Stat 545 Part II Homework # 2

Fall 2015 Rice University

Instructor:	Liang Li <LLi15@mdanderson.org>
Teaching Assistant:	Yinsen Miao <ym11@rice.edu>
Date of Assignment:	Nov 4, 2015
Due Date:	November 18, 2015

Instructions: Your answer to Problems # 3 and # 4 must be prepared in Microsoft Word (with Equation Editor) or Latex. Please include your SAS (preferred) or R code, copy and paste the raw output from the software, and interpret the results.

Problem # 1 (*10 points*). Please finish Exercise 8.32 on page 336 of the textbook.

Problem # 2 (*10 points*). Please finish Exercise 8.37 on page 337 of the textbook.

Problem # 3 (*20 points*). The data set horseshoecrab.csv is the data used for the example in Table 4.4 of the textbook. In this data set, `satell` is the number of satellites, `weight_kg` is the weight in kilograms. Fit a Poisson regression model with `satell` as the outcome and `weight_kg` as the covariate.

(a) Estimate the effect of weight, and interpret it.

(b) Is there any evidence of overdispersion. Explain.

(c) Fit the same Poisson model but with a scale parameter adjusting for overdispersion. Compare the result with (a). Explain the differences.

(d) Fit a negative binomial regression, and compare the result with that in (a)

Problem # 4 (*20 points*). The data set GSS.csv is the data used for the example in Table 8.5 of the textbook. In this data set, `trauma` is the number of traumatic events, `happy` is 3 = not too happy, 2 = pretty happy, 1 = very happy. `race` is 0=white, 1=black. Fit a cumulative logit model with `happy` as the outcome and `trauma` and `race` as covariates.

(a) Estimate the effect of each covariate, and provide explanation.

(b) How many intercepts are there in the model? How to interpret them.

(c) Provide a latent variable interpretation of the cumulative logit model. What are the parameters in that latent variable model?

(d) Plot the predicted probabilities of the three levels of `happy`, as a function of both `trauma` and `race`, similar to Figure 8.6.