Textbook:
   Chapter 12: #16, 21 (read part (a) but there’s nothing you need to do for it),
   28(b), 35 (33, sixth edition), 40, 44, 49, 53, 59 (skip part e), 70.
   Chapter 13: #9, 19, 29.

P.S. In reading Chapter 13.1 – 13.3 you should skip the subsection on “Logistic
Regression”. We’ll come back to this topic at the end of the semester.

Extra Problems:
   A. It is sometimes claimed that giving school-wide tests in elementary school and
   publicizing the results of the tests causes the poorer performing schools to do better the
   following year. The data set “3rd grade reading scores” contains class-wide averages for
   test results from 40 schools in a metropolitan area. The tests were given in 1982. These
   results were publicized for each school. The test was then given again (in an alternate
   form) in these same schools 1983. To see whether the claim might be correct the
   following analysis was performed:
      The 20 schools that scored in the lower half of all 1982 scores were chosen. Their
      reading scores were examined again in 1983. The following table summarizes some of
      the results for
      “Difference” = 1983 score – 1982 score
      for these 20 schools. Note that the mean difference is positive and the difference is
      statistically significant at the level alpha = 0.05, and somewhat below this as well.
      
      | Moment for “Difference” | 11.10 |
      |-------------------------|-------|
      | Mean                    | 11.10 |
      | Std Dev                 | 19.24 |
      | Std Err Mean            | 4.303 |
      | upper 95% Mean          | 20.11 |
      | lower 95% Mean          | 2.094 |
      | N                       | 20    |

      Does this significantly positive Difference tend to confirm the claim made at the
      beginning of this problem? Explain your answer carefully. The data is accessible on our
      website if you want to examine it and/or perform other statistical procedures. (Obviously
      this is only a small data set, so results here could not come near to proving or disproving
      the claim; I want to know only whether they tend to confirm the claim.)
      Can you think of a better way to conduct a study and/or to analyze the resulting
      data that would provide a better type of evidence for or against the claim? (Again, it’s
      obvious that more data would be better; that’s not the type of answer I’m looking for.)

   B. The data set “NBA salaries” gives the starting salaries of NBA first round draft picks in
   1991. Picks #15 and 25 did not sign first year contracts in the NBA. We’ll focus on the
   question of how large a salary pick #15 should receive, based just on this data.
      a. Begin by doing the usual linear regression analysis. (Use JMP.) Give the scatterplot
      and the usual regression table.
b. What is the 95% confidence interval for the salary a typical pick #15 should expect to receive? What is the appropriate 95% interval for the salary that this particular pick #15 would receive?

c. The linear regression fits fairly well, and the linear regression analysis seems reasonably well justified. But there is/are some indication(s) that transformation and/or alternate analyses of the data might improve the linear regression analysis. What is/are the indication(s)?

d. I tried three reasonably successful alternate analyses. #A.1 involved the transformation $Y \rightarrow \log(Y)$. #A.2 was a quadratic regression. #A.3 involved the transformation $X \rightarrow \log(X)$. Perhaps there are some other good possibilities as well. Choose at least one of these 3 – one you think should perform very well, or best. (Or you may use a fourth style if you feel it may be preferable.) Give the resulting regression table, and also use your alternate analysis to re-answer question b. Is your new answer more appropriate than the original one? Why?

e. Suppose the goal had instead been to predict the salary for pick #25. Would that have changed your answer to part d? In particular, would the quadratic regression suggested for choice A.2 have been a good idea for this purpose?

-- END --