- 1. Problem 2.1 from KNNL
- 2. For this problem use the "plastic hardness" data described in the text of the problem 1.22 on page 36. It is available from the file CH01PR22.txt on the course webpage. Make sure you understand which column is X and which is Y and read in the data accordingly.
- a) Generate a scatterplot using your statistical software (PROC GPLOT in SAS). Include the fitted regression line on the plot (i=rl option). Does it appear that the relationship between these two variables is approximately linear? Explain your answer.
- b) Run the linear regression to predict hardness from time and state the estimated regression equation. Give a 95% confidence interval for the slope. Explain to someone not familiar with statistics what this confidence interval means. Remember that time is measured in hours and hardness is measured in Brinell units.
- c) Describe the results of the significance test for the slope that you get in your software output. State the hypotheses being tested, the test statistic with degrees of freedom, the P-value, and your conclusion in a brief sentence.
- d) Explain why or why not inference on the intercept is reasonable (i.e., of interest) in this case.
- e) Continue with the same dataset. Give an estimate of the **mean** hardness that you would expect after 36 and 43 hours, and a 95% confidence interval for each estimate. Which confidence interval is wider and why is it wider?
- f) Again, using the same dataset, give a prediction for the hardness that you would expect for an **individual** piece of plastic after 43 hours; and a 95% prediction interval for this quantity.

3. You fit the simple linear model to a data set and obtain estimates b0 = 1, b1 = 3, and s = 3.0

(a) Suppose the data consisted of n = 18 cases and the standard deviation of b1, s(b1), is equal to 1. Construct the 95% CI for  $\beta_1$  (HINT: Use Table B.2 to get the critical value tc)

(b) Do you have statistical evidence that X helps predict Y? Explain.

(c) The 95% confidence interval for E(Y) when X = 5 is [13.88, 18.12]. Find the 95% prediction interval when X = 5.

4. KNNL Problem 2.26 (do not do parts b) and d); you can reply to just the first question in part c))