Department of Mathematics, IIT Madras
MA 2040 : Probability, Statistics and Stochastic Processes
Additional Problem Set - II Solutions

1. (a) Let X be the number of modems in use. For $k < 50$, the probability that $X = k$ is the same as the probability that k out of 1000 customers need a connection:

$$p_X(k) = \binom{1000}{k}(0.01)^k(0.99)^{1000-k}, \quad k = 0, 1, \ldots, 49.$$

The probability that $X = 50$, is the same as the probability that 50 or more out of 1000 customers need a connection:

$$p_X(50) = \sum_{k=50}^{1000} \binom{1000}{k}(0.01)^k(0.99)^{1000-k}.$$

(b) Let A be the event that there are more customers needing a connection than there are modems. Then,

$$P(A) = \sum_{k=51}^{1000} \binom{1000}{k}(0.01)^k(0.99)^{1000-k}.$$

2. The number of guests that have the same birthday as you is binomial with $p = 1/365$ and $n = 499$. Thus the probability that exactly one other guest has the same birthday is

$$\binom{499}{1}\frac{1}{365}\left(\frac{364}{365}\right)^{498} \approx 0.3486.$$

3. Let random variable X be the number of trials you need to open the door, and let K_i be the event that the i-th key selected opens the door.

(a) In case (i), we have

$$\begin{aligned} p_X(1) &= P(K_1) = 1/5 \\ p_X(2) &= P(K_1^c)P(K_2|K_1^c) = \frac{4}{5}\frac{1}{4} = 1/5 \\ p_X(3) &= P(K_1^c)P(K_2^c|K_1^c)P(K_3|K_1^c \cap K_2^c) = \frac{4}{5}\frac{3}{4}\frac{1}{3} = 1/5. \end{aligned}$$

Similarly, we see that the PMF of X is

$$p_X(x) = 1/5, \qquad x = 1, 2, 3, 4, 5.$$

In case (ii), X is a geometric random variable with $p = 1/5$, and its PMF is $p_X(k) = \frac{1}{5}(\frac{4}{5})^{k-1}, \quad k \geq 1$.

(b) In case (i), we have

$$\begin{aligned}
p_X(1) &= P(K_1) = 2/10 \\
p_X(2) &= P(K_1^c)P(K_2|K_1^c) = \frac{8}{10}\frac{2}{9} \\
p_X(3) &= P(K_1^c)P(K_2^c|K_1^c)P(K_3|K_1^c \cap K_2^c) = \frac{8}{10}\frac{7}{9}\frac{2}{8}.
\end{aligned}$$

Proceeding similarly, we see that the PMF of X is

$$p_X(x) = \frac{2(10-x)}{90}, \qquad x = 1, 2, 3, 4, 5.$$

In case (ii), X is a geometric random variable with $p = 1/5$. $p_X(k) = \frac{1}{5}(\frac{4}{5})^{k-1}, \qquad k \geq 1.$

4. (a) We will use the PMF for the number of girls among the natural children together with the formula for the PMF of a function of a random variable. Let N be the number of natural children that are girls. Then N has a binomial PMF

$$p_N(k) = \begin{cases} \binom{5}{k}(\frac{1}{2})^5, & 0 \leq k \leq 5 \\ 0, & \textit{otherwise.} \end{cases}$$

Let G be the number of girls out of the 7 children, so that $G = N + 2$. By applying the formula for the PMF of a function of a random variable, we have $p_G(g) = \sum_{\{n|n+2=g\}} p_N(n) = p_N(g-2)$. Thus

$$p_G(g) = \begin{cases} \binom{5}{g-2}(\frac{1}{2})^5, & 2 \leq g \leq 7 \\ 0, & \text{otherwise.} \end{cases}$$

(b) The random variable Y takes the values $k \ln a$, where $k = 1, \ldots, n$, if and only if $X = a^k$ or $X = a^{-k}$. Furthermore, Y takes the value 0, if and only if $X = 1$. Thus, we have

$$p_Y(y) = \begin{cases} \frac{2}{2n+1}, & y = \ln a, 2\ln a, \ldots, k \ln a \\ \frac{1}{2n+1}, & y = 0 \\ 0, & \textit{otherwise.} \end{cases}$$

5. The expected value of the gain for a single game is infinite since if X is your gain, then

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} 2^k 2^{-k} = \infty.$$

6. (a) There are 21 integer pairs (x, y) in the region $R = \{(x,y) | -2 \leq x \leq 4, -1 \leq y - x \leq 1\}$, so that the joint PMF of X and Y is

$$p_{X,Y}(x,y) = \begin{cases} \frac{1}{21}, & (x,y) \in R \\ 0, & \text{otherwise.} \end{cases}$$

For each x in the range $[-2, 4]$, there are three possible values of Y. Thus, we have

$$p_X(x) = \begin{cases} \frac{3}{21}, & x = -2, -1, 0, 1, 2, 3, 4. \\ 0, & \text{otherwise.} \end{cases}$$

The mean of X is the midpoint of the range $[-2, 4]$: $\mathbb{E}[X] = 1$. The marginal PMF of Y is

$$p_Y(y) = \begin{cases} \frac{1}{21}, & y = -3 \\ \frac{2}{21}, & y = -2 \\ \frac{3}{21}, & y = -1, 0, 1, 2, 3 \\ \frac{2}{21}, & y = 4 \\ \frac{1}{21}, & y = 5 \\ 0, & \text{otherwise.} \end{cases}$$

The mean of Y is

$$E[Y] = \frac{1}{21}(-3+5) + \frac{2}{21}(-2+4) + \frac{3}{21}(-1+1+2+3) = 1.$$

(b) The profit is given by $P = 100X + 200Y$, so that

$$\mathbb{E}[P] = 100\mathbb{E}[X] + 200\mathbb{E}[Y] = 100 + 200 = 300.$$

7. (a) Since all possible values of (I, J) are equally likely, we have

$$p_{I,J}(i,j) = \begin{cases} \frac{1}{\sum_{k=1}^{n} m_k}, & j \le m_i \\ 0, & \text{otherwise.} \end{cases}$$

The marginal PMFs are given by

$$\begin{aligned} p_I(i) &= \sum_{j=1}^{m} p_{I,J}(i,j) = \frac{m_i}{\sum_{k=1}^{n} m_k}, & i = 1, 2, \ldots, n, \\ p_J(j) &= \sum_{i=1}^{n} p_{I,J}(i,j) = \frac{l_j}{\sum_{k=1}^{n} m_k}, & j = 1, 2, \ldots, m. \end{aligned}$$

where l_j is the number of students that have answered question j, i.e., students i with $j \le m_i$.

(b) The expected value of the score of student i is the sum of the expected values $p_{ij}a + (1-p_{ij})b$ of the scores on questions j with $j = 1, 2, \ldots, m_i$. i.e.,

$$\sum_{j=1}^{m_i} (p_{ij}a + (1-p_{ij})b).$$

8. The marginal PMF p_Y is given by the binomial formula

$$p_Y(y) = \binom{4}{y}\left(\frac{1}{6}\right)^y\left(\frac{5}{6}\right)^{4-y}, \quad y = 0, 1, \ldots, 4.$$

To compute the conditional PMF $p_{X|Y}$, note that given that $Y = y$, X is the number of 1s in the remaining $4-y$ rolls, each of which can take the 5 values 1, 3, 4, 5, 6 with equal probability 1/5. Thus, the conditional PMF $p_{X|Y}$ is binomial with parameters $4-y$ and $p = 1/5$:

$$p_{X|Y}(x|y) = \binom{4-y}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{4-y-x},$$

for all nonnegative integers x and y such that $0 \le x + y \le 4$. The joint PMF is now given by

$$\begin{aligned} p_{X,Y}(x,y) &= p_Y(y)p_{X|Y}(x|y) \\ &= \binom{4}{y}\left(\frac{1}{6}\right)^y\left(\frac{5}{6}\right)^{4-y}\binom{4-y}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{4-y-x}, \end{aligned}$$

for all nonnegative integers x and y such that $0 \le x + y \le 4$. For other values of x and y, we have $p_{X,Y}(x,y) = 0$.

9. One possibility here is to calculate the PMF of X, the number of tosses until the game is over, and use it to compute $\mathbb{E}[X]$. However, with an unfair coin, this turns out to be cumbersome, so we argue by using the total expectation theorem and a suitable partition of the sample space. Let H_k (or T_k) be the event that a head (or a tail, respectively) comes at the kth toss. and let p (respectively, q) be the probability of H_k (respectively, T_k) Since H_1 and T_1 form a partition of the sample space, and $P(H_1) = p$ and $P(T_1) = q$, we have

$$\mathbb{E}[X] = p\mathbb{E}[X|H_1] + q\mathbb{E}[X|T_1]$$

Using again the total expectation theorem, we have

$$\mathbb{E}[X|H_1] = p\mathbb{E}[X|H_1 \cap H_2] + q\mathbb{E}[X|H_1 \cap T_2] = 2p + q(1 + \mathbb{E}[X|T_1]),$$

where we have used the fact $\mathbb{E}[X|H_1 \cap H_2] = 2$ (since the game ends after two successive heads), and

$$\mathbb{E}[X|H_1 \cap T_2] = 1 + \mathbb{E}[X|T_1]$$

(since if the game is not over, only the last toss matters in determining the number of additional tosses up to termination). Similarly, we obtain $\mathbb{E}[X|T_1] = 2q + p(1 + \mathbb{E}[X|H_1])$ Combining the above two relations, collecting terms, and using the fact $p + q = 1$, we obtain after some calculation $\mathbb{E}[X|T_1] = \frac{2+p^2}{1-pq}$ and similarly $\mathbb{E}[X|H_1] = \frac{2+q^2}{1-pq}$. Thus,

$$\mathbb{E}[X] = p.\frac{2+p^2}{1-pq} + q.\frac{2+q^2}{1-pq}$$

and finally, using the fact $p+q=1$, $\mathbb{E}[X] = \frac{2+pq}{1-pq}$.

In the case of a fair coin ($p = q = 1/2$), we obtain $\mathbb{E}[X] = 3$. It can also be verified that $2 \leq [X] \leq 3$ for all values of p.

10. Note that

$$P(X=i|X+Y=n) = \frac{P(X=i, X+Y=n)}{P(X+Y=n)} = \frac{P(X=i)P(Y=n-i)}{P(X+Y=n)} \quad and$$

$$P(X=i) = p(1-p)^{i-1}, \quad for \quad i \geq 1 \quad and \quad P(Y=n-i) = p(1-p)^{n-i-1}, \quad for \quad n-i \geq 1.$$

It follows that

$$P(X=i)P(Y=n-i) = \begin{cases} p^2(1-p)^{n-2}, & i = 1, 2, \ldots, n-1 \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, $P(X+Y=n) = (n-1)p^2(1-p)^{n-2}$. Hence,

$$P(X=i|X+Y=n) = \tfrac{1}{n-1}, \quad i = 1, 2, \ldots, n-1.$$