Explanatory Solution of Assignment 9-9

1. Consider a group of 4 people. Each person randomly sends a like to one of the rest 3 people (i.e. with probability $\frac{1}{3}$ each). We say a person is popular if the person received at least 2 likes. Let N denote the number of popular people among the four.

   (a) Find $E[N]$

   (b) Find $Var(N)$

   (c) Find $P(N=0)$

   Hint: Let X_i denote the event that i is a popular person, $i = 1,2,3,4$. Let $I_i \triangleq I(X_i)$ denote the indicator random variable that equals 1 if X_i is true and 0 otherwise. By definition, $N = I_1 + I_2 + I_3 + I_4$.

   Solution: Let four people named A,B,C,D. Then it is operated as

   - A send a like to one of (B,C,D)
   - B send a like to one of (A,C,D)
   - C send a like to one of (A,B,D)
   - D send a like to one of (A,B,C)

   There are different choices of likes, and the total number of outcomes is $3 \times 3 \times 3 \times 3 = 81$.

   If we take the perspective of the receivers, then we find that

   - A might receive 0,1,2, or 3 likes from (B,C,D), let L_A be the number of likes that A receives.
   - B might receive 0,1,2, or 3 likes from (A,C,D), let L_B be the number of likes that B receives
   - C might receive 0,1,2, or 3 likes from (A,B,D),let L_C be the number of likes that C receives.
   - D might receive 0,1,2, or 3 likes from (A,B,C),let L_D be the number of likes that D receives.

   Take L_A as an example. The results of L_A is determined by the choices of (B,C,D), and they have totally $3 \times 3 \times 3 = 27$ outcomes. Then it can be calculated that

$$p(L_A = 0) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{8}{27}$$
$$p(L_A = 1) = 3 \times \left(\frac{1}{3} \times \frac{2}{3} \times \frac{2}{3}\right) = \frac{12}{27}$$
$$p(L_A = 2) = 3 \times \left(\frac{1}{3} \times \frac{1}{3} \times \frac{2}{3}\right) = \frac{6}{27}$$
$$p(L_A = 3) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27}$$

   This probability also works for L_B, L_C, and L_D. It should be noted that L_A, L_B, L_C, and L_D, and it is difficult to compute the joint pmf of (L_A, L_B, L_C, L_D).

   In the problem, N denotes the number of popular people, i.e.,

$$N = I_{L_A \geq 2} + I_{L_B \geq 2} + I_{L_C \geq 2} + I_{L_D \geq 2} \tag{1}$$

   where I_* is the indicator function here (it equals 1 if the conditions $*$ hold, and 0 otherwise).

   (a) (4 points) Therefore,

$$\begin{aligned} E[N] &= E[I_1 + I_2 + I_3 + I_4] \\ &= E[I_1] + E[I_2] + E[I_3] + E[I_4] \\ &= P(X_1) + P(X_2) + P(X_3) + P(X_4). \end{aligned}$$

We compute $P(X_1)$. X_1 is true $\Leftrightarrow$ Person 1 received at least two likes $\Leftrightarrow$ Person 1 received likes from $(2,3),(2,4),(3,4)$ or $(2,3,4)$. $(2,3)$ likes 1 and 4 does not like 1 happens with probability $\frac{1}{3}\times\frac{1}{3}\times\frac{2}{3}=\frac{2}{27}$. This is the same for $(2,4)$ and $(3,4)$. $(2,3,4)$ is with probability $\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{27}$. Therefore, $P(X_1)=\frac{7}{27}$. Since X_i are symmetric across $i=1,2,3,4$, we thus have $E[N]=4\times\frac{7}{27}=\frac{28}{27}$.

We find that N is a well-defined random variable. From (1), it is shown that the pmf of N depends on the joint pmf of (L_A, L_B, L_C, L_D), and it is hard to compute. However, the expectation of N is simple to compute by using the linearity of the expectation. This is the result of part (a). Note $P(X_1)=p(L_A=2)+p(L_A=3)$.

(b) (4 points) $Var(N)=E[N^2]-(E[N])^2$. Let's compute $E[N^2]$. In particular,

$$\begin{aligned} E[N^2] &= E[(I_1+I_2+I_3+I_4)^2] \\ &= E[I_1^2]+E[I_2^2]+E[I_3^2]+E[I_4^2]+2\sum_{1\le i<j\le 4}E[I_i\times I_j] \\ &= E[I_1]+E[I_2]+E[I_3]+E[I_4]+2\sum_{1\le i<j\le 4}E[I_i\times I_j] \\ &= E[N]+12E[I_1\times I_2] \end{aligned}$$

where the last equality is due to symmetry. Observe that $I_1\times I_2 \Leftrightarrow I(X_1\cap X_2)$. Therefore $E[I_1\times I_2]=P(X_1\cap X_2)$. Both Person 1 and 2 are popular, it must be the two of them received all 4 likes. In particular, Person 1 and person 2 liked each other, and person 3 and person 4 liked person 1 and person 2 (or person 2 and person 1), respectively. Thus $P(X_1\cap X_2)=\frac{2}{81}$. Combining the above, $E[N^2]=\frac{28}{27}+12\times\frac{2}{81}=\frac{4}{3}$. And $Var(N)=\frac{4}{3}-(\frac{28}{27})^2=\frac{188}{729}$

The linearity of the expectation can also be applied here, but we need to compute $\mathbb{E}[I_1\times I_2]$ (A and B are popular). In this case, the number of likes for (A,B,C,D) must be (2,2,0,0). How many outcomes (totally $3\times3\times3\times3=81$ as we show in the beginning) would lead to (2,2,0,0)?

- A must like B
- B must like A
- C must like A or B
- D must the remaining one of A or B

Then it has 2 outcomes, it accounts for $P(X_1\cap X_2)=\frac{2}{81}$.

(c) (2 points) $N=0$ if and only if everyone received exactly one like. We count the number of samples that satisfy this. There are two cases. Either there's only one cycle, in which case we have $(1\to2\to3\to4)$ type of patterns. There are in total $3\times2\times1=6$ possibilities. Or there are two cycles, in which case we have $(1\leftrightarrow2, 3\leftrightarrow4)$ type of patterns. There are in total $\binom{4}{2}/2=3$ possibilities. In all there are 9 samples out of 81 that nobody is popular. Thus $P(N=0)=\frac{9}{81}=\frac{1}{9}$.

Concluding remark: Sometimes the joint pmf/pdf is hard to compute, and then the indicator function is a widely used technique to compute the mean or variance with the marginal pmf rather than the joint pmf. This is a very interesting technique and seems to be very counter-intuitive. You can find many other examples online by searching the key word "indicator" and "linearity of expectation". We also provide a similar problem in Tutorial 11-4.