# Solution of Problem Set 2

## Problem 1:

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

$$R_X = \{0, 1, 2\}.$$

(b) The event $X \geq 1.5$ can happen only if $X$ is 2. Thus,

$$\begin{aligned} P(X \geq 1.5) &= P(X = 2) \\ &= P_X(2) = \frac{1}{6} \end{aligned} \tag{1}$$

(c) Similarly, we have

$$\begin{aligned} P(0 < X < 2) &= P(X = 1) \\ &= P_X(1) = \frac{1}{3} \end{aligned} \tag{2}$$

(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

$$\begin{aligned} P(X = 0|X < 2) &= \frac{P(X = 0, X < 2)}{P(X < 2)} \\ &= \frac{P(X = 0)}{P(X < 2)} \\ &= \frac{P_X(0)}{P_X(0) + P_X(1)} \\ &= \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{3}} = \frac{3}{5} \end{aligned}$$

## Problem2:

Let $A, B, C$ be, respectively, the events that Alice, Bob, Charlie finds a butterfly. Then

$$\begin{aligned} p_X(0) &= P(A^c)P(B^c)P(C^c) = 0.3424 \\ p_X(1) &= P(A)P(B^c)P(C^c) + P(A^c)P(B)P(C^c) + P(A^c)P(B^c)P(C) = 0.4644 \\ p_X(2) &= P(A)P(B)P(C^c) + P(A)P(B^c)P(C) + P(A^c)P(B)P(C) = 0.1741 \\ p_X(3) &= P(A)P(B)P(C) = 0.0191 \end{aligned}$$

**Problem3:**

The range of $X$ is possible values for the number of cars being repaired. Based on the information:

$$R_X = \{0, 1, 2, 3\}.$$

$$\begin{aligned}
P_X(1) &= P_X(2) \\
P_X(0) &= P_X(3) \\
P_X(2) &= \frac{1}{2}P_X(3) \\
P_X(1) + P_X(2) &= \frac{1}{2}(P_X(0) + P_X(3)) \\
&= \frac{1}{2}(2 \times P_X(0)) = \frac{1}{2}(2 \times P_X(3)) \qquad (3) \\
&= P_X(0) \qquad (4)
\end{aligned}$$

Let $P_X(1) = \alpha$.

Then:

$$\begin{cases} P_X(1) &= P_X(2) = \alpha \\ P_X(0) &= P_X(3) = 2\alpha \end{cases}$$

Then, we have the following equation:

$$\begin{aligned}
\sum_{k=0}^{3} P_X(k) = 1 \quad &\longrightarrow \quad 2\alpha + \alpha + \alpha + 2\alpha = 1 \\
&\longrightarrow \quad \alpha = \frac{1}{6}
\end{aligned} \qquad (5)$$

$$\begin{cases} P_X(0) &= P_X(3) = \frac{1}{3} \\ P_X(1) &= P_X(2) = \frac{1}{6} \end{cases}$$

## Problem 4:

(a) $X$ and $Y$ are two independent random variables. So:

$$\begin{aligned}
P(X \le 2 \text{ and } Y \le 2) &= P(X \le 2) \cdot P(Y \le 2) \\
&= (P_X(1) + P_X(2)) \cdot (P_Y(1) + P_Y(2)) && (23) \\
&= (\frac{1}{4} + \frac{1}{8})(\frac{1}{6} + \frac{1}{6}) && (24) \\
&= \frac{3}{8} \cdot \frac{1}{3} = \frac{1}{8} && (25)
\end{aligned}$$

(b) Using $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ and the fact that $X$ and $Y$ are two independent random variables:

$$\begin{aligned}
P(X > 2 \text{ or } Y > 2) &= P(X > 2) + P(Y > 2) - P(X > 2 \text{ and } Y > 2) && (26) \\
&= (\frac{1}{8} + \frac{1}{2}) + (\frac{1}{3} + \frac{1}{3}) - (\frac{1}{8} + \frac{1}{2}) \cdot (\frac{1}{3} + \frac{1}{3}) && (27) \\
&= \frac{5}{8} + \frac{2}{3} - \frac{5}{8} \cdot \frac{2}{3} = \frac{21}{24} = \frac{7}{8} && (28)
\end{aligned}$$

(c) Since $X$ and $Y$ are two independent random variables:

$$\begin{aligned}
P(X > 2 | Y > 2) &= P(X > 2) \\
&= \frac{1}{8} + \frac{1}{2} = \frac{5}{8}
\end{aligned}$$

## Problem 5:

If $X$ is the number of the cars owned by 50 students in the dormitory, then:

$X \sim Binomial(50, \frac{1}{2})$

Thus:

$$\begin{aligned}
P(X > 30) &= \sum_{k=31}^{50} \binom{50}{k} (\frac{1}{2})^k (\frac{1}{2})^{50-k} \\
&= \sum_{k=31}^{50} \binom{50}{k} (\frac{1}{2})^{50} && (31) \\
&= (\frac{1}{2})^{50} \sum_{k=31}^{50} \binom{50}{k} && (32)
\end{aligned}$$

## Problem 6:

a)

$$P(A) = \frac{\binom{20}{4}\binom{30}{6}}{\binom{50}{10}}$$

b)

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(A)}{P(B)} \end{aligned} \tag{106}$$

$$P(B) = \sum_{k=3}^{10} \frac{\binom{20}{k}\binom{30}{10-k}}{\binom{50}{10}} \tag{107}$$

Therefore:

$$P(B|A) = \frac{\binom{20}{4}\binom{30}{6}}{\sum_{k=3}^{10}\binom{20}{k}\binom{30}{10-k}}$$

## **Problem 7:**

a)

$$T = 4 \times 60 = 240 \text{ min}$$
$$\lambda = 240 \times \frac{1}{30} = 8$$

Thus $X \sim Poisson(\lambda = 8)$

$$P(X = 0) = e^{-\lambda} = e^{-8} \tag{108}$$

b)

Let $D$ be the event that a weekday is chosen and let $E$ be the event that a Saturday or Sunday is chosen. Then:

$$\begin{aligned} P(D) &= \frac{5}{7} \\ P(E) &= \frac{2}{7} \end{aligned} \tag{109}$$

## **Problem 8:**

(a)

$$
\begin{aligned}
EX &= \sum_k x_k P_x(x_k) \\
&= 1 \times 0.5 + 2 \times 0.3 + 3 \times 0.2 = 0.5 + 0.6 + 0.6 = 1.7 \qquad (113)
\end{aligned}
$$

(b)

$$
\begin{aligned}
EX^2 &= \sum_k x_k^2 P_x(x_k) \text{ LOTUS} \\
&= (1)^2 \times 0.5 + (2)^2 \times 0.3 + (3)^2 \times 0.2 = 0.5 + 1.2 + 1.8 = 3.5 \qquad (114)
\end{aligned}
$$

Thus,

$$var(X) = EX^2 - (EX)^2 = 3.5 - (1.7)^2 = 0.61$$

$$SD(X) = \sqrt{var(X)} = \sqrt{0.61} = 0.781$$

(c)

$$
\begin{aligned}
E[\frac{2}{X}] &= \sum_k (\frac{2}{x_k}) P_x(x_k) \quad \text{by LOTUS} \\
&= (\frac{2}{1}) \times 0.5 + (\frac{2}{2}) \times 0.3 + (\frac{2}{3}) \times 0.2 \qquad (115) \\
&= 1 + 0.3 + \frac{0.4}{3} = 1 + \frac{3}{10} + \frac{2}{15} = \frac{43}{30} \qquad (116)
\end{aligned}
$$

**Problem 9:**

$$Y = -2X + 3 \qquad (139)$$

$$EY = -2EX + 3 \quad \text{linearity of expectation} \qquad (140)$$

$$1 = -2EX + 3 \quad \rightarrow \qquad EX = 1 \qquad (141)$$

$$
\begin{aligned}
&\mathrm{Var}(Y) = 4 \times \mathrm{Var}(X) = EY^2 - (EY)^2 = 9 - 1 = 8 \qquad (142) \\
&\rightarrow \quad \mathrm{Var}(X) = 2 \qquad (143)
\end{aligned}
$$

## Problem10:

$$f(\alpha) = E(X^2 - 2\alpha X + \alpha^2) \tag{169}$$
$$= EX^2 - 2\alpha EX + \alpha^2 \tag{170}$$

Thus:

$$f(\alpha) = \alpha^2 - 2(EX)\alpha + EX^2 \tag{171}$$
$$\tag{172}$$

$f(\alpha)$ is a polynomial of degree 2 with positive coefficient for $\alpha^2$

$$\frac{\partial f(\alpha)}{\partial \alpha} = 0 \quad \rightarrow \quad 2\alpha - 2EX = 0 \tag{173}$$
$$\rightarrow \quad \alpha = EX \tag{174}$$

## Problem 11:

(a) Since we are counting the number of red balls we draw, we are summing up many Bernoullis (each of which is a single draw – and 1,0 is represented by red or not red). Each Bernoulli has parameter p = 4/9 , the probability of getting a red ball. This means X is a binomial random variable with parameters (n, p) = (10, 4/9).

(b) E[X] = np = 40/9.
Var(X) = np(1 − p) = 40/9 · 5/9 = 200/81.

(c)

$$\begin{aligned} P(\{X \le 3\}) &= P(\{X=0\}) + P(\{X=1\}) + P(\{X=2\}) + P(\{X=3\}) \\ &= \binom{10}{0}\left(\frac{4}{9}\right)^0\left(\frac{5}{9}\right)^{10} + \binom{10}{1}\left(\frac{4}{9}\right)^1\left(\frac{5}{9}\right)^9 \\ &+ \binom{10}{2}\left(\frac{4}{9}\right)^2\left(\frac{5}{9}\right)^8 + \binom{10}{3}\left(\frac{4}{9}\right)^3\left(\frac{5}{9}\right)^7 \\ &\sim .278 \end{aligned}$$

## Problem 12:

The mass of $X$ is $p_X(0) = \frac{\binom{4}{0}\binom{4}{4}}{\binom{8}{4}} = 1/70$, $p_X(1) = \frac{\binom{4}{1}\binom{4}{3}}{\binom{8}{4}} = 8/35$, $p_X(2) = \frac{\binom{4}{2}\binom{4}{2}}{\binom{8}{4}} = 18/35$, $p_X(3) = \frac{\binom{4}{3}\binom{4}{1}}{\binom{8}{4}} = 8/35$, $p_X(4) = \frac{\binom{4}{4}\binom{4}{0}}{\binom{8}{4}} = 1/70$. So the CDF of $X$, as $x = 0, 1, 2, 3, 4$, is $F_X(0) = 1/70$, $F_X(1) = 17/70$, $F_X(2) = 53/70$, $F_X(3) = 69/70$, $F_X(4) = 1$.

**Problem 13:**

**a.** We have $P(3 \leq X \leq 5) = (\frac{2}{7})(\frac{5}{7})^2 + (\frac{2}{7})(\frac{5}{7})^3 + (\frac{2}{7})(\frac{5}{7})^4 = 5450/16807 = 0.3243$.

**b.** We have $P(a \leq X \leq b) = \sum_{x=a}^{b}(\frac{2}{7})(\frac{5}{7})^{x-1} = (\frac{2}{7})(\frac{5}{7})^{a-1}\sum_{x=0}^{b-a}(\frac{5}{7})^x = (\frac{2}{7})(\frac{5}{7})^{a-1}\frac{(1-(\frac{5}{7})^{b-a+1})}{(1-\frac{5}{7})} = (\frac{5}{7})^{a-1}(1-(\frac{5}{7})^{b-a+1})$.

**Problem 14:**

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 15:**

Let's first make sure we understand what Var$(2X-Y)$ and Var$(X+2Y)$ mean. They are Var$(Z)$ and Var$(W)$, where the random variables $Z$ and $W$ are defined as $Z=2X-Y$ and $W=X+2Y$. Since $X$ and $Y$ are independent random variables, then $2X$ and $-Y$ are independent random variables. Also, $X$ and $2Y$ are independent random variables. Thus, we can write

$$\mathrm{Var}(2X-Y) = \mathrm{Var}(2X)+\mathrm{Var}(\text{-Y}) = 4\mathrm{Var}(X)+\mathrm{Var}(\mathrm{Y}) = 6,$$
$$\mathrm{Var}(X+2Y) = \mathrm{Var}(X)+\mathrm{Var}(2\mathrm{Y}) = \mathrm{Var}(X)+4\mathrm{Var}(\mathrm{Y}) = 9.$$

By solving for Var$(X)$ and Var$(Y)$, we obtain Var$(X)=1$ and Var$(Y)=2$.

**Problem 16:**

$$R_X = \{1, 2, 3, \ldots\} \tag{117}$$

$$P_X(k) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-1}, \qquad \text{for } k = 1, 2, 3, \ldots \tag{118}$$

$$R_Y = \{|X-5| \big| X \in R_X\} = 0, 1, 2, ... \tag{119}$$

Thus,

$$\begin{aligned} P_Y(0) &= P(Y=0) = P(|X-5|=0) = P(X=5) \\ &= (\frac{2}{3})^4(\frac{1}{3}) \end{aligned} \tag{120}$$

For $k = 1, 2, 3, 4$

$$\begin{aligned} P_Y(k) &= P(Y=k) = P(|X-5|=k) = P(X=5+k \quad \text{or} \quad X=5-k) \\ &= P_X(5+k) + P_X(5-k) = [(\frac{2}{3})^{4+k} + (\frac{2}{3})^{4-k}](\frac{1}{3}) \end{aligned} \tag{121}$$

For $k \geq 5$,

$$\begin{aligned} P_Y(k) &= P(Y=k) = P(|X-5|=k) = P(X=5+k) \\ &= P_X(5+k) = (\frac{2}{3})^{4+k}(\frac{1}{3}) \end{aligned} \tag{122}$$

So, in summary:

$$P_Y(k) = \begin{cases} (\frac{2}{3})^{k+4}(\frac{1}{3}) & \text{for } k = 0, 5, 6, 7, 8, ... \\ ((\frac{2}{3})^{k+4} + (\frac{2}{3})^{4-k})(\frac{1}{3}) & \text{for } k = 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

**Problem 17:**

$$R_X = \{-10, -9, ..., 9, 10\} \tag{123}$$

$Y = g(X)$ Thus:

$$R_Y = \{0, 1, 2, 3, 4, 5\} \tag{124}$$

Thus,

$$P_Y(0) = P(X \leq 0) = \sum_{k=-10}^{0} P_X(k) = \frac{11}{21}$$

For $k = 1, 2, 3, 4$

$$\begin{aligned}
P_Y(1) &= P(Y = 1) = P(X = 1) = \frac{1}{21} \\
P_Y(2) &= P_Y(3) = P_Y(4) = \frac{1}{21} \\
P_Y(5) &= P(X \geq 5) \\
&= P_X(5) + P_X(6) + P_X(7) + P_X(8) + P_X(9) + P_X(10) &(125) \\
&= \frac{6}{21} &(126) \\
& &(127)
\end{aligned}$$

Thus:

$$P_Y(k) = \begin{cases} \frac{11}{21} & \text{for } k = 0 \\ \frac{1}{21} & \text{for } k = 1, 2, 3, 4 \\ \frac{6}{21} & \text{for } k = 5 \\ 0 & \text{otherwise} \end{cases}$$

**Problem 18:**

(a) $m = 2$, since

$$P(X \geq 2) = 0.6 \text{ and } P(X \leq 2) = 0.7 \tag{183}$$

(b)

$$\begin{aligned}
P_X(k) &= \frac{1}{6} \text{ for } k = 1, 2, 3, 4, 5, 6 &(184) \\
&\rightarrow 3 \leq m \leq 4 &(185)
\end{aligned}$$

Thus, we conclude $3 \leq m \leq 4$. Any value $\in [3, 4]$ is a median for $X$.

**Problem19:**

(a) In case (1), we have

$$p_X(1) = \mathbf{P}(K_1) = \frac{1}{5},$$
$$p_X(2) = \mathbf{P}(K_1^c)\mathbf{P}(K_2 \mid K_1^c) = \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{5},$$
$$p_X(3) = \mathbf{P}(K_1^c)\mathbf{P}(K_2^c \mid K_1^c)\mathbf{P}(K_3 \mid K_1^c \cap K_2^c) = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{5}.$$

Proceeding similarly, we see that the PMF of $X$ is

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

(b)

We can also view the problem as ordering the keys in advance and then trying them in succession, in which case the probability of any of the five keys being correct is $1/5$.

In case (2), $X$ is a geometric random variable with $p = 1/5$, and its PMF is

$$p_X(k) = \frac{1}{5} \cdot \left(\frac{4}{5}\right)^{k-1}, \qquad k \geq 1.$$

**Problem 20:**

$$\mathbf{E}[Y] = 32 + 9\mathbf{E}[X]/5 = 32 + 18 = 50.$$

Also

$$\text{var}(Y) = (9/5)^2 \text{var}(X),$$

where $\text{var}(X)$, the square of the given standard deviation of $X$, is equal to 100. Thus, the standard deviation of $Y$ is $(9/5) \cdot 10 = 18$. Hence a normal day in Fahrenheit is one for which the temperature is in the range $[32, 68]$.

## Problem 21:

(a) Let $X$ be the number of tosses until the game is over. Noting that $X$ is geometric with probability of success

$$\mathbf{P}\big(\{HT,TH\}\big) = p(1-q) + q(1-p),$$

we obtain

$$p_X(k) = \big(1 - p(1-q) - q(1-p)\big)^{k-1}\big(p(1-q)+q(1-p)\big), \qquad k = 1, 2, \ldots$$

Therefore

$$\mathbf{E}[X] = \frac{1}{p(1-q)+q(1-p)}$$

and

$$\mathrm{var}(X) = \frac{pq + (1-p)(1-q)}{\big(p(1-q)+q(1-p)\big)^2}.$$

(b) The probability that the last toss of the first coin is a head is

$$\mathbf{P}\big(HT \,|\, \{HT,TH\}\big) = \frac{p(1-q)}{p(1-q)+(1-q)p}.$$

## Problem 22:

**a.** We compute the following: $p_Y(0) = \binom{2}{0}\binom{6}{2}/\binom{8}{2} = 15/28$, $p_Y(1) = \binom{2}{1}\binom{6}{1}/\binom{8}{2} = 3/7$, and $p_Y(2) = \binom{2}{2}\binom{6}{0}/\binom{8}{2} = 1/28$.

**b.** We have $\mathbb{E}(Y) = (0)(15/28) + (1)(3/7) + (2)(1/28) = 1/2$.

**c.** Yes, the probability mass functions of $X$ and $Y$ are the same, so their expected values are the same too.

## Problem 23:

Let p be the probability of success. In this case, p=1/100.

$$P(\text{success at the } k^{th} \text{ trial}) = p\,(1-p)^{k-1}$$

$$P_X(k) = (0.01)(0.99)^{k-1},\ k=1,2,\ldots..$$

Therefore,

$$P(\text{at least 19 failures}) = P_X(k >= 20)$$

Thus, we can write

$$\sum_{k=20}^{\infty} p(1-p)^{k-1} = \frac{p}{1-p} \times \frac{(1-p)^{20}}{1-(1-p)} = (1-p)^{19} = (0.99)^{19} = 0.8262$$

## Problem 24:

Let $X$ denote the number of raindrops that fall during the next one minute. Then $X$ is Poisson with $\lambda = 6$.

**a.** We have $P(X=5) = e^{-6}6^5/5! = 0.1606$.

**b.** We have $P(X=0) = e^{-6}6^0/0! = 0.002479$.

**c.** We have $P(X \geq 4) = 1 - P(X \leq 3) = 1 - e^{-6}6^0/0! - e^{-6}6^1/1! - e^{-6}6^2/2! - e^{-6}6^3/3! = 1 - 61e^{-6} = 0.8488$.

## Problem 25:

**a.** We let $X$ denote the number of errors in such a book chapter. Then we compute $P(X \leq 4) = \sum_{x=0}^{4} P(X = x) = \sum_{x=0}^{4} e^{-2.8} 2.8^x / x! = 0.8477$.

**b.** We compute $P(X \leq 3 \mid P \leq 5) = \frac{P(X \leq 3 \ \& \ X \leq 5)}{P(X \leq 5)} = \frac{P(X \leq 3)}{P(X \leq 5)} = \frac{\sum_{x=0}^{3} e^{-2.8} 2.8^x / x!}{\sum_{x=0}^{5} e^{-2.8} 2.8^x / x!} = 0.6919/0.9349 = 0.74$.

**Problem 26:**

```
import random
import numpy as np

n_iter = 10000
x = np.zeros(n_iter)
p=0.25
for i in range(n_iter):
   k = 0
   while x[i] == 0:
      k = k + 1
      if random.random() < p:
         x[i] = k

count = np.sum(x>3)
probability_estimate = count/n_iter
print(probability_estimate)
```

**Problem 27:**

```
import random
import numpy as np
import matplotlib.pyplot as plt

n_iter = 10000
x = np.zeros(n_iter)
p=0.25
for i in range(n_iter):
   k = 0
   while x[i] == 0:
      k = k + 1
      if random.random() < p:
         x[i] = k

probability_estimate=[]
for i in range(1,21):
   probability_estimate.append(np.sum(x-i==0)/n_iter)
```

```
true_probability = np.zeros(20)
for k in range(1,21):
   true_probability[k-1] = (1-p)**(k-1)*p

plt.stem(range(1,21),probability_estimate, basefmt=" ", label = 'estimate')
plt.stem(range(1,21),true_probability, 'r', basefmt=" ", markerfmt='ro', label = 'true')
plt.axis([0.5,20.5,0,0.3]) # adjust the axis limits
plt.legend()
plt.xlabel('k')
plt.ylabel('P(k)')
plt.title('Geometric PMF')
plt.show()
```

**Problem 28:**

```
from scipy import stats
import numpy as np

xi = np.array([-1,1])
pX = np.array([1/4,3/4])
custm = stats.rv_discrete(name='custm', values=(xi, pX))
x = custm.rvs(size=10000)

print('estimated mean of X', np.mean(x))
print('estimated variance of X', np.var(x))
```

**Problem 29:**

```
from scipy.stats import binom

#(a)
binom.pmf(3, 10, 0.1)

#(b)
sum([binom.pmf(x, 10, 0.1) for x in range(3)])
```