STAT 203 PROBLEM SET 4

Due date: Thursday, March 5.

(1) The O-rings in the booster rockets used in space launching play an important part in preventing rockets from exploding. Probabilities of O-ring failures are thought to be related to temperature. A detailed discussion of the background of the problem is found in the Flight of the Space Shuttle Challenger (pp. 33-35) in Chatterjee, Handcock, and Simonoff (1995). Each flight has six O-rings that could potentially be damaged in a particular flight. The data from 23 flights are given in the data file *Orings.txt*. For each flight, we have the number of O-rings damaged (out of 6) and the temperature of the launch.
   (a) Fit a logistic regression connecting the probability of an O-ring failure with temperature. Interpret the coefficients.
   (b) The data for Flight 18 that was launched when the launch temperature was 75 was thought to be problematic, and was deleted. Fit a logistic regression to the reduced data set, and interpret the coefficients.
   (c) From the fitted model, find the probability of an O-ring failure when the temperature at launch was 31 degrees. This was the temperature forecast for the day of the launching of the fatal Challenger flight on Jan 20, 1986.
   (d) Would you have advised the launching on that particular day?

(2) Field-goal-kicking data for the entire American Football League (AFL) and National Football League (NFL) for the 1969 season are given in the data file *NFL.txt*. Let $\pi(X)$ denote the probability of kicking a field goal from a distance of X yards.
   (a) For each of the leagues, fit the model

$$\pi(X) = \frac{e^{\beta_0+\beta_1 X+\beta_2 X^2}}{1+e^{\beta_0+\beta_1 X+\beta_2 X^2}}.$$

   (b) Let Z be an indicator variable representing the league, that is,

$$Z = \begin{cases} 1, & \text{for the AFL;} \\ 0, & \text{for the NFL.} \end{cases}$$

   Fit a single model combining the data from both leagues by extending the model to include the indicator variable Z; that

is, fit

$$\pi(X, Z) = \frac{e^{\beta_0+\beta_1 X+\beta_2 X^2+\beta_3 Z}}{1 + e^{\beta_0+\beta_1 X+\beta_2 X^2+\beta_3 Z}}.$$

(c) Does the quadratic term contribute significantly to the model?
(d) Are the probabilities of scoring field goals from a given distance the same for each league?

(3) A researcher in geriatrics designed a prospective study to investigate the effects of two interventions on the frequency of falls. One hundred subjects were randomly assigned to one of the two interventions: education only ($X_1 = 0$) and education plus aerobic exercise training ($X_1 = 1$). Subjects were at least 65 years of age and in reasonably good health. Three variables considered to be important as control variables were gender (X_2 : 0 = female; 1 = male), a balance index (X_3), and a strength index (X_4). The higher the balance index, the more stable is the subject; and the higher the strength index, the stronger the subject. Each subject kept a diary recording the number of falls (Y) during the six months of the study.
(a) Fit a Poisson regression model with the response function

$$\mu(X, \beta) = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4).$$

State the estimated regression coefficients, their estimated standard deviations, and the estimated response function.
(b) Obtain the model deviance and test the goodness of fit of the Poisson regression model at level $\alpha = 0.05$. State the alternatives, decision rule, and conclusion.
(c) Assuming that the fitted model is appropriate, use the relevant partial deviance to test whether gender (X_2) can be dropped from the model; control α at level 0.05. State the alternatives, decision rule and conclusion.