Probability & Statistics: Assignment 5

Due on November 23, 2018 at 8:30am

Prof. D. Eynard

A. Romanelli / A. Vicini

Problem 1

Given $X = Y = \{0, 1, 2\}$ and the following probability distribution:

${}_{Y}{}^{X}$	0	1	2	$P(Y)$
0				
1				
2				
$P(X)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	

Knowing that:

- $P_{Y|X}(0,0) = 1$
- $P_{Y|X}(0,1) = P_{Y|X}(1,1) = \frac{1}{2}$
- $P_{Y|X}(0,2) = P_{Y|X}(1,2) = P_{Y|X}(2,2) = \frac{1}{3}$

We'll compute the following probabilities:

We can reverse the formula:

$$P_{Y|X} = \frac{P_{XY}(x,y)}{P_X(x)}$$

to obtain:

$$P_{XY} = P_X(x) \cdot P_{Y|X}(y,x)$$

and thus:

- $P_{XY}(0,0) = P_X(0) \cdot P_{Y|X}(0,0) = \frac{1}{2}$
- $P_{XY}(0,1) = P_X(1) \cdot P_{Y|X}(1,0) = 0$
- $P_{XY}(0,2) = P_X(2) \cdot P_{Y|X}(2,0) = 0$
- $P_{XY}(1,0) = P_X(1) \cdot P_{Y|X}(0,1) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$
- $P_{XY}(2,0) = P_X(2) \cdot P_{Y|X}(0,2) = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$
- $P_{XY}(1,1) = P_X(1) \cdot P_{Y|X}(1,1) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$
- $P_{XY}(1,2) = P_X(1) \cdot P_{Y|X}(2,1) = 0$
- $P_{XY}(2,1) = P_X(2) \cdot P_{Y|X}(1,2) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}$
- $P_{XY}(2,2) = P_X(2) \cdot P_{Y|X}(2,2) = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$

Which gives us back the following table:

${}_{Y}{}^{X}$	0	1	2	$P(Y)$
0	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{17}{24}$
1	0	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{5}{24}$
2	0	0	$\frac{1}{12}$	$\frac{1}{12}$
$P(X)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	

$P(Y)$ was easily findable by computing the respecting sums:
$P_Y(0) = \frac{1}{2} + \frac{1}{8} + \frac{1}{12} = \frac{17}{27}$, $P_Y(1) = \frac{1}{8} + \frac{1}{12} = \frac{5}{24}$ $P_Y(2) = \frac{1}{12}$

Now we'll have to compute $P_{X|Y}$, knowing that:

$$P_{X|Y} = \frac{P_{XY}(x,y)}{P_Y(y)}$$

Problem 1 continued on next page...

- $P_{X|Y}(0,0) = \frac{1}{2} \cdot \frac{24}{17} = \frac{12}{17}$
- $P_{X|Y}(0,1) = 0$
- $P_{X|Y}(0,2) = 0 \cdot 12 = 0$
- $P_{X|Y}(1,0) = \frac{1}{8} \cdot \frac{24}{17} = \frac{3}{17}$
- $P_{X|Y}(2,0) = \frac{1}{12} \cdot \frac{24}{17} = \frac{2}{17}$
- $P_{X|Y}(2,1) = \frac{1}{12} \cdot \frac{24}{5} = \frac{2}{5}$
- $P_{X|Y}(2,2) = \frac{1}{12} \cdot 12 = 1$
- $P_{X|Y}(1,1) = \frac{1}{8} \cdot \frac{24}{5} = \frac{3}{5}$
- $P_{X|Y}(1,2) = 0$

We shall now compute $\text{Var}(X)$ and $\text{Var}(Y)$ which we know to be given by: $\text{E}[X^2] - \text{E}[X]^2$ and $\text{E}[Y^2] - \text{E}[Y]^2$ respectively.

$$\text{E}[X] = \sum_{i=0}^{2} i \cdot P_X(i) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} = \frac{3}{4}$$

$$\text{E}[X^2] = \sum_{i=0}^{2} i^2 \cdot P_X(i) = 0 + \frac{1}{4} + 4 \cdot \frac{1}{4} = \frac{5}{4}$$

$$\text{E}[Y] = \sum_{i=0}^{2} i \cdot P_Y(i) = 0 + \frac{5}{24} + 2 \cdot \frac{1}{12} = \frac{3}{8}$$

$$\text{E}[Y^2] = \sum_{i=0}^{2} i^2 \cdot P_Y(i) = 0 + \frac{5}{24} + 4 \cdot \frac{1}{12} = \frac{13}{24}$$

If we plug in the values we get:

$$\text{Var}(X) = \frac{5}{4} - \left(\frac{3}{4}\right)^2 = \frac{11}{16}$$

$$\text{Var}(Y) = \frac{13}{24} - \left(\frac{3}{8}\right)^2 = \frac{77}{192}$$

We can compute the $\text{Cov}(X, Y)$ as:

$$\text{E}[XY] - \text{E}[X]\text{E}[Y]$$

We can obtain

$$\text{E}[XY] = \sum_{i=0}^{2} xy \cdot P_{XY}(x, y) = \frac{5}{8}$$

Substituting we get:

$$\text{Cov}(X, Y) = \text{E}[XY] - \text{E}[X]\text{E}[Y] = \frac{5}{8} - \frac{3}{4} \cdot \frac{3}{8} = \frac{11}{32}$$

Problem 2

If we let X be a random variable with a range $\mathcal{X} = \{1, 2, \dots, k\}$ and $P_X(k) = \frac{1}{k}$, then:

$$\mathrm{E}[X] = \sum_{i=1}^{k} i \cdot P_X(i)$$

Under the previous assumption of a uniform distribution, we can say that $P_X(i)$ is a constant and thus:

$$\mathrm{E}[X] = \frac{1}{k} \sum_{i=0}^{k} i = \frac{1}{k} \cdot \frac{k(k+1)}{2} = \frac{k+1}{2}$$

We'll also need $\mathrm{E}[X^2]$ for the next step:

$$\mathrm{E}[X^2] = \frac{1}{k} \sum_{i=0}^{k} i^2 = \frac{1}{6k} k(k+1)(2k+1) = \frac{1}{6}(k+1)(2k+1)$$

And now we should be finally able to obtain $\mathrm{Var}(X)$:

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \frac{1}{6}(k+1)(2k+1) - \left(\frac{k+1}{2}\right)^2 = \frac{1}{12}(k^2 - 1)$$

Problem 3

We can start by recalling what the Chebyshev inequality tells us that, given any $\epsilon > 0$ which we can associate with the probability of being farther out than ϵ from the mean bounded by the variance as:

$$P(|X - \mathrm{E}[X]| \geq \epsilon) \leq \frac{\mathrm{Var}(X)}{\epsilon^2}$$

For a symmetric distribution like a Binomial one we can say:

$$P(X - \mathrm{E}[X] \geq \epsilon) \leq \frac{1}{2}\frac{\mathrm{Var}(X)}{\epsilon^2}$$

Since $P(|X - \mathrm{E}[X]| \geq \epsilon) = 2 \cdot P(X - \mathrm{E}[X] \geq \epsilon)$

To then compute $\mathrm{Var}(X)$, we'll need the expectation values of the outcomes of the roulette for each game. Let then X_i be the random variable associated with a given game:

$$\mathrm{E}[X_i] = \sum P_{X_i}(x) \cdot x = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}, \qquad x \in \{0, 1\}$$

To compute the expectation value of X_n, that is, the mean of all the X_i:

$$X_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\mathrm{E}[X_n] = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[X_i] = \frac{1}{n} \cdot n \cdot \frac{1}{2} = \frac{1}{2}$$

We'll now compute $\mathrm{Var}(X_i)$, which we'll need to compute $\mathrm{Var}(X_n)$:

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

Since the values 0,1 do not change when squared, we'll simply have to compute

$$\mathrm{E}[X_i] - \mathrm{E}[X_i]^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

And finally to compute $\mathrm{Var}(X_n)$ we do

$$\mathrm{Var}(X_n) = \mathrm{E}[X_n^2] - \mathrm{E}[X_n]^2 = \left(\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[X_i]\right)^2 - \frac{1}{4} = \left(\frac{1}{n^2}\right) \cdot n \cdot \frac{1}{4} = \frac{1}{4n}$$

a) Out of 40 rounds, to have at least 70% black results means that: $n = 40, p = 0.7, \mathrm{Var}(X_{40}) = \frac{1}{160}$, which gives us:

$$\begin{aligned} & & P(X_{40} - \mathrm{E}[X_{40}] \geq \epsilon) &\leq \tfrac{1}{2}\tfrac{\mathrm{Var}(X_{40})}{\epsilon^2} \\ \text{If we let: } X_{40} - \mathrm{E}[X_{40}] &= \epsilon & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.2) &\leq \tfrac{1}{2}\tfrac{1}{160 \cdot (\frac{1}{5})^2} \\ \tfrac{1}{40} \cdot 40 \cdot 0.7 - \tfrac{1}{2} &= \epsilon & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.2) &\leq \tfrac{25}{320} \\ \tfrac{7}{10} - \tfrac{5}{10} &= \epsilon & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.2) &\leq 0.078125 \\ \tfrac{1}{5} &= \epsilon & & \end{aligned}$$

Problem 3 continued on next page. . .

b) Out of 40 rounds, having at least 50.1% black results would mean: $n = 40, p = 0.501, \mathrm{Var}(X_{40}) = \frac{1}{160}$

$$\begin{aligned} & P(X_{40} - \mathrm{E}[X_{40}] \geq \epsilon) \leq \tfrac{1}{2}\tfrac{\mathrm{Var}(X_{40})}{\epsilon^2} \\ \text{If we let: } X_{40} - \mathrm{E}[X_{40}] = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{1}{2}\tfrac{1}{160\cdot(0.001)^2} \\ 0.501 - 0.5 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{1000000}{320} \\ 0.001 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq 3125 \end{aligned}$$

c) Out of 1500 rounds, having at least 60% black results means: $n = 1500, p = 0.6, \mathrm{Var}(X_{1500}) = \frac{1}{6000}$

$$\begin{aligned} & P(X_{1500} - \mathrm{E}[X_{1500}] \geq \epsilon) \leq \tfrac{1}{2}\tfrac{\mathrm{Var}(X_{1500})}{\epsilon^2} \\ \text{If we let: } X_{1500} - \mathrm{E}[X_{1500}] = \epsilon \quad & P(X_{1500} - \mathrm{E}[X_{1500}] \geq 0.1) \leq \tfrac{1}{2}\tfrac{1}{6000\cdot(0.1)^2} \\ 0.6 - 0.5 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{100}{12000} \\ 0.1 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{1}{120} \approx 0.0083 \end{aligned}$$

d) Out of 1500 rounds, having at least 50.1% black results means: $n = 1500, p = 0.501, \mathrm{Var}(X_{1500}) = \frac{1}{6000}$

$$\begin{aligned} & P(X_{1500} - \mathrm{E}[X_{1500}] \geq \epsilon) \leq \tfrac{1}{2}\tfrac{\mathrm{Var}(X_{1500})}{\epsilon^2} \\ \text{If we let: } X_{1500} - \mathrm{E}[X_{1500}] = \epsilon \quad & P(X_{1500} - \mathrm{E}[X_{1500}] \geq 0.001) \leq \tfrac{1}{2}\tfrac{1}{6000\cdot(0.001)^2} \\ 0.501 - 0.5 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{1000000}{12000} \\ 0.001 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{1000}{12} \approx 83.333 \end{aligned}$$

e) Out of 10^6 rounds, having at least 70% black results would mean: $n = 10^6, p = 0.7, \mathrm{Var}(X_{10^6}) = \frac{1}{4\cdot 10^6}$

$$\begin{aligned} & P(X_{10^6} - \mathrm{E}[X_{10^6}] \geq \epsilon) \leq \tfrac{1}{2}\tfrac{\mathrm{Var}(X_{10^6})}{\epsilon^2} \\ \text{If we let: } X_{10^6} - \mathrm{E}[X_{10^6}] = \epsilon \quad & P(X_{10^6} - \mathrm{E}[X_{10^6}] \geq 0.2) \leq \tfrac{1}{2}\tfrac{1}{4\cdot 10^6\cdot(0.2)^2} \\ 0.7 - 0.5 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.2) \leq \tfrac{25}{8\cdot 10^6} \approx 3.125\cdot 10^{-6} \\ 0.2 = \epsilon \quad & \end{aligned}$$

f) Out of 10^6 rounds, having at least 50.1% black results means: $n = 10^6, p = 0.501, \mathrm{Var}(X_{10^6}) = \frac{1}{4\cdot 10^6}$

$$\begin{aligned} & P(X_{10^6} - \mathrm{E}[X_{10^6}] \geq \epsilon) \leq \tfrac{1}{2}\tfrac{\mathrm{Var}(X_{10^6})}{\epsilon^2} \\ \text{If we let: } X_{10^6} - \mathrm{E}[X_{10^6}] = \epsilon \quad & P(X_{10^6} - \mathrm{E}[X_{10^6}] \geq 0.001) \leq \tfrac{1}{2}\tfrac{1}{4\cdot 10^6\cdot(0.001)^2} \\ 0.501 - 0.5 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{10^6}{8\cdot 10^6} \\ 0.001 = \epsilon \quad & P(X_{40} - \mathrm{E}[X_{40}] \geq 0.001) \leq \tfrac{10^6}{8\cdot 10^6} = \tfrac{1}{8} = 0.125 \end{aligned}$$

Problem 4

Given the PDF of the standard normal distribution $\mathcal{N}(\mu, \sigma^2)$ we want to compute the following:

a) $P(X > 2) = 0.0228$ as it follows from the symmetry of the distribution.
$P(X \leq -2) \equiv P(X > 2)$, since $P(X = 2) = 0$

b) $P(X \geq -1)$ We can see that we can compute this as $1 - P(X \leq -1)$ and know by symmetry that $P(X \leq -1) \equiv P(X \geq 1)$, which gives us: $1 - 0.1587 = 0.8413$

c) $P(|X - \mu| < 1)$ given that $\mu = 0$ we need to compute the probability of $P(-1 < X < 1)$, which we can calculate as: $1 - P(X \leq -1) - P(X \geq 1)$, again, because of symmetry, $P(X \leq -1) \equiv P(X \geq 1)$ and thus we get: $1 - 2P(X \geq 1) = 1 - 2 \cdot 0.1587 = 0.6826$

d) $P(|X - \mu| > 2)$ which again is the same as $P(X < -2 \cup X > 2) = P(X < -2) + P(X > 2)$ thus we can easily compute, again by symmetry $P(|X - \mu| > 2) = 2 \cdot P(X > 2) = 2 \cdot 0.0228 = 0.0456$

e) $P(X \neq 1)$, knowing that $P(X = 1) = 0$, we can find its complement by subtracting it from the total probability, which returns us $P(X \neq 1) = 1 - P(X = 1) = 1 - 0 = 1$

Problem 5