Stat 545 Part II Homework # 1

Fall 2015 Rice University

Instructor:	Liang Li <LLi15@mdanderson.org>
Teaching Assistant:	Yinsen Miao <ym11@rice.edu>
Date of Assignment:	October 26, 2015
Due Date:	November 11, 2015

Instructions: Your answer to Problem # 1 must be prepared in Microsoft Word (with Equation Editor) or Latex. Please include your SAS code, copy and paste the raw output from the software, and interpret the result. The plot should also be prepared using software.

Problem # 1. The data set basketball.csv shows, for each game in the 2010-2011 season of the National Basketball Association in which Rajon Rondo of the Boston Celtics played. Let x = the number of assists he recorded and y = whether the Celtics won (1 = yes). Using software, do the following

(a) show that the logistic model fitted to these data gives $logit[\hat{P}(Y = 1)] = -2.235 + 0.294x$

(b) show that $\hat{P}(Y = 1)$ increases from 0.21 to 0.99 over the observed range of x from 3 to 24

(c) construct a $\alpha = 0.05$ level significance test and 95% confidence interval about the effect of x in the conceptual population that these games represent

(d) Plot a figure of the fitted model, where the horizontal axis is the number of assists, and the vertical axis is the probability of winning the game.

(e) Perform the Hosmer-Lemeshow test to check the goodness of fit

(f) Perform a likelihood ratio test on the effect of x. Report the two log-likelihood values, test statistic, degree of freedom, and p-value.

Problem # 2. Prove that logistic regression model (5.1) has the steepest slope where $\pi(x) = 1/2$. Generalize the result to the model with multiple covariates.

Problem # 3. Construct the log-likelihood function for the model $logit[\pi(x)] = \alpha + \beta x$ with independent binomial outcomes of y_0 successes in n_0 trials at $x = 0$ and y_1 successes in

n_1 trials at $x = 1$. Derive the likelihood equations, and show that $\hat{\beta}$ is the sample log odds ratio.

Problem # 4. Consider the likelihood equations (5.18) for a logistic regression model. Using the equation resulting from the intercept parameter, show that the overall sample proportion of successes equals the sample mean of the fitted success probabilities.

Problem # 5. Please finish Exercise 5.18 on page 202 of the textbook.