STAT 8004 – Statistical Methods II

Spring 2015

Homework Assignment 9 (Due on 4/9/2015 before the end of the day)

- Reading assignment
  - All in-class coverage about generalized linear models
  - Please read material in Chapter 6 of Faraway (2006) on R Generalized linear models.
- Comprehensive coverage for generalized linear models is
  - McCullagh, P. and Nelder, J. A. (1989). Generalized Linear Models. Chapman and Hall.
- The following exercises are to be collected. Please upload your homework to the Blackboard. Following the requirement of STAT 8003, please typeset your homework with Latex and upload both the pdf and latex files.

1. Data is generated from the exponential distribution with density $f(y) = \lambda \exp(-\lambda y)$ where $\lambda, y > 0$.
   (a) Show that it belongs to the exponential family distributions by identifying θ, $b(\theta)$, ϕ, $a(\phi)$ and $c(y; \phi)$.
   (b) What is the canonical link and variance functions for a GLM with the response following the exponential distribution?
   (c) Is there any practical difficulty for using the canonical link in practice?
   (d) Express the deviance as a function of y_i and fitted mean μ_i $(i = 1, \ldots, n)$.

2. Consider the Orings data of Faraway (2006) where the number of damaged ones out of six orings and corresponding temperatures of space shuttle launches are recorded.

```
library(faraway)
data(orings)
```

   (a) Construct the appropriate test statistic for testing the effect of the temperature. State the appropriate null distribution and give the p-value.
   (b) Does it affect the conclusion by changing the link function to probit and other link functions.
   (c) Creating a new column of response as the indicator on whether or not there is some oring damage in that launch. Refit this binary response to a GLM with the logit link.
   (d) Which model do you prefer, the one in part (a) or part (c)? Why?