

# Statistics 431: Statistical Inference

## Fall 2006

### Problem Set 7

Due 13 Nov 2006

---

Devore Chapter 13: #8, 10, 18, 20, 26, 30, 34.

Additionally:

A. The data set “NBA salaries” (`problem-set-07-nba-salaries.jmp`) 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  $Y \mapsto \log(Y)$ . #A.2 was a quadratic regression. #A.3 involved the transformation  $X \mapsto \log(X)$ . There are other possibilities. Choose at least one of these three – one you think should perform well, or best. (Or, use a fourth style if you think it 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?