Homework 6

S320/520

Due by 4pm, Thursday 15th October

Please write "S320" or "S520" at the top of your homework. Trosset question numbers refer to the hardcover textbook. Include R code as an appendix to your answers. Data for the questions from chapter 7 can be found at

`http://mypage.iu.edu/~mtrosset/StatInfeR.html`

For this homework, you may either upload your answers to Canvas (recommended) or submit on paper in class. If you upload to Canvas, your answers must be typed and in PDF format.

1. Trosset exercise 7.7.4.
2. Trosset exercise 7.7.6.
3. Trosset exercise 8.4.4.
4. In the Powerball lottery, there are 59 white balls, numbered 1 to 59. Each week, five of the white balls are drawn, without replacement. In the past, the most frequently occurring white ball has been 23.

   (a) In the next lottery, will the probability of drawing the number 23 be greater than 5/59, less than 5/59, or equal to 5/59?

   Before one season, the Oakland A's were considered to be an average major league baseball team, predicted to win half (81) of their 162 games. They win the first six games of the season.

   (b) True or false: After the first six games, the best prediction of the total number of games the Oakland A's win that season is 162 out of 162.

   (c) True or false: After the first six games, the best prediction of the total number of games the Oakland A's win that season is still 81 out of 162.

   I survey a simple random sample of 1000 U.S. households and find out their income.

   (d) True or false: By the Central Limit Theorem, the incomes in the population will have an approximately normal distribution.

   (e) True or false: By the Central Limit Theorem, the incomes in the sample will have an approximately normal distribution.

5. Let X be a discrete random variable with probability mass function

$$P(X = x) = \begin{cases} 0.3 & x = -2 \\ 0.6 & x = -1 \\ 0.1 & x = 12 \\ 0 & \text{otherwise.} \end{cases}$$

Let $X_1, \ldots, X_n$ be an iid sequence of random variables with the same distribution as X. Let $\bar{X}$ be the sample mean (of $X_1, \ldots, X_n$.)

(a) Find EX.

(b) Find $\text{Var}(X)$.

(c) What is the expected value of $\bar{X}$?

(d) What is the variance of $\bar{X}$? (Note: This will depend on n.)

(e) Suppose $n = 100$. Use the R function `pnorm()` to find the approximate probability that $\bar{X}$ is greater than 0.5.

6. (Extra credit for everyone.) I want to find out the average number of people per household in the U.S. I survey a simple random sample of U.S. households and obtain the results displayed in the following table.

Household size	Number of households
1	27
2	34
3	16
4	13
5	6
6	3
7	1

(a) Lacking any other information, our best estimate for the population mean household size is the sample mean. What is the sample mean of our data?

(b) What is our estimate for the standard deviation of household sizes?

(c) What is the estimated standard error of the sample mean? (That is, based on our answer to (b), what is our estimate for the standard deviation of the distribution of the sample mean?)

(d) Our error is the difference between the sample mean and the population mean. Using the normal distribution, find the approximate probability that the absolute value of the error in a survey of this form and size is less than 0.5.

(e) Can we be reasonably sure that the average household size for all U.S. households is between 2 and 3?