STA 580 — Spring 2011 — Dr. Charnigo

Written Assignment 6

This assignment is due on Thursday 28 April at 5:30 p.m. You may work in self-selected groups of two or three, in which case you may hand in one copy of the assignment for the group.

[70] 1. In Lecture 7 we examined the data set {FEV.xls} to determine how forced expiratory volume might be related to smoking. Now you will examine this data set to determine how forced expiratory volume might be related to age. Let X denote age and Y denote forced expiratory volume. Students in Spring 2009 considered a linear regression model of the form

$$Y_i = \alpha + \beta x_i + \epsilon_i.$$

In contrast, you will consider a linear regression model of the form

$$\log Y_i = \alpha + \beta \log x_i + \epsilon_i.$$

To do so, put $v_i := \log y_i$ and $u_i := \log x_i$. Then replace the x_i's and y_i's in the usual formulas with u_i's and v_i's respectively. You may proceed as if the ϵ_i were independent normal random variables with mean 0 and unknown but common variance σ^2.

[10] a. Report the least squares estimates of α and β.

[10] b. Report 95% confidence intervals for α and β.

[10] c. Test $H_0 : \beta = 0$ by constructing an ANOVA table and calculating an f statistic.

[10] d. Test $H_0 : \beta = 0$ by calculating a t statistic based on the least squares estimate of β.

[10] e. What fraction of the variability in the logarithm of forced expiratory volume is accounted for by the logarithm of age?

[10] f. Define $\widehat{y}_i := \exp[\widehat{v}_i]$. Provide an explicit formula for $\widehat{y}_i$, simplified as much as possible, in terms of x_i and the least squares estimates of α and β.

[10] g. Suppose that forced expiratory volume will be measured tomorrow for a child who is 11 years old. Noting that $\log 11 = 2.398$, obtain a 95% prediction interval for the logarithm of the forced expiratory volume of this child. Then exponentiate the endpoints of that interval to obtain a 95% prediction interval for the forced expiratory volume of this child. Compare this prediction interval to that calculated by students in Spring 2009.

[30] 2. Reconsider "Infectious Disease" on page 454, previously the subject of an exercise in Written Assignment 4.

[10] a. Report point and 95% interval estimates for the risk difference $p_1 - p_2$.

[10] b. Report point and 95% interval estimates for the relative risk p_1/p_2.

[10] c. Report point and 95% interval estimates for the odds ratio $\{p_1/(1-p_1)\}/\{p_2/(1-p_2)\}$.