Weekly Assignment 5

Caroline Liu
STU3009 - Statistical Methods for Computer Science
Trinity College Dublin

February 16, 2020

Declaration concerning plagiarism

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at `https://www.tcd.ie/calendar/`

I have completed the Online Tutorial in avoiding plagiarism 'Ready, Steady, Write', located at `http://tcd-ie.libguides.com/plagiarism/ready-steady-write`

Student Name - Caroline Liu
Student Number - 16327123

Question 1

A box contains 5 red and 5 blue marbles. Two marbles are withdrawn randomly. If they are the same color, then you win \$1.10; if they are different colors, then you lose \$1.00. Calculate:

Question 1(a)

The expected value of the amount you win

Solution -

$$-0.0667$$

Justification -

- Let X denote the amount of winnings.
- The probability of winning when drawing 2 balls of the same colour is - $\frac{\binom{2}{1}\cdot\binom{5}{2}}{\binom{10}{2}} = \frac{4}{9}$
- Winning: $1.10 \times \frac{4}{9}$, Losing: $-1 \times (1 - \frac{4}{9})$
- $\mathbb{E}(X) = 1.10 \times \frac{4}{9} - 1 \times (1 - \frac{4}{9}) = -\frac{1}{15} \approx -0.0667$

Question 1(b)

The variance of the amount you win

Solution -

$$1.0889$$

Justification -

- Using the formula: $Var(X) = \mathbb{E}[X^2]\mathbb{E}[X]^2$
- $\mathbb{E}(X^2) = 1.10^2 \times \frac{4}{9} + (-1)^2 \times (1 - \frac{4}{9}) = \frac{82}{75}$.
- $Var(X) = \frac{82}{75} - (-\frac{1}{15})^2 = \frac{82}{75} - \frac{1}{225} = \frac{49}{45} \approx 1.0889$

Question 2

Suppose you carry out a poll following an election. You do this by selecting n people uniformly at random and asking whether they voted or not, letting $X_i = 1$ if person i voted and $X_i = 0$ otherwise. Suppose the probability that a person voted is 0.6. Let $Y = \sum_{i=1}^{n} X_i$.
Hints: use linearity of the expectation and the fact that people are sampled independently.

Question 2(a)

Calculate $\mathbb{E}[X_i]$ and $Var(X_i)$.

Solution -

$$\mathbb{E}[X_i] = 0.6$$

$$Var(X_i) = 0.24$$

Justification -

- According to the question, $X_i = 1$ if a person voted, and $X_i = 0$ otherwise. Hence $\mathbb{E}[X_i]$ would only take a value within 1 and 0.
- According to the question, the probability that a person votes is 0.6, hence $\mathbb{E}[X_i] = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = P(X_i = 1) = 0.6$.
- Therefore $Var(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = 1^2 \times P(X_i = 1) + 0^2 \times P(X_i = 0) = 0.6 - (0.6)^2 = 0.6 - 0.36 = 0.24$

Question 2(b)

*Somehow there is no part (b)

Question 2(c)

What is $\mathbb{E}[Y]$? Is it the same as $\mathbb{E}[X]$ or different, and why ?

Solution - Assuming that X is the sum of all X_i's from $i = 1$ to $i = n$, $i.e. \mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$.

$$\mathbb{E}[Y] = 0.6n = \mathbb{E}[X]$$

Justification -

- For the below solution, it is assumed that X is the sum of all X_i's from $i = 1$ to $i = n$, $i.e. \mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$as it is not stated in the question.
- $\mathbb{E}[Y] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} 0.6 = 0.6n$ using the linearity of the expectation (so $\mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i]$)

- also considering the fact that these students are sampled independently, $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} 0.6 = 0.6n$
- Therefore, $\mathbb{E}[Y]$ is the same as $\mathbb{E}[X]$

Question 2(d)

What $\mathbb{E}[\frac{1}{n}Y]$

Solution -

$$0.6$$

Justification -

- According to part (b), $\mathbb{E}[Y] = 0.6n$, hence $\mathbb{E}[\frac{1}{n}Y] = \frac{1}{n}\mathbb{E}[Y] = 0.6$.

Question 2(e)

What is the variance of $\frac{1}{n}Y$ (express in terms of $Var(X)$) ?

Solution -

$$Var(\frac{1}{n}Y) = \frac{1}{n^2}Var(X)$$

Justification -

- $Var(\frac{1}{n}Y) = \frac{1}{n^2}Var(Y) = \frac{1}{n^2}Var(\sum_{i=1}^{n} X_i)$.
- Since the students are sampled independently, $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = nVar(X)$.
- Hence, $Var(\frac{1}{n}Y) = \frac{1}{n^2}Var(X)$

Question 3

Suppose that 2 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let X_i equal 1 if the $i'th$ ball selected is white, and let it equal 0 otherwise.

Question 3(a)

Give the joint probability mass function of X_1 and X_2.

Solution -

- Note that $P(X_1, X_2) = P(X_1 \cap X_2)$ in the below solution

- For the first ball chosen to be white $P(X_1 = 1) = \frac{5}{13}$, therefore for the second ball chosen to be white $P(X_2 = 1) = \frac{4}{12}$, hence

$$P(X_1 = 1, X_2 = 1) = \frac{5}{13} \cdot \frac{4}{12} = \frac{5}{39}$$

- For the first ball chosen to be white $P(X_1 = 1) = \frac{5}{13}$, therefore for the second ball chosen to be red $P(X_2 = 0) = \frac{8}{12}$, hence

$$P(X_1 = 1, X_2 = 0) = \frac{5}{13} \cdot \frac{8}{12} = \frac{10}{39}$$

- For the first ball chosen to be red $P(X_1 = 0) = \frac{8}{13}$, therefore for the second ball chosen to be white $P(X_2 = 1) = \frac{5}{12}$, hence

$$P(X_1 = 0, X_2 = 1) = \frac{8}{13} \cdot \frac{5}{12} = \frac{10}{39}$$

- For the first ball chosen to be $P(X_1 = 0) = \frac{8}{13}$, therefore for the second ball chosen to be red $P(X_2 = 0) = \frac{7}{12}$, hence

$$P(X_1 = 0, X_2 = 0) = \frac{8}{13} \cdot \frac{7}{12} = \frac{14}{39}$$

Question 3(b)

Are X_1 and X_2 independent ?(Use the formal definition of independence to determine this)

Solution -

X_1 and X_2 are not independent

Justification -

- According to the definition, for two random variables A and B, they would be independent if $P(A = a \cap B = b) = P(A = a) \cdot P(B = b)$ for all a, b.
- Intuitively, two random variables A and B are independent if knowing the value of one of them does not change the probabilities for the other, $i.e. P(B = b \mid A = a) = P(B = b)$ for all a, b.
- Events X_1 and X_2 in this case are not independent as $P(X_2)$ changes depending on the result of X_1 whether a white ball has be taken or not during the first draw.

$$P(X_1 = 1) = \frac{5}{39} + \frac{10}{39} = \frac{15}{39}$$

$$P(X_2 = 1) = \frac{5}{39} + \frac{10}{39} = \frac{15}{39}$$

$$P(X_1 = 1)P(X_2 = 1) = \frac{15}{39} \cdot \frac{15}{39} = \frac{25}{169}$$

According to part (a),

$$P(X_1 = 1, X_2 = 1) = P(X_1 = 1 \cap X_2 = 1) = \frac{5}{39} \neq \frac{25}{169}$$

- Hence $P(X_1 = 1)P(X_2 = 1) \neq P(X_1 = 1 \cap X_2 = 1)$ therefore variables X_1 and X_2 are not independent.

Question 3(c)

Calculate $\mathbb{E}[X_2]$

Solution -

$$0.3846$$

Justification -

$$\mathbb{E}[X_2] = 1 \times \frac{5}{39} + 0 \times \frac{10}{39} + 1 \times \frac{10}{39} + 0 \times \frac{14}{39} = \frac{5}{39} + \frac{10}{39} = \frac{15}{39} = \frac{5}{13} \approx 0.3846$$

Only considering cases where $X_2 = 1$

Question 3(d)

Calculate $\mathbb{E}[X_2 \mid X_1 = 1]$

Solution -

$$0.333$$

Justification -

- Since according to part (b), X_1 and X_2 are dependent, we obtain that

$$P(X_2 = 1 \mid X_1 = 1) = \frac{4}{12} = \frac{1}{3}$$

$$P(X_2 = 0 \mid X_1 = 1) = \frac{8}{12} = \frac{2}{3}$$

- hence, to obtain the conditional expectation in the question,

$$\mathbb{E}[X_2 \mid X_1 = 1] = (1 \times P(X_1 = 1 \mid X_2 = 1)) + (0 \times P(X_1 = 1 \mid X_2 = 0))$$

$$\mathbb{E}[X_2 \mid X_1 = 1] = 1 \times \frac{1}{3} + 0 \times \frac{2}{3} = \frac{1}{3} \approx 0.333$$