

# Homework 12

S520

**Upload your answers as a PDF or HTML file through the Assignments tab on Canvas by 4 pm, Thursday 28th April. No late homeworks at all will be accepted this week.**

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. 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.<sup>1</sup>

Is there a simpler explanation for the decreased performance following praise? What does this have to do with chapter 15?

2. Trosset chapter 15.7 exercise 4
3. Trosset chapter 15.7 exercise 8
4. Apply the bivariate normal to the baseball wins data in `baseball-wins.txt` to estimate the following:
  - (a) The probability that a randomly selected team wins at least 84.5 games.
  - (b) The probability that a team that won 95 games one season wins at least 84.5 games the next season.
  - (c) The probability that a team that won 75 games one season wins at least 84.5 games the next season.
5. The file `examanxiety.txt` on Canvas contains information on a number of variables:
  - **Exam**: score on a math exam
  - **Revise**: hours spend revising for the math exam
  - **Anxiety**: “math anxiety” on a scale from 0 to 100 (100 is most anxious)

---

<sup>1</sup>Reprinted in *Judgement Under Uncertainty: Heuristics and Biases* (1982).

- (a) Find the regression line to predict exam score from anxiety. Write down your answer as an equation (do not just paste R output.)
- (b) Which of the following regression assumptions are met?
- Linearity
  - Independence
  - Equal variance (homoskedasticity)
  - Normality of errors
- (c) Suppose we want to make probabilistic predictions of a student's exam score given their math anxiety. Should we use the bivariate normal? Why or why not?
6. When 100 children at a certain school begin a grade, they're given an aptitude test. The top 20% go to class A, the next 20% to class B, and so on down to the bottom 20%, who go to class E. At the end of the year, they're given a similar aptitude test. The data is given in the file `testscores.txt`.

The Board of Education, none of whom have taken a statistics course, wants to measure teacher performance by looking at the average change (second test minus first test.) They get the following results:

- Class A: Average change:  $-2.15$
- Class B: Average change:  $3.45$
- Class C: Average change:  $7.6$
- Class D: Average change:  $14.5$
- Class E: Average change:  $21.9$

Based on these results, the Board is deeply unhappy with the teacher of Class A. But did Class A really perform the worst compared to expectations? Do calculations and give a conclusion that the Board can understand.