

Problem set 11

S520

Upload your answers through the Assignments tab on Canvas by 11:59 pm, Thursday 30th November. Draw graphs in R and include code.

Trosset question numbers refer to the hardcover textbook. Draw all graphs in R and include all R code. You may work with others, but you must write up your homework independently — you should not have whole sentences in common with other students or other sources.

1. (5 points.) Trosset chapter 14.6 exercise 3.

Data: <https://mtrosset.pages.iu.edu/StatInfer/Data/unusual.dat> . You can read it in with:

```
unusual <- matrix(scan("https://mtrosset.pages.iu.edu/StatInfer/Data/unusual.dat"),
  ncol = 2, byrow = TRUE)
```

2. (5 points.) The psychologists Daniel Kahneman and Amos Tversky described the following situation:

The instructors in a flight school adopted a policy of consistent positive reinforcement recommended by psychologists. They verbally reinforced each successful execution of a flight maneuver. After some experience with this training approach, the instructors claimed that contrary to psychological doctrine, high praise for good execution of complex maneuvers typically results in a decrement of performance on the next try.¹

Is there a simpler explanation for the decreased performance following praise? (Hint: This is the regression chapter, so it probably has something to do with that.)

3. (5 points.) Trosset chapter 15.7 exercise 8.
4. (5 points.) A Major League Baseball team plays 162 games each season. There are 30 teams. Each season, the number of wins by Major League Baseball teams has an approximately normal distribution with mean 81 and standard deviation 11.7. The correlation between a team's wins one season and their wins the next season is 0.54.
 - (a) Suppose a baseball executive believed the best prediction of a team's wins in 2015 should be equal to their wins in 2014. For example, he predicts that the Los Angeles Angels, who had the most wins in 2014 with 98, would have 98 wins in 2015. Using the data given, explain to the baseball executive (who knows very little statistics) why this particular prediction is likely too high.

¹Reprinted in *Judgement Under Uncertainty: Heuristics and Biases* (1982).

- (b) Use regression to predict the Los Angeles Angels' 2015 wins using only the above data.
 - (c) The executive looks at the regression predictions for all MLB players and sees that no team is predicted to win more than 91 games. The executive suspects the predictions are too low, because in every full season since 1961, at least one team has won at least 96 games. Explain to the executive, who knows very little statistics, why his suspicions are misplaced.
5. (10 points.) Using the file `adults.txt`, which gives heights (in cm) and weights (in kg) of U.S. adults:
- (a) Draw a scatterplot with `Height` on the x -axis and `Weight` on the y -axis, and add the least squares regression line.
 - (b) Draw a scatterplot with `Weight` on the x -axis and `Height` on the y -axis, and add the least squares regression line.
 - (c) Predict the weight of a U.S. adult who is 180 cm tall. Call this predicted weight w .
 - (d) Predict the height of a U.S. adult who weighs w kg.
 - (e) Is your answer to (d) 180 cm? If not, explain why not.