Tutorial 5

October 22, 2020

Question 1

A discrete random variable N is uniformly distributed on $\{1, 2, 3, \dots, 10\}$.

Let X be the indicator of the event $\{N \leq 5\}$.

Let Y be the indicator of the event $\{N \text{ is even }\}$.

(a) Are X and Y independent?

$$X = \begin{cases} 1 & N \leq 5 \\ 0 & \text{otherwise} \end{cases} \qquad Y = \begin{cases} 1 & N \text{ even} \\ 0 & \text{otherwise} \end{cases}$$

We can check for independence of X and Y by checking

$$\mathbf{P}\left(X = 1 \cap Y = 1\right) \stackrel{?}{=} \mathbf{P}\left(X = 1\right) \cdot \mathbf{P}\left(Y = 1\right)$$

$$\begin{aligned} \mathbf{P}\left(X = 1\right) &= \mathbf{P}\left(N \leq 5\right) \\ &= \frac{5}{10} \end{aligned}$$

$$\begin{aligned} \mathbf{P}\left(Y = 1\right) &= \mathbf{P}\left(N \text{ even}\right) \\ &= \frac{5}{10} \end{aligned}$$

$$\begin{aligned} \mathbf{P}\left(X = 1 \cap Y = 1\right) &= \mathbf{P}\left(N \leq 5 \cap N \text{ even}\right) \\ &= \frac{2}{10} \end{aligned}$$

Since $\mathbf{P}\left(X = 1 \cap Y = 1\right) \neq \mathbf{P}\left(X = 1\right) \cdot \mathbf{P}\left(Y = 1\right)$, X and Y are not independent.

(b) Find $\mathbf{E}\left((X + Y)^2\right)$.

First note that

$$XY = \begin{cases} 1 & N \leq 5 \cap N \text{ even} \\ 0 & \text{otherwise} \end{cases}$$

It follows that

$$\begin{aligned} \mathbf{E}\left(XY\right) &= 1 \cdot \mathbf{P}\left(N \leq 5 \cap N \text{ even}\right) \; + \; 0 \cdot \mathbf{P}\left((N \leq 5 \cap N \text{ even})^c\right) \\ &= \mathbf{P}\left(N \leq 5 \cap N \text{ even}\right) \end{aligned}$$

It should also be noted that the square of an indicator variable is identical to the original indicator variable.

$$\begin{aligned}
\mathbf{E}\left((X+Y)^2\right) &= \mathbf{E}\left(X^2+2XY+Y^2\right) \\
&= \mathbf{E}\left(X^2\right) + 2\mathbf{E}\left(XY\right) + \mathbf{E}\left(Y^2\right) \\
&= \mathbf{E}\left(X\right) + 2\mathbf{E}\left(XY\right) + \mathbf{E}\left(Y\right) \\
&= \mathbf{P}\left(N \leq 5\right) + 2\mathbf{P}\left(N \leq 5 \cap N \text{ even}\right) + \mathbf{P}\left(N \text{ even}\right) \\
&= \frac{5}{10} + 2\left(\frac{2}{10}\right) + \frac{5}{10} \\
&= \frac{14}{10} = 1.4
\end{aligned}$$

Question 2

13 cards are drawn at random without replacement from an ordinary deck of playing cards. If X is the number of spades in these 13 cards, find the PMF of X. If, in addition, Y is the number of hearts in these 13 cards, find the probability $\mathbf{P}\left(X=2,\, Y=5\right)$. What is the joint PMF of X and Y?

Let X be the number of spades in the 13 cards drawn at random without replacement from an ordinary deck of 52 cards. As there are 13 spades and 39 non-spades in an ordinary deck, the PMF of X is

$$\mathbf{P}\left(X=x\right) = \begin{cases} \dfrac{\binom{13}{x}\binom{39}{13-x}}{\binom{52}{13}} & 0 \leq x \leq 13 \\ 0 & \text{otherwise} \end{cases}$$

Now, let Y be the number of hearts contained in these 13 cards. There are 13 hearts in an ordinary deck, leaving 26 cards that are non-spade and non-heart. The joint probability mass function is

$$\mathbf{P}\left(X=x,\, Y=y\right) = \begin{cases} \dfrac{\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}}{\binom{52}{13}} & 0 \leq x \leq 13,\, 0 \leq y \leq 13-x \\ 0 & \text{otherwise} \end{cases}$$

for $x,\, y \in \mathbb{Z}$.

The desired probability is computed as:

$$\mathbf{P}\left(X=2,\, Y=5\right) = \frac{\binom{13}{2}\binom{13}{5}\binom{26}{13-7}}{\binom{52}{13}}$$

$$= \frac{\frac{13!}{2!11!}\,\frac{13!}{5!8!}\,\frac{26!}{6!20!}}{\frac{52!}{13!39!}}$$

$$= 3.647 \times 10^{-7}$$

Question 3

Consider the multinomial distribution:

- $m \geq 2$ categories
- $n \geq 1$ items chosen at random, **with** replacement
- $p_k = \mathbf{P}\,(\text{Item of type } k \text{ chosen}),\ k = 1, \ldots, m$
- $X_k = \text{Number of type } k \text{ chosen},\ k = 1, \ldots, m$

(a) Compute $\mathbf{P}\,(X_1 = x_1,\, X_2 = x_2,\, \ldots,\, X_m = x_m)$.

Computing this probability is equivalent to finding the joint PMF.

$$\mathbf{P}\,(X_1 = x_1,\, X_2 = x_2,\, \ldots,\, X_m = x_m)$$

$$= \binom{n}{x_1}\binom{n - x_1}{x_2} \cdots \binom{n - x_1 - \ldots - x_{m-1}}{x_m} p_1^{x_1}\, p_2^{x_2} \cdots p_m^{x_m}$$

$$= \frac{n!}{x_1!\, x_2! \cdots x_m!}\, p_1^{x_1}\, p_2^{x_2} \cdots p_m^{x_m}$$

provided that $\sum_{i=1}^{m} x_i = n$.

The full form of the joint PMF is:

$$p\,(x_1,\, x_2,\, \ldots,\, x_m) = \begin{cases} \dfrac{n!}{x_1!\, x_2! \cdots x_m!}\, p_1^{x_1}\, p_2^{x_2} \cdots p_m^{x_m} & \sum_{i=1}^{m} x_i = n \\ \\ 0 & \text{otherwise} \end{cases}$$

(b) Find the marginal distribution of X_k for each k. Are X_i and X_j independent?

Intuitively, we expect that X_k will have a binomial distribution with parameters $(n,\, p_k)$. This can be shown formally. Without loss of generality, take $k = 1$. We know that to find the marginal distribution of X_1, we will need to sum over everything except X_1, i.e. X_2 through X_m.

$$
\begin{aligned}
p_{X_1}(x_1) &= \sum_{\substack{x_k \\ 2 \le k \le m}} \frac{n!}{x_1!\, x_2! \,\cdots\, x_m!}\, p_1^{x_1}\, p_2^{x_2} \,\cdots\, p_m^{x_m} \\
&= \sum_{\substack{x_k \\ 2 \le k \le m}} \frac{n!}{(n-x_1)!\, x_1!} \cdot \frac{(n-x_1)!}{x_2!\, x_3! \,\cdots\, x_m!}\, p_1^{x_1}\, p_2^{x_2} \,\cdots\, p_m^{x_m} \\
&= \frac{n!}{(n-x_1)!\, x_1!}\, p_1^{x_1} \sum_{\substack{x_k \\ 2 \le k \le m}} \frac{(n-x_1)!}{x_2!\, x_3! \,\cdots\, x_m!}\, p_2^{x_2}\, p_3^{x_3} \,\cdots\, p_m^{x_m} \\
&= \binom{n}{x_1} p_1^{x_1}\, (p_2 + p_3 + \ldots + p_m)^{n-x_1} && \text{(Multinomial Theorem)} \\
&= \binom{n}{x_1} p_1^{x_1}\, (1-p_1)^{n-x_1}
\end{aligned}
$$

Thus we have

$$
p_{X_1}(x_1) \;=\; \begin{cases} \dbinom{n}{x_1} p_1^{x_1}\, (1-p_1)^{n-x_1} & 0 \le x_1 \le n \\ 0 & \text{otherwise} \end{cases}
$$

Iterating through the remaining $k = 2, \ldots m$, it can be shown that each X_k will have a binomial distribution with parameters $(n,\, p_k)$.

While the results of the individual trials still remain independent, X_i will not be independent of X_j since

$$
\mathbf{P}\left(X_i = n,\, X_j = n\right) \;=\; 0 \;\ne\; \mathbf{P}\left(X_i = n\right) \cdot \mathbf{P}\left(X_j = n\right)
$$

Question 4

Let $(X_1,\, X_2,\, X_3) \;\sim\; \text{Multi}(n,\, p_1,\, p_2,\, p_3)$. Find the conditional distribution of X_1 given that $X_3 = x_3$. Intuitively, we expect that

$$
X_1 \,|\, X_3 = x_3 \;\sim\; \text{Binomial}\left(n - x_3,\, \frac{p_1}{p_1 + p_2}\right)
$$

$$
\begin{aligned}
\mathbf{P}\left(X_1 \,|\, X_3 = x_3\right) &= \frac{\mathbf{P}\left(X_1 = x_1,\, X_3 = x_3\right)}{\mathbf{P}\left(X_3 = x_3\right)} \\
&= \frac{\mathbf{P}\left(X_1 = x_1,\, X_2 = n - x_1 - x_3,\, X_3 = x_3\right)}{\mathbf{P}\left(X_3 = x_3\right)} \\
&= \frac{\dfrac{n!}{x_1!\,(n-x_1-x_3)!\, x_3!}\, p_1^{x_1}\, p_2^{n-x_1-x_3}\, p_3^{x_3}}{\dfrac{n!}{(n-x_3)!\, x_3!}\, p_3^{x_3}\, (1-p_3)^{n-x_3}}
\end{aligned}
$$

$$= \frac{(n-x_3)!}{(n-x_1-x_3)!\,x_1!}\,\frac{p_1^{x_1}\,p_2^{x_2}}{(p_1+p_2)^{x_1+x_2}}$$

$$= \binom{n-x_3}{x_1}\left(\frac{p_1}{p_1+p_2}\right)^{x_1}\left(\frac{p_2}{p_1+p_2}\right)^{x_2}$$

$$= \binom{n-x_3}{x_1}\left(\frac{p_1}{p_1+p_2}\right)^{x_1}\left(\frac{p_2}{p_1+p_2}\right)^{n-x_3-x_1}$$

for $0 \le x_1 \le n - x_3$, and zero elsewhere.

As expected,

$$X_1 \,|\, X_3 = x_3 \sim \text{Binomial}\left(n - x_3,\, \frac{p_1}{p_1+p_2}\right)$$

Question 5

Suppose $X \sim \text{Bin}(N, p)$, where the number of trials, N, is also a random variable (but independent of the trials themselves). Then conditioned on the fact that $N = n$, the number of successes, X, would have distribution $\text{Bin}(n, p)$. What can be said about the unconditional distribution of X, in particular the case when N is a Poisson random variable?

This is actually the exact same question as Tutorial 3 Question 1, but worded slightly differently... We saw in Tutorial 3 Question 1 that if

$$N \sim \text{Poisson}(\lambda)$$

and the conditional distribution of $X \,|\, N = n$ was

$$X \,|\, N = n \sim \text{Binomial}(n,\, p)$$

then applying the law of total probability, the unconditional distribution of X was

$$X \sim \text{Poisson}(\lambda\, p)$$

Question 6

Following the setup of the previous question, let $Y = N - X$ represent the number of failures. It is implied that Y has (unconditional) distribution $\text{Poisson}(\lambda \cdot (1-p))$. Show that X and Y are independent. [Note that this is strongly due to the Poisson distribution of N, and does not happen otherwise (i.e. with deterministic N).]

We can show that X and Y are independent by finding their joint PMF. Before beginning, we recall that

$$X \,|\, N = n \sim \text{Binomial}(n,\, p)$$

$$Y \,|\, N = n \sim \text{Binomial}(n,\, (1-p))$$

$$
\begin{aligned}
\mathbf{P}(X = x, Y = y) &= \mathbf{P}(X = x, Y = y \,|\, N = n) \cdot \mathbf{P}(N = n) \\
&= \mathbf{P}(X = x, Y = y \,|\, N = x + y) \cdot \mathbf{P}(N = x + y) \\
&= \frac{(x+y)!}{x!\,y!} p^x (1-p)^y \cdot e^{-\lambda} \frac{\lambda^{x+y}}{(x+y)!} \\
&= e^{-\lambda} \cdot \frac{\lambda^x p^x}{x!} \cdot \frac{\lambda^y (1-p)^y}{y!} \\
&= e^{-\lambda + \lambda p - \lambda p} \cdot \frac{(\lambda p)^x}{x!} \cdot \frac{(\lambda(1-p))^y}{y!} \\
&= e^{-\lambda p - \lambda(1-p)} \cdot \frac{(\lambda p)^x}{x!} \cdot \frac{(\lambda(1-p))^y}{y!} \\
&= e^{-\lambda p} \frac{(\lambda p)^x}{x!} \cdot e^{-\lambda(1-p)} \frac{(\lambda(1-p))^y}{y!}
\end{aligned}
$$

for $x \geq 0$, $y \geq 0$, and zero otherwise.

Since the joint PMF of X and Y is a product of their (unconditional) marginal PMFs, X and Y are independent.