STAT 270 Assignment 2

Upload to Crowdmark by 5pm on Wednesday, June 10, 2020

Instructions: Be sure to label the axes appropriately on all graphs. Also, you must provide any code you use (final computed values and plots are not sufficient). **I strongly recommend that you create a Jupyter Notebook for Question 1 of this assignment**, embedding your written comments in Markdown cells. For every question, you should create a separate file (in .jpg, .pdf, or .png format) for uploading to Crowdmark.

1. (8 marks) The data in the file `fishdata.txt` consist of the weights (in grams) and diagonal lengths (in cm) of 159 fish observed at local fish markets.

   a. (3 marks) Create a plot that shows the relationship between weight and diagonal length.
   b. (1 mark) Compute the sample correlation of weight and diagonal length.
   c. (2 marks) Provide an interpretation of the sample correlation you computed in b).
   d. (2 marks) Is the sample correlation a good representation of the relationship between weight and diagonal length suggested by your plot in a)? Explain.

2. (4 marks) A researcher conducts an experiment using a closed tank of gas. She considers n different temperatures of the gas (recording them as $x_1, x_2, \ldots, x_n$ in ºC) and, at each one, measures the pressure inside the tank (recording the pressures as $y_1, \ldots, y_n$, in pascals). She then converts the temperatures to ºF, recording the new values as $x_1^*, x_2^*, \ldots, x_n^*$. Show that the sample correlation between temperature and pressure of the gas is the same regardless of whether the temperature is specified in ºC or ºF. Recall: Temperature can be converted from ºC to ºF via the following formula:

$$x_i^* = \frac{9}{5}x_i + 32.$$

3. (8 marks) One test for COVID-19 has a detection rate of approximately 80% and a false positive rate of approximately 0.2%.

   a. (3 marks) Explain what *detection rate* and *false positive rate* mean in this context.
   b. (3 marks) Assume that the prevalence of COVID-19 in a particular region is 1%. Compute the *negative predictive value* (the probability that an individual who receives a negative test result does not have COVID-19) of the test.
   c. (2 marks) Is the probability that two individuals who do **not** have COVID-19 test positive equal to $2 \times 0.2\% = 0.4\%$? Explain.

4. (6 marks) Assume that the probability of having a daughter is the same as the probability of having a son (each 50%). Sam has two children.
   a. (3 marks) Given that Sam's first child was a boy, what is the probability that both of his children are boys?
   b. (3 marks) Given that at least one of Sam's children is a boy, what is the probability that both of his children are boys?

5. (10 marks) Erica has 20 outfits (including one purple pantsuit and two other purple outfits) and 10 pairs of shoes (including one pair of silver flip-flops and three pairs of purple shoes). She decides to pick one outfit and one pair of shoes at random to wear to a party.

   a. (2 marks) What is the probability that she chooses her purple pantsuit and flip-flops?
   b. (2 marks) What is the probability she chooses her purple pantsuit or her flip-flops?
   c. (3 marks) What is the probability that she wears neither a purple outfit nor purple shoes?
   d. (3 marks) What is the probability that she chooses exactly one purple item (either an outfit or shoes)?

6. (4 marks) Today is Sunday. Based on the weather forecast, the probability of rain on Monday, Tuesday, and Wednesday is 20%, 10%, and 60%, respectively. The probability of rain on both Monday and Tuesday is 5%. If it rains on Tuesday, the probability of rain on Wednesday is 75%.

   a. (2 marks) Are the events "rain on Monday" and "rain on Tuesday" independent? Explain.
   b. (2 marks) Are the events "rain on Tuesday" and "rain on Wednesday" independent? Explain.

7. (8 marks) Lei has 12 candles (6 blue, 4 white, and 2 yellow), identical except for colour, in a box.

   a. (2 marks) In how many unique ways can Lei arrange all 12 candles in a single row in her window?
   b. (3 marks) If Lei randomly draws 5 candles from the box (without replacement), what is the probability that she will draw 2 blue, 2 white, and 1 yellow candle?
   c. (3 marks) Say Lei has chosen 2 blue, 2 white, and 1 yellow candle. She randomly arranges them in a line in her window. What is the probability that the 2 blue candles are side by side?

8. (4 marks) I roll a fair, 6-sided die two times. I define a random variable, X, as the number of prime numbers (defined as 2, 3, or 5) that I roll. What is the probability mass function of X?

9. (4 marks) I have a loaded (unfair) 6-sided die. Using this die, the probability of rolling a 1 is 0.25, while the probabilities of the other outcomes are equal. I roll the die once. I define a random variable, Y, as the outcome of my roll.

    a. (3 marks) What is the probability mass function of Y?
    b. (1 mark) What is the probability of rolling a 1 or a 3?