Homework 1: due Tuesday, 13 September

Please attach a printout of your SAS program(s) as an appendix to your HW paper.

I expect each student to write their own code. Using someone else's SAS program is academic dishonesty.

Homework is due at the start of class. We will discuss the homework problems as soon as they have been collected. Please photocopy your homework paper before class so you'll be able to check your answers.

Please do not attach extra output at the back of your homework—attach it to your own copy of the homework for reference.

1. **Child IQ.** For this problem, please download the spreadsheet ChildIQ.xls from the course website. and import the data into SAS (see *LSB* §2.3).

   a. Fit the regression of child's IQ (the response) on these predictors: mother's IQ, mother's age, and mother's high-school education status. (This is the model in the SAS program for the first class.)

      Give the output table of regression coefficients, standard errors, and the associated tests. Which predictors are not significant?

   b. Drop the non-significant predictors from the model and re-fit the regression. Give the output table of regression coefficients, standard errors, and the associated tests. Write a sentence interpreting each regression coefficient in terms of child's IQ.

   c. Print and attach the "Fit Diagnostics" panel of plots for your model in (b). The first two plots show residuals and studentized residuals against predicted values. How are these two plots supposed to look if the usual assumptions for linear regression hold ?

      Do they show any problems?

   d. Find Spearman correlation between mother IQ and child IQ. Use Google ("SAS Spearman correlation") to find out how to do this.

2. **Blood lead studies.** Children in a Minneapolis neighborhood were recruited into two different studies simultaneously. The first study, called LPP, was a trial of education to prevent lead poisoning; children were randomized to education (intervention = 1) or control (intervention = 0). The second study, called PBIQ, was an observational study of the effects of lead on child's IQ. Indicator variables LPP and PBIQ identify those recruited to each study.

For this problem, please download the spreadsheet Lead-data.xls from the course website. and import the data into SAS.

   a. How many children were in LPP alone, PBIQ alone, and in both?

   b. Compare mean birthweight between males and females in the PBIQ study.