Statistics - Week 5 Questions

Owen Burke, 15316452

March 6, 2019

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 Two marbles are withdrawn randomly.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:

a)The expected value of the amount you win

Firstly, let's define the expected value.
This is given as :

$$\text{E[X]} = \sum_{i=1}^{n}(X_\text{i})(P(X = x_\text{i}))$$

Let's say the possible values of X are 1.1 and -1 (from the question - losing \$1.00 is the same as saying X = -1).
So, from the formula, for each X, we must find the probability of X = P(X).
This is also P(winning) and P(losing). The probability of winning is the P(two reds) + P(two blues).
Let's start with P(two reds).
This is equal to $\text{P}(\text{R}_1 \cap \text{R}_2)$, where R_1 is the event of the first ball being red, and so on. This is given as :

$$\text{P}(\text{R}_1 \cap \text{R}_2) = \text{P}(\text{R}_2 \mid \text{R}_1)\text{P}(\text{R}_1)$$

We know $\text{P}(\text{R}_1) = \frac{5}{10}$ as there are 5 red balls out of 10 balls.
We also now $\text{P}(\text{R}_2 \mid \text{R}_1) = \frac{4}{9}$ as after the first red ball, there are 4 red balls left out of 9 balls. This gives :

$$\text{P}(\text{R}_1 \cap \text{R}_2) = \frac{5}{10} * \frac{4}{9} = \frac{2}{9}$$

Given that there are an equal number of blues and reds, P(two reds) = P(two blues).
By this P(winning) = $\frac{2}{9} + \frac{2}{9} = \frac{4}{9}$ and P(losing) = 1 - $\frac{4}{9} = \frac{5}{9}$.
We can now find E[X] as :

$$\text{E[X]} = (1.1)(\tfrac{4}{9}) + (\text{-}1)(\tfrac{5}{9}) = \mathbf{-0.06666667}.$$

b)The variance of the amount you win.

Given that Var(X) = $\text{E}[\text{X}^2]$ - $(\text{E[X]})^2$, and we can find $\text{E}[\text{X}^2]$ as $((1.1^2)(\frac{4}{9}) + ((\text{-}1)^2)(\frac{5}{9})) = 1.0933333333$.
We also know E[X] = -0.0666667, so Var(X) is given as :

$$\text{Var(X)} = 1.093333333 - ((-0.06666667)^2) = \mathbf{1.088888}$$

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.

a)Calculate $E[X_i]$ and $Var(X_i)$.

Given that X_i can only be either 0 or 1, and we know the formula for the expected value is

$$E[X] = \sum_{i=1}^{n} (X_i)(P(X = x_i))$$

and we also know the probabilities of voting and not voting, we can find $E[X_i]$ by

$$E[X_i] = (1)(0.6) + (0)(1\text{-}0.6) = \mathbf{0.6}$$

We also know the formula for variance is

$$Var(X_i) = E[{X_i}^2] - (E[X_i])^2$$

So $Var(X_i)$ is given as

$$Var(X_i) = ((1^2)(0.6) + (0^2)(0.4)) - (0.6)^2 = \mathbf{0.24}$$

Let $Y = \sum_{i=1}^{n} X_i$

c)What is E[Y]? Is it the same as E[X] or different, and why?

Given that $Y = \sum_{i=1}^{n} X_i$, then we can say $E[Y] = E[\sum_{i=1}^{n} X_i]$
Given that we know $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$, and that $E[X_i] = 0.6$ we can conclude

$$E[Y] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} 0.6 = (0.6)n$$

From this we conclude that E[Y] is the expected number of people who voted. This is **not** the same as E[X] (which is the expected value of whether an individual voted or not). We can also see this from the formula, as $0.6 \neq (0.6)n$, except where $n = 1$.

d)What is $E[\frac{1}{n}Y]$?

From $E[Y] = E[\sum_{i=1}^{n} X_i]$ and part c), we know $E[\sum_{i=1}^{n} X_i] = E[n(X_i)]$.
Therefore, $E[\frac{1}{n}Y] = E[(\frac{1}{n})(n(X_i))]$. In this case, the $\frac{1}{n}$ and multiplication by n cancel out and we get

$$E[\frac{1}{n}Y] = E[(\frac{1}{n})(n(X_i))] = E[X_i] = 0.6$$

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

Firstly, let's define Var($\frac{1}{n}$Y).

$$\mathrm{Var}(\tfrac{1}{n}\mathrm{Y}) = \mathrm{E}[(\tfrac{1}{n}\mathrm{Y})^2] - (\mathrm{E}[\tfrac{1}{n}\mathrm{Y}])^2$$

Given from d), that E[$\frac{1}{n}$Y] = E[X_i], then

$$\mathrm{Var}(\tfrac{1}{n}\mathrm{Y}) = \mathrm{E}[{X_i}^2] - (\mathrm{E}[X_i])^2$$

From a) we know that E[${X_i}^2$] - (E[X_i])^2 = Var(X_i). Therefore Var($\frac{1}{n}$Y) in terms of Var(X_i) is

$$\mathrm{Var}(\tfrac{1}{n}\mathrm{Y}) = \mathrm{Var}(X_i)$$

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 ith ball selected is white, and let it equal 0 otherwise.

a) Give the joint probability mass function of X_1 and X_2

X_i = 1, if the ball is white.
X_i = 0, if the ball is not white (red).
To find the joint probability mass function, we must find the intersections for all different values of X_1 and X_2 (ie when $X_1 = 0$ and $X_2 = 0$, when $X_1 = 0$ and $X_2 = 1$, and so on) The intersection is given by the formula :

$$P(X_1 \cap X_2) = P(X_2 \mid X_1)P(X_1)$$

So we must now find the values that we will need to get the intersections, given by :
$P(X_1 = 0) = \frac{8}{13}$, for 8 red balls from 13.
$P(X_1 = 1) = \frac{5}{13}$, for 5 white balls from 13.
$P(X_2 = 0 \mid X_1 = 0) = \frac{7}{12}$, for 7 red balls (one having been taken already) from 12 remaining balls.
$P(X_2 = 1 \mid X_1 = 0) = \frac{5}{12}$, for 5 white balls (no white balls having been taken already) from 12 remaining balls.
$P(X_2 = 0 \mid X_1 = 1) = \frac{8}{12}$, for 8 red balls (no red balls having been taken already) from 12 remaining balls.
$P(X_2 = 1 \mid X_1 = 1) = \frac{4}{12}$, for 4 white balls (one having been taken already) from 12 remaining balls.

We can now find the intersections given by :
$P(X_1 = 0 \cap X_2 = 0) = \frac{7}{12} * \frac{8}{13} = \frac{14}{39}$
$P(X_1 = 1 \cap X_2 = 0) = \frac{8}{12} * \frac{5}{13} = \frac{10}{39}$

$\mathrm{P}(\mathrm{X}_1 = 0 \cap \mathrm{X}_2 = 1) = \frac{5}{12} * \frac{8}{13} = \frac{10}{39}$
$\mathrm{P}(\mathrm{X}_1 = 1 \cap \mathrm{X}_2 = 1) = \frac{4}{12} * \frac{5}{13} = \frac{5}{39}$

By these calculations, the joint probability mass function is given by :

	$\mathbf{X_1 = 0}$	$\mathbf{X_1 = 1}$	$\mathbf{P(X_2 = x)}$
$\mathbf{X_2 = 0}$	$\frac{14}{39}$	$\frac{10}{39}$	$\frac{8}{13}$
$\mathbf{X_2 = 1}$	$\frac{10}{39}$	$\frac{5}{39}$	$\frac{5}{13}$
$\mathbf{P(X_1 = x)}$	$\frac{8}{13}$	$\frac{5}{13}$	1

b) Are $\mathbf{X_1}$ and $\mathbf{X_2}$ independent? (Use the formal definition of independence to determine this)

The formal definition of independence states two events are independent if

$$\mathrm{P}(\mathrm{E} \cap \mathrm{F}) = \mathrm{P}(\mathrm{E})\mathrm{P}(\mathrm{F})$$

So, taking $\mathrm{X}_1 = 1$ and $\mathrm{X}_2 = 1$ (from part a):
$\mathrm{P}(\mathrm{X}_1 = 1 \cap \mathrm{X}_2 = 1) = \frac{5}{39}$
$\mathrm{P}(\mathrm{X}_1 = 1) = \frac{5}{13}$
$\mathrm{P}(\mathrm{X}_2 = 1) = \frac{5}{13}$
From the definition of independence :

$$\frac{5}{39} \neq \frac{5}{13} * \frac{5}{13}$$

Therefore X_1 and X_2 are **not** independent.

c) Calculate $\mathbf{E[X_2]}$

$\mathrm{E}[\mathrm{X}_2] = \sum_{i=1}^{n}(X_{2\mathrm{i}})(P(X = X_{2\mathrm{i}}))$
This is given as :

$$\mathrm{E}[\mathrm{X}_2] = (0)(\tfrac{8}{13}) + (1)(\tfrac{5}{13}) = \frac{5}{13}$$

(the probabilites were calculated in part a)

d) Calculate $\mathbf{E[X_2 \mid X_1 = 1]}$

$\mathrm{E}[\mathrm{X}_2 \mid \mathrm{X}_1 = 1] = \sum_{i=1}^{n}(X_{2\mathrm{i}}|X_1 = 1)(P(X_{2\mathrm{i}}|X_1 = 1))$
This is given as :

$$\mathrm{E}[\mathrm{X}_2 \mid \mathrm{X}_1 = 1] = (0)(\tfrac{8}{12}) + (1)(\tfrac{4}{12}) = \frac{4}{12}$$

;(the probabilites were calculated in part a)