3.1.6 Solved Problems: Discrete Random Variables

Problem 1

Let X be a discrete random variable with the following PMF

$$P_X(x) = \begin{cases} 0.1 & \text{for } x = 0.2 \\ 0.2 & \text{for } x = 0.4 \\ 0.2 & \text{for } x = 0.5 \\ 0.3 & \text{for } x = 0.8 \\ 0.2 & \text{for } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

a. Find R_X, the range of the random variable X.
b. Find $P(X \leq 0.5)$.
c. Find $P(0.25 < X < 0.75)$.
d. Find $P(X = 0.2|X < 0.6)$.

Solution

a. The range of X can be found from the PMF. The range of X consists of possible values for X. Here we have

$$R_X = \{0.2, 0.4, 0.5, 0.8, 1\}.$$

b. The event $X \leq 0.5$ can happen only if X is $0.2, 0.4$, or 0.5. Thus,

$$\begin{aligned} P(X \leq 0.5) &= P(X \in \{0.2, 0.4, 0.5\}) \\ &= P(X = 0.2) + P(X = 0.4) + P(X = 0.5) \\ &= P_X(0.2) + P_X(0.4) + P_X(0.5) \\ &= 0.1 + 0.2 + 0.2 = 0.5 \end{aligned}$$

c. Similarly, we have

$$\begin{aligned} P(0.25 < X < 0.75) &= P(X \in \{0.4, 0.5\}) \\ &= P(X = 0.4) + P(X = 0.5) \\ &= P_X(0.4) + P_X(0.5) \\ &= 0.2 + 0.2 = 0.4 \end{aligned}$$

d. This is a conditional probability problem, so we can use our famous formula $P(A|B) = \frac{P(A\cap B)}{P(B)}$. We have

$$\begin{aligned} P(X=0.2|X<0.6) &= \frac{P\big((X=0.2) \text{ and } (X<0.6)\big)}{P(X<0.6)} \\ &= \frac{P(X=0.2)}{P(X<0.6)} \\ &= \frac{P_X(0.2)}{P_X(0.2)+P_X(0.4)+P_X(0.5)} \\ &= \frac{0.1}{0.1+0.2+0.2} = 0.2 \end{aligned}$$

Problem 2

I roll two dice and observe two numbers X and Y.

a. Find R_X, R_Y and the PMFs of X and Y.
b. Find $P(X=2, Y=6)$.
c. Find $P(X>3|Y=2)$.
d. Let $Z = X+Y$. Find the range and PMF of Z.
e. Find $P(X=4|Z=8)$.

Solution

a. We have $R_X = R_Y = \{1,2,3,4,5,6\}$. Assuming the dice are fair, all values are equally likely so

$$P_X(k) = \begin{cases} \frac{1}{6} & \text{for } k=1,2,3,4,5,6 \\ 0 & \text{otherwise} \end{cases}$$

Similarly for Y,

$$P_Y(k) = \begin{cases} \frac{1}{6} & \text{for } k=1,2,3,4,5,6 \\ 0 & \text{otherwise} \end{cases}$$

b. Since X and Y are independent random variables, we can write

$$\begin{aligned} P(X=2, Y=6) &= P(X=2)P(Y=6) \\ &= \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}. \end{aligned}$$

c. Since X and Y are independent, knowing the value of X does not impact the probabilities for Y,

$$\begin{aligned} P(X > 3|Y = 2) &= P(X > 3) \\ &= P_X(4) + P_X(5) + P_X(6) \\ &= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{2}. \end{aligned}$$

d. First, we have $R_Z = \{2, 3, 4, \ldots, 12\}$. Thus, we need to find $P_Z(k)$ for $k = 2, 3, \ldots, 12$. We have

$$\begin{aligned} P_Z(2) &= P(Z = 2) = P(X = 1, Y = 1) \\ &= \tfrac{1}{6} \cdot \tfrac{1}{6} = \tfrac{1}{36}; \\ P_Z(3) &= P(Z = 3) = P(X = 1, Y = 2) + P(X = 2, Y = 1) \\ &= P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1) \\ &= \tfrac{1}{6} \cdot \tfrac{1}{6} + \tfrac{1}{6} \cdot \tfrac{1}{6} = \tfrac{1}{18}; \\ P_Z(4) &= P(Z = 4) = P(X = 1, Y = 3) + P(X = 2, Y = 2) + P(X = 3, Y = 1) \\ &= 3 \cdot \tfrac{1}{36} = \tfrac{1}{12}. \end{aligned}$$

We can continue similarly:

$$\begin{aligned} P_Z(5) &= \tfrac{4}{36} = \tfrac{1}{9}; \\ P_Z(6) &= \tfrac{5}{36}; \\ P_Z(7) &= \tfrac{6}{36} = \tfrac{1}{6}; \\ P_Z(8) &= \tfrac{5}{36}; \\ P_Z(9) &= \tfrac{4}{36} = \tfrac{1}{9}; \\ P_Z(10) &= \tfrac{3}{36} = \tfrac{1}{12}; \\ P_Z(11) &= \tfrac{2}{36} = \tfrac{1}{18}; \\ P_Z(12) &= \tfrac{1}{36}. \end{aligned}$$

It is always a good idea to check our answers by verifying that $\sum_{z \in R_Z} P_Z(z) = 1$. Here, we have

$$\begin{aligned} \textstyle\sum_{z \in R_Z} P_Z(z) &= \tfrac{1}{36} + \tfrac{2}{36} + \tfrac{3}{36} + \tfrac{4}{36} + \tfrac{5}{36} + \tfrac{6}{36} \\ &+ \tfrac{5}{36} + \tfrac{4}{36} + \tfrac{3}{36} + \tfrac{2}{36} + \tfrac{1}{36} \\ &= 1. \end{aligned}$$

e. Note that here we cannot argue that X and Z are independent. Indeed, Z seems to completely depend on X, $Z = X + Y$. To find the conditional probability $P(X = 4|Z = 8)$, we use the formula for conditional probability

$$
\begin{aligned}
P(X=4|Z=8) &= \frac{P(X=4, Z=8)}{P(Z=8)} \\
&= \frac{P(X=4, Y=4)}{P(Z=8)} \\
& \frac{\frac{1}{6} \cdot \frac{1}{6}}{\frac{5}{36}} \\
&= \frac{1}{5}.
\end{aligned}
$$

Problem 3

I roll a fair die repeatedly until a number larger than 4 is observed. If N is the total number of times that I roll the die, find $P(N=k)$, for $k = 1, 2, 3, \ldots$.

Solution

In each trial, I may observe a number larger than 4 with probability $\frac{2}{6} = \frac{1}{3}$. Thus, you can think of this experiment as repeating a Bernoulli experiment with success probability $p = \frac{1}{3}$ until you observe the first success. Thus, N is a geometric random variable with parameter $p = \frac{1}{3}$, $N \sim Geometric(\frac{1}{3})$. Hence, we have

$$
P_N(k) = \begin{cases} \frac{1}{3}(\frac{2}{3})^{k-1} & \text{for } k = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}
$$

Problem 4

You take an exam that contains 20 multiple-choice questions. Each question has 4 possible options. You know the answer to 10 questions, but you have no idea about the other 10 questions so you choose answers randomly. Your score X on the exam is the total number of correct answers. Find the PMF of X. What is $P(X > 15)$?

Solution

Let's define the random variable Y as the number of your correct answers to the 10 questions you answer randomly. Then your total score will be $X = Y + 10$. First, let's find the PMF of Y. For each question your success probability is $\frac{1}{4}$. Hence, you perform 10 independent $Bernoulli(\frac{1}{4})$ trials and Y is the number of successes. Thus, we conclude $Y \sim Binomial(10, \frac{1}{4})$, so

$$P_Y(y) = \begin{cases} \binom{10}{y}(\frac{1}{4})^y(\frac{3}{4})^{10-y} & \text{for } y = 0, 1, 2, 3, \ldots, 10 \\ 0 & \text{otherwise} \end{cases}$$

Now we need to find the PMF of $X = Y + 10$. First note that $R_X = \{10, 11, 12, \ldots, 20\}$. We can write

$$\begin{aligned} P_X(10) &= P(X = 10) = P(Y + 10 = 10) \\ &= P(Y = 0) = \binom{10}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^{10-0} = \left(\frac{3}{4}\right)^{10}; \\ P_X(11) &= P(X = 11) = P(Y + 10 = 11) \\ &= P(Y = 1) = \binom{10}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^{10-1} = 10\frac{1}{4}\left(\frac{3}{4}\right)^9. \end{aligned}$$

So, you get the idea. In general for $k \in R_X = \{10, 11, 12, \ldots, 20\}$,

$$\begin{aligned} P_X(k) &= P(X = k) = P(Y + 10 = k) \\ &= P(Y = k - 10) = \binom{10}{k-10}\left(\frac{1}{4}\right)^{k-10}\left(\frac{3}{4}\right)^{20-k}. \end{aligned}$$

To summarize,

$$P_X(k) = \begin{cases} \binom{10}{k-10}(\frac{1}{4})^{k-10}(\frac{3}{4})^{20-k} & \text{for } k = 10, 11, 12, \ldots, 20 \\ 0 & \text{otherwise} \end{cases}$$

In order to calculate $P(X > 15)$, we know we should consider $y = 6, 7, 8, 9, 10$

$$P_Y(y) = \begin{cases} \binom{10}{y}(\frac{1}{4})^y(\frac{3}{4})^{10-y} & \text{for } y = 6, 7, 8, 9, 10 \\ 0 & \text{otherwise} \end{cases}$$

$$P_X(k) = \begin{cases} \binom{10}{k-10}(\frac{1}{4})^{k-10}(\frac{3}{4})^{20-k} & \text{for } k = 16, 17, \ldots, 20 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} P(X > 15) &= P_X(16) + P_X(17) + P_X(18) + P_X(19) + P_X(20) \\ &= \binom{10}{6}(\frac{1}{4})^6(\frac{3}{4})^4 + \binom{10}{7}(\frac{1}{4})^7(\frac{3}{4})^3 + \binom{10}{8}(\frac{1}{4})^8(\frac{3}{4})^2 \\ &+ \binom{10}{9}(\frac{1}{4})^9(\frac{3}{4})^1 + \binom{10}{10}(\frac{1}{4})^{10}(\frac{3}{4})^0. \end{aligned}$$

Problem 5

Let $X \sim Pascal(m,p)$ and $Y \sim Pascal(l,p)$ be two independent random variables. Define a new random variable as $Z = X + Y$. Find the PMF of Z.

Solution

This problem is very similar to Example 3.7, and we can solve it using the same methods. We will show that $Z \sim Pascal(m+l,p)$. To see this, consider a sequence of Hs and Ts that is the result of independent coin tosses with $P(H) = p$, (Figure 3.2). If we define the random variable X as the number of coin tosses until the mth heads is observed, then $X \sim Pascal(m,p)$. Now, if we look at the rest of the sequence and count the number of heads until we observe l more heads, then the number of coin tosses in this part of the sequence is $Y \sim Pascal(l,p)$. Looking from the beginning, we have repeatedly tossed the coin until we have observed $m+l$ heads. Thus, we conclude the random variable Z defined as $Z = X + Y$ has a $Pascal(m+l,p)$ distribution.

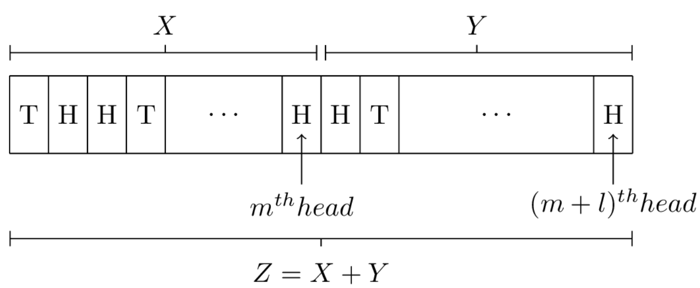

Fig.3.2 - Sum of two Pascal random variables.

In particular, remember that $Pascal(1,p) = Geometric(p)$. Thus, we have shown that if X and Y are two independent $Geometric(p)$ random variables, then $X + Y$ is a $Pascal(2,p)$ random variable. More generally, we can say that if $X_1, X_2, X_3, \ldots, X_m$ are m independent $Geometric(p)$ random variables, then the random variable X defined by $X = X_1 + X_2 + \ldots + X_m$ has a $Pascal(m,p)$ distribution.

Problem 6

The number of customers arriving at a grocery store is a Poisson random variable. On average 10 customers arrive per hour. Let X be the number of customers arriving from $10am$ to $11:30am$. What is $P(10 < X \leq 15)$?

Solution

We are looking at an interval of length 1.5 hours, so the number of customers in this interval is $X \sim Poisson(\lambda = 1.5 \times 10 = 15)$. Thus,

$$\begin{aligned}
P(10 < X \leq 15) &= \sum_{k=11}^{15} P_X(k) \\
&= \sum_{k=11}^{15} \frac{e^{-15}15^k}{k!} \\
&= e^{-15}\left[\frac{15^{11}}{11!} + \frac{15^{12}}{12!} + \frac{15^{13}}{13!} + \frac{15^{14}}{14!} + \frac{15^{15}}{15!}\right] \\
&= 0.4496
\end{aligned}$$

Problem 7

Let $X \sim Poisson(\alpha)$ and $Y \sim Poisson(\beta)$ be two independent random variables. Define a new random variable as $Z = X + Y$. Find the PMF of Z.

Solution

First note that since $R_X = \{0, 1, 2, ..\}$ and $R_Y = \{0, 1, 2, ..\}$, we can write $R_Z = \{0, 1, 2, ..\}$. We have

$$\begin{aligned}
P_Z(k) &= P(X + Y = k) \\
&= \sum_{i=0}^{k} P(X + Y = k | X = i)P(X = i) \quad \text{(law of total probability)} \\
&= \sum_{i=0}^{k} P(Y = k - i | X = i)P(X = i) \\
&= \sum_{i=0}^{k} P(Y = k - i)P(X = i) \\
&= \sum_{i=0}^{k} \frac{e^{-\beta}\beta^{k-i}}{(k-i)!}\frac{e^{-\alpha}\alpha^i}{i!} \\
&= e^{-(\alpha+\beta)} \sum_{i=0}^{k} \frac{\alpha^i\beta^{k-i}}{(k-i)!i!} \\
&= \frac{e^{-(\alpha+\beta)}}{k!} \sum_{i=0}^{k} \frac{k!}{(k-i)!i!}\alpha^i\beta^{k-i} \\
&= \frac{e^{-(\alpha+\beta)}}{k!} \sum_{i=0}^{k} \binom{k}{i}\alpha^i\beta^{k-i} \\
&= \frac{e^{-(\alpha+\beta)}}{k!}(\alpha + \beta)^k \quad \text{(by the binomial theorem).}
\end{aligned}$$

Thus, we conclude that $Z \sim Poisson(\alpha + \beta)$.

Problem 8

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} \frac{1}{4} & \text{for } k=-2 \\ \frac{1}{8} & \text{for } k=-1 \\ \frac{1}{8} & \text{for } k=0 \\ \frac{1}{4} & \text{for } k=1 \\ \frac{1}{4} & \text{for } k=2 \\ 0 & \text{otherwise} \end{cases}$$

I define a new random variable Y as $Y=(X+1)^2$.

a. Find the range of Y.
b. Find the PMF of Y.

Solution

Here, the random variable Y is a function of the random variable X. This means that we perform the random experiment and obtain $X=x$, and then the value of Y is determined as $Y=(x+1)^2$. Since X is a random variable, Y is also a random variable.

a. To find R_Y, we note that $R_X=\{-2,-1,0,1,2\}$, and

$$\begin{aligned} R_Y &= \{y=(x+1)^2 | x \in R_X\} \\ &= \{0,1,4,9\}. \end{aligned}$$

b. Now that we have found $R_Y=\{0,1,4,9\}$, to find the PMF of Y we need to find $P_Y(0), P_Y(1), P_Y(4)$, and $P_Y(9)$:

$$\begin{aligned} P_Y(0) &= P(Y=0) = P((X+1)^2=0) \\ &= P(X=-1) = \tfrac{1}{8}; \\ P_Y(1) &= P(Y=1) = P((X+1)^2=1) \\ &= P\big((X=-2) \text{ or } (X=0)\big); \\ & P_X(-2)+P_X(0) = \tfrac{1}{4}+\tfrac{1}{8}=\tfrac{3}{8}; \\ P_Y(4) &= P(Y=4) = P((X+1)^2=4) \\ &= P(X=1) = \tfrac{1}{4}; \\ P_Y(9) &= P(Y=9) = P((X+1)^2=9) \\ &= P(X=2) = \tfrac{1}{4}. \end{aligned}$$

Again, it is always a good idea to check that $\sum_{y \in R_Y} P_Y(y) = 1$. We have

$$\sum_{y \in R_Y} P_Y(y) = \frac{1}{8} + \frac{3}{8} + \frac{1}{4} + \frac{1}{4} = 1.$$
