# Problem set 7

Online S520

## Upload your answers through the Assignments tab on Canvas

1. (10 points) Trosset exercise 8.4.6.

   Note:

   - "Make a profit" implies the change in your investment is greater than \$0.
   - Recall that for a continuous probability distribution, the probability of assigning one single value is 0.
   - Review Theorem 5.2 (sum of independent normal RVs) and the linear transformation of normal RVs in the supplemental materials in Module 6

2. (10 points) 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.

3. (10 points) 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 table below.

| 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?

Note:

- You can either enter data in R and compute the sample mean, or view the results in the table as a discrete probability distribution.
- If you would like to enter the data in R, the R function rep() might be useful.
- Some examples:

```
> rep(c(1,2), each = 5)
[1] 1 1 1 1 1 2 2 2 2 2

> rep(c(1,2), times = 5)
[1] 1 2 1 2 1 2 1 2 1 2

> rep(c(1,2), times = c(2,5))
[1] 1 1 2 2 2 2 2
```

(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?)

Note:

- Refer to Chapter 8 lecture notes page 8 for the definition.

(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.

Note:

- Read lecture notes page 8 for some definitions and results
- What is the approximate distribution of the difference between the sample mean and the population mean?
- Review the concept of absolute value: If $|X| < 3$, then $-3 < X < 3$.

(e) Can we be reasonably sure that the average household size for all U.S. households is between 2 and 3? (Hint: Use results from parts (a) and (d).)