Assignment 5 Solutions

April 16, 2021

Question 1

Part a)
X and Y are independent random variables if and only for all pairs of real numbers x and y the following equality if verified

$$Pr(X = x \wedge Y = y) = Pr(X = x) \cdot Pr(Y = y)$$

Therefore, to show that two random variables are not independent, we need one pair of values for x and y such that the equation above is not verified.
We first observe that the range of X is $\{-1, 0, 1\}$ and the range of Y is $\{0, 1\}$.
Now notice that $Pr(X = -1 \wedge Y = 0) = 0$: if $X = -1$, then $Y = 1$. Moreover, we know that $Pr(X = -1) \neq 0$ and $Pr(Y = 0) \neq 0$ since each value lies in the range of X and Y respectively.
Thus, $Pr(X = -1 \wedge Y = 0) = Pr(X = -1) \cdot Pr(Y = 0)$ does not hold and X and Y are not independent.

Part b)
The range of XY is $\{-1, 0, 1\}$. We also have the following

$$Pr(XY = -1) = Pr(X = -1 \wedge Y = 1) = Pr(X = -1)$$

The last equality comes from the fact that when $X = -1$ then $Y = 1$ is true. In other words, the event $Y = 1$ contains $X = -1$.
Hence, $Pr(XY = -1) = 1/3$. In a similar fashion we can find that

$$Pr(XY = 1) = Pr(X = 1 \wedge Y = 1) = Pr(X = 1) = 1/3$$

We can conclude that $Pr(XY = 0) = 1/3$ which yields

$$E(X) = E(XY) = 0 \cdot 1/3 + 1 \cdot 1/3 - 1 \cdot 1/3 = 0$$

And thus the equality follows.

Question 2

We first note that $Pr(\text{the 7 games are played}) = Pr(\text{in the first 6 games both teams win 3 games})$
.
We can observe that this probability is equal to $\binom{6}{3} \cdot (3/4)^3 \cdot (1/4)^3$ The $\binom{6}{3}$ corresponds to the number of ways to choose three out of the first six games (that are won by the Sens). $(3/4)^3$ corresponds to the probability of those three chosen games being won by the Sens (since the results of the games are mutually independent). Similarly, $(1/4)^3$ corresponds to the probability of the three other games being lost by the Sens.

Question 3

Let G be a graph with m edges.
Let $k \neq j$ for some $1 \leq k, j \leq m$. We show that X_k and X_j are independent random variables. Consult the proof of the Theorem 7.1.1 to find the definitions of those random variables. In order to show the independence of X_k and X_j we have to show the following four equalities:

- $Pr(X_k = 1 \wedge X_j = 1) = Pr(X_k = 1) \cdot Pr(X_j = 1)$
- $Pr(X_k = 0 \wedge X_j = 1) = Pr(X_k = 0) \cdot Pr(X_j = 1)$
- $Pr(X_k = 1 \wedge X_j = 0) = Pr(X_k = 1) \cdot Pr(X_j = 0)$
- $Pr(X_k = 0 \wedge X_j = 0) = Pr(X_k = 0) \cdot Pr(X_j = 0)$

Note that from the proof we already know that $Pr(X_m = 1) = 1/2$ for any m. Let e_k and e_j be the edges of the graph G corresponding to the indicator variables X_k and X_j.

Case 1 : e_k and e_j share an endpoint. So we can write $e_k = \{u, v\}$ and $e_j = \{u, w\}$.

Let (x_1, x_2, x_3) denote the flips for u , v and w respectively.

- If $X_k = 1$ and $X_j = 1$ then (x_1, x_2, x_3) is (H, T, T) or (T, H, H) out of the possible 8 outcomes. This entails that $Pr(X_k = 1 \wedge X_j = 1) = 2/8 = 1/4$ as desired.
- If $X_k = 0$ and $X_j = 1$ then $(x_1, x_2, x_3) = (H, H, T)$ or (T, T, H) and so $Pr(X_k = 0 \wedge X_j = 1) = 1/4$.
- If $X_k = 1$ and $X_j = 0$ then $(x_1, x_2, x_3) = (H, T, H)$ or (T, H, T) and so $Pr(X_k = 1 \wedge X_j = 0) = 1/4$.
- If $X_k = 0$ and $X_j = 0$ then $(x_1, x_2, x_3) = (H, H, H)$ or (T, T, T) and so $Pr(X_k = 0 \wedge X_j = 0) = 1/4$.

Therefore, all equations are verified in this case.

Case 2 : e_k and e_j do not share an endpoint. We can write $e_k = \{u, v\}$ and $e_j = \{w, x\}$.

Let (x_1, x_2, x_3, x_4) denote the flips for u , v,w and x respectively.

- If $X_k = 1$ and $X_j = 1$ then (x_1, x_2, x_3, x_4) is (H, T, H, T),(T, H, T, H), (H, T, T, H) or (T, H, H, T) out of the possible 16 outcomes. This entails that $Pr(X_k = 1 \wedge X_j = 1) = 4/16 = 1/4$ as desired.
- If $X_k = 0$ and $X_j = 1$ then $(x_1, x_2, x_3, x_4) = (H, H, H, T)$, (H, H, T, H), (T, T, H, T) or (T, T, T, H) and so $Pr(X_k = 0 \wedge X_j = 1) = 1/4$.
- If $X_k = 1$ and $X_j = 0$ then $(x_1, x_2, x_3, x_4) = (H, T, H, H)$, (T, H, H, H), (H, T, T, T) or (T, H, T, T, T) and so $Pr(X_k = 1 \wedge X_j = 0) = 1/4$.
- If $X_k = 0$ and $X_j = 0$ then $(x_1, x_2, x_3, x_4) = (H, H, H, H))$, (T, T, T, T), (H, H, T, T) or (T, T, H, H) and so $Pr(X_k = 0 \wedge X_j = 0) = 1/4$.

Thus all equations are also verified in this case.

Consider a triangle made from vertices u, v and w and edges $e_1 = \{u, v\}$, $e_2 = \{u, w\}$, and $e_3 = \{v, w\}$.
Let (x_1, x_2, x_3) represent the sequence of flips for u, v and w respectively.

If $X_1 = 1$ and $X_2 = 1$ then v and w must be in the same subset. Therefore $Pr(X_1 = 1 \wedge X_2 = 1 \wedge X_3 = 1) = 0$. It follows that X_1, X_2, and X_3 are not mutually independent (since $Pr(X_i) \neq 0$).

Question 4

Let G be a planar graph with at least two connected components.
We give a direct proof. A proof by contradiction is possible but boils down to arguing in a similar way based on the following two observations:

1. Each connected component with exactly one vertex does not contribute to the number of edges or faces of G.
2. Each connected component with exactly with two vertices contributes exactly by one to the number of edges of G, by two to the number of vertices, and zero to the number of faces.

Let e, f, v denote the number of edges, faces, and vertices of G. Consider the subgraph G' of G consisting only of the connected components of G with three or more vertices. Let v' , e' and f' be the number of vertices, edges and face of G' respectively.

We observe that we can apply Theorem 7.5.4 on each of the connected components of G', which results in the following bounds[1]

$$e' \leq 3v' - 6 \text{ and } f' \leq 2v' – 4$$

Now let c_1 be the number of connected components of G with exactly one vertex, c_2 be the number of connected components of G with exactly two vertices, and c_3 be the number of connected components of G with three or more vertices. We have that

$$\begin{aligned} e &= c_2 + e' \quad \text{by observations 1. and 2.} \\ &\leq c_2 + 3v' – 6 \\ &\leq 3v – 6 \quad \text{since } v = v' + 2c_2 + c_1 \end{aligned}$$

Similarly $f = f'$ by observation 1. and 2. So $f = f' \leq 2v' – 4 \leq 2v – 4$ since $v' \leq v$.

Question 5

K_5 has $\binom{5}{2} = 10$ edges. If K_5 has a plane embedding, then that embedding must have at most $3 \cdot 5 – 6 = 9$ edges by Theorem 7.5.4. Therefore, K_5 has no planar embedding.

Question 6

Let G be a graph with m edges.
For each vertex u of G we roll a fair three sided fair die, each roll is independent from the other.
If the die lands on 1 then u is added to A, if the die lands on 2 then u is added to B, and if the die lands on 3 u is added C.

Let X be the random variable corresponding to the number of edges of G between A, B and C. We compute $E(X)$. First we arbitrarily number the m edges of G as $e_1, \ldots, e_m$.
For $1 \leq k \leq m$ we define the following indicator variables (one for each edge of G):

$X_k = 1$ if e_k is an edge between A and B , or between A and C or between B and C. Otherwise $X_k = 0$.

We can observe that $X = \sum_{k=1}^{m} X_k$ (think about what both sides are counting).

[1] This requires a bit of work. The math checks out but there are some tiny cases to be careful about.

Therefore, $E(X) = \sum_{k=1}^{m} E(X_k)$ by linearity of expected value.

By definiton of expected value, $E(X_k) = 0 \cdot Pr(X_k = 0) + 1 \cdot Pr(X_k = 1) = Pr(X_k = 1)$.
Consider the two endpoints forming the edge e_k. Each endpoint goes into one of the 3 sets A, B or C. In particular there are nine different outcomes for both endpoints in total. Out of these nine outcomes, three of them result in the two endpoints being in the same set. For the other six, the endpoints are in different subsets and thus defining an edge between either A and B or A and C or B and C. Hence $Pr(X_k = 1) = 6/9 = 2/3$ which entails that $E(X) = \sum_{k=1}^{m} E(X_k) = \sum_{k=1}^{m} 2/3 = 2m/3$.

Now suppose that the claim does not hold. Then any partition of the edge set of G into three parts A, B and C contains less than $2m/3$ edges between A,B and C. In particular, $E(X)$ must be less than $2m/3$ which contradicts what we found above.

Question 7

Let $R(Z)$ denote the range of Z.
$E(Z) = \sum_{z \in R(Z)} zPr(Z = z)$ by definition. We first observe that $R(Z) = \{0, .., n-1\}$. We can thus rewitre $E(Z)$ as follows[2]:

$$E(Z) = \sum_{k=0}^{n-1} k \cdot Pr(Z = k) = \sum_{k=1}^{n-1} kPr(Z = k)$$

We can describe the sample space $S = \{(x_1, x_2) : x_1, x_2 \in \{1, .., n\}\}$. By the product rule we can see that $|S| = n^2$. We now show that $Pr(Z = k) = 2(n-k)/n^2$ when $k \geq 1$.

We are counting the elements in S whose absolute difference is equal to k.
Case 1: Count the elements of S such that $|x_1 - x_2| = k$ and $x_1 < x_2$.

If $x_1 < x_2$ then $|x_1 - x_2| = x_2 - x_1 = k$. This entails that $x_2 - k = x_1$. Thus we have $n - k$ options for x_1. Once x_1 is chosen, there only one value x_2 such that $x_2 - x_1 = k$.

Case 2: Count the elements of S such that $|x_1 - x_2| = k$ and $x_2 < x_1$.

By symmetry we get also $n - k$ options.
By the sum rule, there is $2(n - k)$ elements in S whose absolute difference is equal to k.
Thus, $E(Z) = \sum_{k=1}^{n-1} k2(n-k)/n^2$. Note that this can be rewritten as $1/n^2\big(2n \sum_{k=1}^{n-1} k - 2\sum_{k=1}^{n-1} k^2\big)$. Using the identities proved in Assignment 2 Question 6 and Question 7, this can be reduced to $n^2 - 1/3n$.

[2]This line is not necessary. However it does make it clear that we only consider integer values.