Statistics 659 - Assignment 8
(due Friday, April 1, 2015, 11:59pm)

Instructions:

- Whether you write out the solutions by hand or in a text document, be sure that they are *neat, legible and in order* (even if you choose to solve them in a different order).

- **Type** your name, email address, course number, section number and assignment number at the top of the first page (or cover page).

- Either scan or print your solutions to a **PDF** file under 15MB in size. It must be in a *single* file, not separate files for separate pages. Name the file using your name (for example, I could use twehrly659hw01.pdf) to avoid confusion with other students and/or assignments. *Do not* take a photo of each page and then paste them into a document – this will make your file too big and the results will generally not be very readable anyway.

- Login to your WebAssign account to upload your file. You must do this by **11:59 pm U.S. Central time**, *according to the WebAssign server*, on the due date. We highly recommend that you start the upload at least 15 minutes earlier. You can make multiple submissions but *only the last submission will be graded*.

This assignment covers the material in Chapter 5 discussed in Lectures 23 through 26.

Assess the fit of the model that you chose in Problem 5.13 in Homework 7.

5.9ad, 5.10c, 5.14, 5.16 (fit the data for problem 4.30), 5.18 (study=city in part a), 5.19, 5.21, 5.22 (use the malformation data from Chapter 2), 5.23, 5.24

Continuation of Variable Selection Problem

Hosmer and Lemeshow's ICU study considers a sample of 200 subjects who were part of a larger study on the survival of patients following admission to an adult intensive care unit (ICU). The goal is to develop a logistic regression model to predict the probability of survival to hospital discharge of these patients.

Reference: Lemeshow et al. (1988), Journal of the American Statistical Association, 348-356.

4. Check for interactions among the variables in the preliminary main effects model that you found in part 2. of this problems. Use the approach outlined in class where you (i) check for interactions added one at a time to the model, (ii) fit a model with the interactions determined in step (i), (iii) eliminate nonsignificant interactions one at a time until all remaining interactions are significant.

5. Assess the fit of the model you developed in Part 4. This should include an overall assessment of fit and also the use of the diagnostic statistics. Identify any subjects (or covariate patterns) that are poorly fit or overly influential.

6. Construct marginal model plots to check the linearity of age and sys in the model that you found in part 4. If you did not find a model that included these variables, use the model `sta=age sys age*sys can typ loc`.

(Only for students having taken STAT 414, 610 or STAT 630) 5.25