Business Analytics (108-1)

Assignment 5

1.

For the *Moth Coloration* data (`case2102`), consider the response count to be the number of light months removed and the binomial denominator to be the total number of months removed (light and dark) at each location.

(a) Fit the logistic regression model with distance as the explanatory variable; then report the estimates and standard errors.

(b) Compute the deviance goodness-of-fit test statistic, and obtain a p-value for testing the adequacy of the model.

2.

The dataset `vitaminC` is from a randomized experiment to assess the effect of large doses of vitamin C on the incidence of colds. The subjects were given tablets to take daily, but neither the subjects nor the doctors who evaluated them were aware of the dose of vitamin C contained in the tablets. The data shows the proportion of subjects in each of the four dose categories who did not report any illness during the study period.

(a) Fit the logistic regression model. Report the estimated coefficients and their standard errors. Summarize the findings of this model.

(b) What can be concluded about the adequacy of the logistic regression model? What evidence is there that the odds of a cold are associated with the dose of vitamin C?

3.

Data `ex1028` in Sleuth2 package shows the number of Atlantic Basin tropical storms and hurricanes for each year from 1950 to 1997. The variable *storm index* is an index of overall intensity of the hurricane season. (It is the average of number of tropical storm, number of hurricanes, the number of days of tropical storms, the number of days of hurricanes, the total number of intense hurricanes, and the number of days they last – when each of these is expressed as a percentage of the average value for that variable. A *storm index* score of 100, therefore, represents, essentially, an average hurricane year.) Also listed are whether the year was cold, warm, or neutral El Nino year, a constructed numerical variable temperature that takes on the values -1, 0, and 1 according to whether the El Nino temperature is cold, neutral, or warm; and a variable indicating whether West Africa was wet or dry that year. It is though that wet years in West Africa often bring more hurricanes.

Use Poisson log-linear regression to describe the distribution of (a) number of storms and (b) number of hurricanes as a function of El Nino temperature and West Africa wetness.