Chapters covered: Chapter 6

Show your work to receive full credit.

- **Problem 1** The following statements contain errors. Identify all of the errors in each statement and explain why they are errors.

  a. The probabilities that a real estate agent will sell 0, 1, 2, 3, or 4 homes during a given month are 0.85, 0.15, 0.09, 0.03, and -0.01.

  b. Suppose you draw 3 cards *without replacement* from a standard deck of playing cards. Let A be the event that the first selected card is a heart. Let B be the event that the first selected card is red. Then $P(A \cap B) = P(A)P(B) = (1/4)(1/2) = 1/8$. Let X be the total number of times that the selected card is a red heart, then $X \sim Bin(3, \frac{1}{8})$.

- **Problem 2** The probability distribution for a random variable X is provided in the following table:

x	0	1	2	3	4
$P(X = x)$	0.1	0.2	0.3	0.1	0.3

  a. Verify that this is a valid probability distribution

  b. Compute the mean of X.

  c. Compute the standard deviation of X.

- **Problem 3** On a school's basketball team, the top free throw shooters usually have probability of about 0.90 of making any given free throw.

  a. During a game, one such player shoots 10 free throws. Let X = number of free throws they make. What must you assume for X to have a binomial distribution?

  b. Specify the values of n and p for the binomial distribution of X in part a.

  c. Find the probability that the player makes (i) exactly 9 free throws, (ii) more than 7 free throws, and (iii) fewer than or equal to 6 free throws. Show your work including formulas with numbers plugged in. You may double check your answers using `dbinom()` or `pbinom()`.

  d. Suppose this player shoots 400 free throws over the course of a season. Find the mean and standard deviation of the probability distribution of the number of free throws he makes.

- **Problem 4** Let Z have a standard normal distribution and let X have a normal distribution with mean 1/2 and standard deviation 2.

  a. Does Z or X exhibit more spread in the values it takes. Explain.

  b. Without using `R`, compute $P(-3.42 \leq X \leq 4.42)$. (Hint: Use that fact that $P(Z \leq 1.96) = 0.975$ and $P(Z \leq -1.96) = 0.025$).

  c. Without using `R`, compute $P(|Z| > 1.96)$ (Hint: Use that fact that $P(Z \leq 1.96) = 0.975$ and $P(Z \leq -1.96) = 0.025$).

  d. You observe $x = 3/2$. Compute the z-score. Is this an unusual observation? Explain why or why not.

- **Problem 5** A college senior who took the Graduate Record Examination scored 620 on the Verbal Reasoning section and 670 on the Quantitative Reasoning section. The mean score for Verbal Reasoning section was 462 with a standard deviation of 119 and the mean score for the Quantitative Reasoning was 584 with a standard deviation of 151. Suppose that both distributions are normal.

  a. Let V = a GRE taker's score on the Verbal Reasoning section and Q = a GRE taker's score on the Quantitative Reasoning section. Write down the statistical notations for the two normal distributions.

  b. What is the student's z-score on the Verbal Reasoning section? On the Quantitative Reasoning section?

  c. What do these z-scores tell you? Relative to others, which section did she do better on?

  d. Find her percentile scores for the two exams. (Hint: use `pnorm()` to help with your calculation.)

  e. What percent of test takers did better than her on the Verbal Reasoning section? On the Quantitative Reasoning section? (Hint: use the result from part d)

  f. In order to be the top 10% on the Quantitative Reasoning section, what is the minimum score she should have had? (Hint: use `qnorm()` to help with your calculation.)

  g. Explain why simply comparing her raw scores from the two sections would lead to the incorrect conclusion that she did better on the Quantitative Reasoning section.

R Problem

In this problem, we will use the SOCR Data Dinov 020108 HeightsWeights data set, which contains 25,000 synthetic records of human heights and weights of 18 years old children.

Import the data using the following command.

```
dat <- read.csv("https://yuyangyy.com/stat3011/data/SOCR_HeightWeight.csv")
```

(a) Construct a histogram and a Q-Q plot of the heights. Include your graphics in your answer. Describe the overall shape of the distribution. Can we safely assume that the height variable follows a normal distribution?

(b) Find the mean and standard deviation of the height variable. Round your answers to three decimal places.

(c) Regardless of your answer in (a) (whether normal or not), let's assume that the height variable is continuous and normally distributed. Based on the mean and standard deviation from (b), (i) find the 90th percentile; (ii) if a male teenager is 70 inches tall, then what percentage of teenagers are taller than him? Show your work.