

Let us reuse the `coinflips` data file. There are two variables: coin number and head. The data is described in the text. Using whatever appropriate method, model the probability of a head based on the coin number. (Not too difficult.)

- (1) What is the proportional reduction in deviance for your model? Does this indicate you have a good model?
- (2) How accurate (absolute accuracy) is your model when using the default threshold of $\tau = 0.500$?
- (3) What is the relative accuracy of your model (with a threshold of $\tau = 0.500$)?
- (4) Which threshold provides the best accuracy, τ_0 ?
- (5) What is the false positive rate (FPR) for the τ_0 threshold?
- (6) What is the false negative rate (FNR) for the τ_0 threshold?
- (7) Produce a ROC curve and appropriately interpret A' , the area under the ROC curve for your model.
- (8) Compare your response regarding the proportional reduction in deviance (above) with your interpretation of A' .