Question 1

1.

Question: Suppose Chester performs two calculations on the same random variable X: (a) he finds the variance of a random variable; (b) he finds the variance of a (random variable times some quantity). In your own words, explain why these two values will not be the same.

Answer

- **Definition:** The variance of a random variable is a measure of how much the values of the random variable differ from the mean. it's the average of the squared differences from the mean. Represented as $\sigma^2 = E[(X - \mu)^2]$

- **In Context:** In a discrete probability distribution, the variance of a random variable is the **Sum** of the squared differences from the mean. Since, it is a sum we need to take each value by themselves.

- **Equation**

$$\sigma^2 = \sum_{i=1}^{n}(x_i - \mu)^2 \cdot P(x_i)$$

Question 2

Mary investigates two random variables, X and Y , where $Y = X^4$. She finds that $E(XY) = 0$, $E(X) = 0$, and $E(Y) = 4$.

(a)

Question: Using this information, find $COV(X, Y)$

Answer

- **Definition:** The covariance of two random variables is a measure of how much the two random variables vary together.

- **Equation**

$$COV(X, Y) = E(XY) - E(X)E(Y)$$
$$\text{or}$$
$$COV(X, Y) = E(XY) - \mu_X \mu_Y$$
$$\text{where } \mu_X = E(X) \text{ and } \mu_Y = E(Y)$$

- **Given:** $E(XY) = 0$, $E(X) = 0$, and $E(Y) = 4$

$$COV(X, Y) = E(XY) - E(X)E(Y)$$
$$= 0 - 0 \cdot 4$$
$$= 0$$

Conclusion: The covariance of X and Y is 0.

(b)

Question: In your onw words, describe the association between X and Y.

Answer: The association between X and Y is 0. This means that the two random variables are not associated with each other.

(c)

Question: Are X and Y independent? Explain your answer using your own words. Include a statement about your covariance calculation in your answer. (You do NOT need to complete any calculations to answer this question.)

Answer: Yes, X and Y are independent. This is because the covariance of X and Y is 0. This means that the two random variables are not associated with each other.

Question 3

Answer the following questions about joint random variables.

(a)

Question: The correlation coefficient can tell us a lot about how two variables are related. What type of relationship does the correlation coefficient measure?

> **Definition:** Correlation coefficient is a measure of the strength and direction of the linear relationship between two variables.

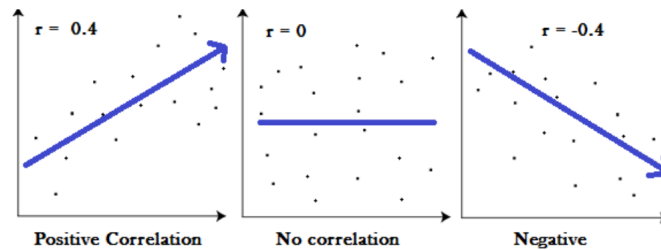> **Answer:** The correlation coefficient measures the strength and direction of the linear relationship between two variables.

Figure 1: Correlation Coefficient [1]

(b)

Question: George starts with the joint distribution (X, Y). He finds the marginal distribution for Y . George checks his work, and finds that the sum of the marginal distribution probabilities is equal to 1.25. Is this correct or was an error made (is there enough information to tell?)? If an error was made, what should be the sum of the marginal distribution probabilities?

> **Answer:** Yes, he made an error because the sum of the marginal distribution should be equal to 1.

> **Explanation:** The reason for this is because if you're looking for the probability of something happening if you had a 100% chance of something happening, that would equal 1.

> **Equation:**

$$\sum_{i=1}^{n} P(Y_i) = 1$$

Question 4

Question: Random Variables:

(a)

Question: What does it mean to "define a random variable"?

> **Answer:** To define a random variable is to assign a numerical value to each outcome of a random experiment.

> **Example:** For example, if you were to roll a die, you could define a random variable X as the number that comes up when you roll the die. and you can define X as the set of all possible outcomes of the die roll. Such that $X = \{1, 2, 3, 4, 5, 6\}$. You could also have another die lets call that Y and $Y = \{1, 2, 3, 4, 5, 6\}$. You can also take those two random variables and add them together to get a new random variable $Z = X + Y$. These are all examples of discrete random variables. So, if you were to roll the die and get a 1, then $X = 1$ and $Y = 1$ and $Z = 2$. The probability of getting $Z = 2$ can be expressed as $P(Z = 2)$.

$$P(Z = 2) = P(X = 1, Y = 1)$$
$$P(Z = 2) = P(X = 1) \cdot P(Y = 1)$$
$$P(Z = 2) = \frac{1}{6} \cdot \frac{1}{6}$$
$$P(Z = 2) = \frac{1}{36}$$

> **Conclusion:** Given our random variables X and Y, I can define a new random variable $Z = X + Y$.

(b)

Question: Explain, in your own words, why it is important that we define a random variable before we do any calculations.

> **Answer:** It is important to define a random variable before we do any calculations because it allows us to keep track of our work and simplifies expressions to make them easier to work with.

(c)

Question: When defining a random variable, we have to state the parameters. Define "parameter" using your own words.

> **Answer:** A parameter is the logical operator that is used to define the what the random variable is including. I always think of a number line and if it is less then, equal to greater than etc.

> **Example:** Lets reuse the example with the dice, again my set is defined by $X = \{1, 2, 3, 4, 5, 6\}$ and $Y = \{1, 2, 3, 4, 5, 6\}$. If I wanted to define a new random variable $Z = X + Y$ However, I wanted to only include the numbers that are less then or equal to 4 for my set of X I would have to rewrite my equation for Z as $Z = X + Y$ such that $X \leq 4$.

$$P(Z = 2) = P(X \leq 4, Y = 1)$$
$$P(Z = 2) = P(X \leq 4) \cdot P(Y = 1)$$
$$P(Z = 2) = \frac{2}{3} \cdot \frac{1}{6}$$
$$P(Z = 2) = \frac{1}{9}$$

> **Conclusion:** Parameters shape the equation and the set of possible outcomes for the random variable.

Question 5

Binomial Distribution). Suppose X Bin(n, p). Marcy defines a new random variable, X/n, where n is the same constant integer as in the distribution for X.

(a)

Question: Use properties of expected value to help Marcy find E(X/n). Fully simplify your answer (should not have an expected value in it).

Answer:

$$E(X/n) = \frac{1}{n} E(X)$$
$$= \frac{1}{n} \cdot n \cdot p$$
$$= p$$

Conclusion: The expected value of X/n is p.

(b)

Question: Use properties of variance to help Marcy find V(X/n). Fully simplify your answer (should not have a variance in it).

Answer:

$$V(X/n) = \frac{1}{n^2} V(X)$$
$$= \frac{1}{n^2} \cdot n \cdot p \cdot (1-p)$$
$$= \frac{p \cdot (1-p)}{n}$$

Conclusion: The variance of X/n is $\frac{p \cdot (1-p)}{n}$.
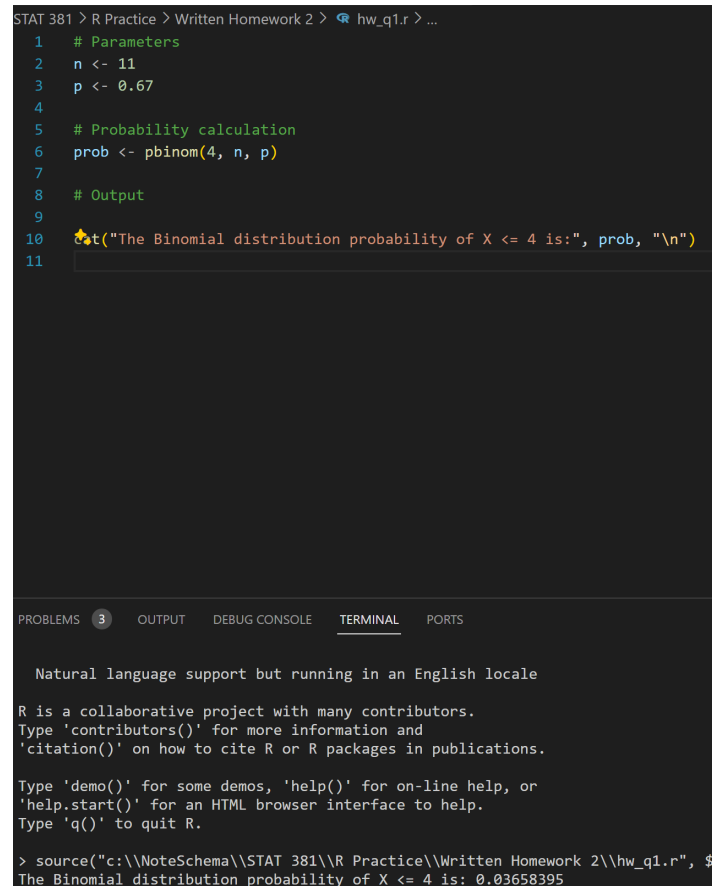
Question 6

For the following problems, write down how you would code the probabilities requested into R. You should use the specialized distribution functions mentioned in the lecture notes from Sections 5.2–5.5. There is one problem for each distribution (Binomial, Poisson, Geometric, and Hypergeometric). You do not need to provide the final answer. It is fine to hand-write your code.

(a)

Question: Suppose X ～ Bin(11, 0.3). Find $P(X < 5)$

    Answer:


```
STAT 381 > R Practice > Written Homework 2 > R hw_q1.r > ...
1    # Parameters
2    n <- 11
3    p <- 0.67
4
5    # Probability calculation
6    prob <- pbinom(4, n, p)
7
8    # Output
9
10   print("The Binomial distribution probability of X <= 4 is:", prob, "\n")
11
```
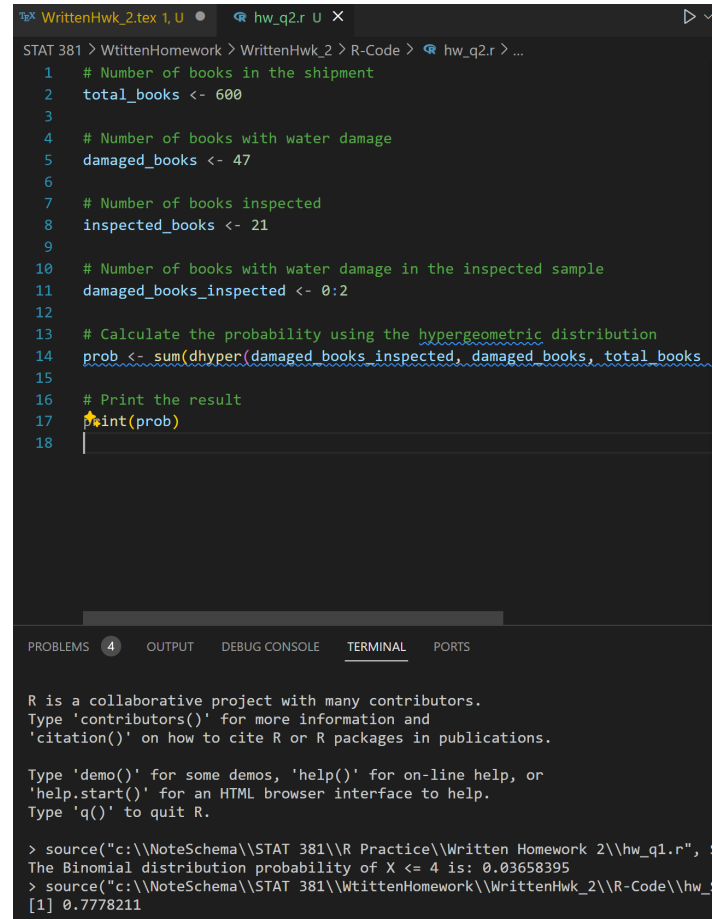
```
PROBLEMS  3    OUTPUT    DEBUG CONSOLE    TERMINAL    PORTS

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> source("c:\\NoteSchema\\STAT 381\\R Practice\\Written Homework 2\\hw_q1.r", $
The Binomial distribution probability of X <= 4 is: 0.03658395
```

Figure 2: R Code for $P(X < 5)$

(b)

Question: Suppose a truck of books shipped to the Chicago Public Library contains 600 books. For some reason, there are 47 books that have water damage. The librarian randomly chooses 21 books for inspection. What is the probability that 2 or less of the inspected books have water damage?

    Answer:

Figure 3: R Code for $P(X \leq 2)$

(c)

Question: Suppose X ~ $Pois(\lambda t = 11)$. Find $P(X \geq 5)$ incorporating the lower tail in your calculation.

Answer:

```
STAT 381 > WtittenHomework > WrittenHwk_2 > R-Code > R hw_q3.r > ...
1    library(stats)
2    ✦✦
3    # Set the parameters
4    lambda <- 11
5    x <- 15
6
7    # Calculate the probability using the lower tail
8    p_lower <- ppois(x - 1, lambda, lower.tail = TRUE)
9
10
11   # Calculate the probability of X ≥ 15
12   p_x_geq_15 <- 1 - p_lower
13
14   # Print the result
15   print(p_x_geq_15)
16
```

```
PROBLEMS  3    OUTPUT    DEBUG CONSOLE    TERMINAL    PORTS


  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> source("c:\\NoteSchema\\STAT 381\\WtittenHomework\\WrittenHwk_2\\R-Code\\hw_$
[1] 0.145956
```

Figure 4: R Code for $P(X \geq 15)$

(d)

Question: Suppose X $\quad Geom(p = 0.11)$. Define X to be like we did in section 5.4 class notes on page 1. We want to fin $P(X = 3)$.

    Answer:

Figure 5: R Code for $P(X = 3)$