UNIT 6

BIVARIATE RANDOM VARIABLE

Structure

6.1 INTRODUCTION

In Units 4 and 5, you studied discrete and continuous univariate random variables and their different distributions like CDF, PMF, PDF, etc. In the present unit, we will discuss bivariate discrete and continuous random variables and their corresponding distributions. First of all, what we mean by a bivariate random variable is discussed in Sec. 6.2. Like univariate case CDF is also defined in the same way in bivariate discrete and continuous worlds of probability theory and is defined in Sec. 6.3. The next two sections, i.e., Sec. 6.4 and Sec. 6.5 discuss discrete and continuous bivariate random variables respectively.

What we have discussed in this unit is summarised in Sec. 6.6. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit some more questions based on the entire unit are given in Sec. 6.7 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1 of this course, solutions of all the SAQs and Terminal Questions are given in Sec. 6.8.

In the next unit, you will study what is expected value of a distribution and moment generating function (MGF) of the random variable.

Expected Learning Outcomes

After completing this unit, you should be able to:

- explain the concept of discrete and continuous bivariate random variables;
- define joint, marginal and conditions probability mass/density functions and CDF in bivariate random variables setting; and
- able to apply the concept of PMF, PDF and CDF to obtain required probabilities in bivariate setting.

6.2 BIVARIATE RANDOM VARIABLE

In Unit 2, you have studied what is a probability space and in Unit 4, you have studied random variable. Recall those concepts because in this unit, we will use many concepts related to probability measure and random variable.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. Let $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ be two random variables. Recall the definition of univariate random variable which associate a unique real number to each members of Ω. Note that both X and Y are defined on the same sample space Ω, so the ordered pair $(X(\omega), Y(\omega))$, of real numbers for each $\omega \in \Omega$ lies in $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. So, the ordered pair random variable (X, Y) which associates a unique ordered pair of real numbers to each member of the sample space is called **bivariate random variable** on Ω, i.e., (X, Y) is a function which can be denoted as follows.

$(X(.), Y(.)) : \Omega \to \mathbb{R}^2$

For example, consider the random experiment of tossing a coin twice. Sample space of this random experiment is $\Omega = \{HH, HT, TH, TT\}$. Let (X, Y) be the bivariate random variable, where X denotes the number of heads in two tosses and Y denotes the number of tails before the first head. So, the bivariate random variable (X, Y) associates ordered pair of real numbers (2, 0) with HH, (1, 0) with HT, (1, 1) with TH and (0, 2) with TT and is visualised in Fig. 6.1 as follows.

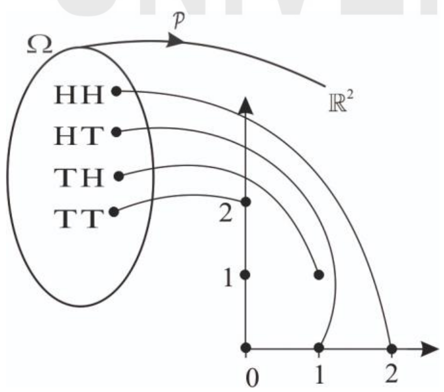


Fig. 6.1: Visualisation of the bivariate random variable (X, Y) as a function from sample space Ω to xy-plane $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$

In Unit 4, you have studied what is a discrete random variable and in Unit 5, you have studied what is a continuous random variable. If both X and Y are discrete random variables then the ordered pair random variable (X, Y) is called a **discrete bivariate random variable**. However, if both X and Y are continuous random variables then it is not necessary that (X, Y) will be joint continuous bivariate random variable. Why it is so its discussion is beyond the scope of the course. The technical word is jointly continuous which is used when we want that bivariate random variable (X, Y) is continuous. So, if both X and Y are **jointly** continuous random variables then the ordered pair random variable (X, Y) is called a **continuous bivariate random variable**. Like univariate cases to study bivariate discrete and continuous random variables, we will need the cumulative distribution function of the random variable (X, Y). So, let us first define the cumulative distribution function (CDF) of the random variable (X, Y) in the next section. … (6.1)

6.3 CDF FOR BIVARIATE RANDOM VARIABLE

Recall that to define CDF of a univariate random variable, we used the Borel measurable sets of the form $(-\infty, x], x \in \mathbb{R}$ which generated Borel σ-field $\mathcal{B}(\mathbb{R})$. If you want you may refer to (3.69). In univariate cases, random variables attain real values so, we needed to define Borel σ-field $\mathcal{B}(\mathbb{R})$. But in bivariate cases, random variable (X, Y) attains ordered pair (x, y) of real values. So, we need to define Borel σ-field $\mathcal{B}(\mathbb{R}^2)$. Like univariate case the collection $\mathcal{C}_4 = \{(-\infty, x] \times (-\infty, y] : x, y \in \mathbb{R}\}$ generates Borel σ-field $\mathcal{B}(\mathbb{R}^2)$. That is $\sigma(\mathcal{C}_4) = \mathcal{B}(\mathbb{R}^2)$. … (6.2)

Like the univariate case, the next concept that is required to define before defining CDF of a bivariate random variable (X, Y) is the probability measure $\mathcal{P}_{X, Y}$ induced by the bivariate random variable (X, Y). Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and $X : \Omega \to \mathbb{R}, Y : \Omega \to \mathbb{R}$ be two random variables then the probability measure $\mathcal{P}_{X, Y}$ induced by the bivariate random variable (X, Y) is defined by (if required you may refer 4.16, 4.27, 4.28, 4.31 and 4.32)

$$\mathcal{P}_{X, Y}(B) = \mathcal{P}\left(\left\{\omega \in \Omega : \left(X(\omega), Y(\omega)\right) \in B\right\}\right), \quad B \in \mathcal{B}(\mathbb{R}^2) \qquad \dots (6.3)$$

or $\mathcal{P}_{X, Y}(B) = \mathcal{P}((X, Y) \in B), \quad B \in \mathcal{B}(\mathbb{R}^2)$ … (6.4)

Following the similar steps as we did in Sec. 4.4 of Unit 4 for the univariate case, we can prove that $\mathcal{P}_{X, Y}$ is a probability measure on $\mathcal{B}(\mathbb{R}^2)$ and so $\left(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathcal{P}_{X, Y}\right)$ is a probability space and known as **induced probability space by the random variable (X, Y)** on the xy-plane $(\mathbb{R}^2)$.

It implies that the probability measure $\mathcal{P}_{X, Y}$ assigns probability to every member of $\mathcal{B}(\mathbb{R}^2)$. But we know that (refer 6.2)

$\sigma(\mathcal{C}_4) = \mathcal{B}(\mathbb{R}^2)$, where $\mathcal{C}_4 = \{(-\infty, x] \times (-\infty, y] : x, y \in \mathbb{R}\}$

So, all the rectangles of the form $(-\infty, x] \times (-\infty, y], x, y \in \mathbb{R}$ are members of the Borel σ-field $\mathcal{B}(\mathbb{R}^2)$. So, by definition of the probability measure $\mathcal{P}_{X, Y}$, you may refer (6.3), we have

$$\mathcal{P}_{X, Y}\left((-\infty, x] \times (-\infty, y]\right) = \mathcal{P}\left(\left\{\omega \in \Omega : \left(X(\omega), Y(\omega)\right) \in (-\infty, x] \times (-\infty, y]\right\}\right) \dots (6.5)$$

$$(-\infty, x] \times (-\infty, y] \in \mathcal{B}(\mathbb{R}^2), \quad x, y \in \mathbb{R}$$

$$\Rightarrow \mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right) = \mathcal{P}\left(\{\omega\in\Omega:\ X(\omega)\le x,\ \ Y(\omega)\le y\}\right) \qquad \text{... (6.6)}$$

$$(-\infty,\ x]\times(-\infty,\ y]\in\mathcal{B}\left(\mathbb{R}^2\right),\ \ x,\ y\in\mathbb{R}$$

$$\left[\because\ a\in(-\infty,\ x]\Rightarrow a\le x\ \text{ and }\ b\in(-\infty,\ y]\Rightarrow b\le y\right]$$

Like univariate case and due to the frequent use in obtaining probability of the random variable (X, Y) lying in different intervals, we use the following simple notation for (6.5) or (6.6).

$$\mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right) = \mathcal{P}\left(X\le x,\ \ Y\le y\right) \qquad \text{... (6.7)}$$

$$(-\infty,\ x]\times(-\infty,\ y]\in\mathcal{B}\left(\mathbb{R}^2\right),\ \ x,\ y\in\mathbb{R}$$

Now, we can define CDF for bivariate random variable (X, Y) as follows.

CDF for Bivariate Random Variable (X, Y): Let $\left(\Omega,\ \mathcal{F},\ \mathcal{P}\right)$ be a probability space and $\left(\mathbb{R}^2,\ \mathcal{B}\left(\mathbb{R}^2\right),\ \mathcal{P}_{X,\,Y}\right)$ be the induced probability space induced by the random variable (X, Y). Then a function $F_{\mathcal{P}_{X,\,Y}}:\mathbb{R}^2\to[0,\ 1]$ defined by

$$\left.\begin{aligned}&F_{\mathcal{P}_{X,\,Y}}(x,\ y) = \mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right),\ \ (x,\ y)\in\mathbb{R}^2\ \text{ where}\\&\mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right) = \mathcal{P}\left(\{\omega\in\Omega:\ X(\omega)\le x,\ \ Y(\omega)\le y\}\right)\ \ \forall\ \ (x,\ y)\in\mathbb{R}^2\end{aligned}\right\}\ (6.8)$$

is called cumulative distribution function (CDF) or simply distribution function of the bivariate random variable (X, Y) corresponding to the induced probability measure $\mathcal{P}_{X,\,Y}$.

Using simplified notation used in (6.7), we can write (6.8) as follows

$$F_{\mathcal{P}_{X,\,Y}}(x,\ y) = \mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right) = \mathcal{P}\left(X\le x,\ \ Y\le y\right),\ \ (x,\ y)\in\mathbb{R}^2 \qquad \text{… (6.9)}$$

We will use CDF very frequent so, we will denote it simply by $F_{X,\,Y}$ instead of $F_{\mathcal{P}_{X,\,Y}}$. So, using this notation (6.9) can be written as

$$F_{X,\,Y}(x,\ y) = \mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right) = \mathcal{P}\left(X\le x,\ \ Y\le y\right),\ \ (x,\ y)\in\mathbb{R}^2 \qquad \text{… (6.10)}$$

In view of (6.6) it can also be written as

$$F_{X,\,Y}(x,\ y) = \mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right) = \mathcal{P}\left(\{\omega\in\Omega:\ X(\omega)\le x,\ \ Y(\omega)\le y\}\right),(x,\ y)\in\mathbb{R}^2 \qquad \text{… (6.11)}$$

Combining (6.8), (6.9), (6.10) and (6.11), we have

$$\left.\begin{aligned}F_{\mathcal{P}_{X,\,Y}}(x,\ y) = F_{X,\,Y}(x,\ y) &= \mathcal{P}_{X,\,Y}\left((-\infty,\ x]\times(-\infty,\ y]\right)\\&= \mathcal{P}\left(\{\omega\in\Omega:\ X(\omega)\le x,\ \ Y(\omega)\le y\}\right)\\&= \mathcal{P}\left(X\le x,\ \ Y\le y\right),\ \ (x,\ y)\in\mathbb{R}^2\end{aligned}\right\} \qquad \text{… (6.12)}$$

Also, we say that $F_{X,\,Y}$ is the distribution function of the bivariate random variable (X, Y) instead of saying that corresponding to the induced probability measure $\mathcal{P}_{X,\,Y}$. … (6.13)

Like univariate case CDF of bivariate random variable (X, Y) also has similar properties listed as follows.

(a) $F_{X,\,Y}(-\infty,\ -\infty) = \lim\limits_{\substack{x\to-\infty\\y\to-\infty}} F_{X,\,Y}\left(x,\ y\right) = 0.$ … (6.14)

Also, $F_{X, Y}(-\infty, y) = \lim_{x \to -\infty} F_{X, Y}(x, y) = 0$ … (6.15)

and $F_{X, Y}(x, -\infty) = \lim_{y \to -\infty} F_{X, Y}(x, y) = 0$ … (6.16)

(b) $F_{X, Y}(\infty, \infty) = \lim_{\substack{x \to \infty \\ y \to \infty}} F_{X, Y}(x, y) = 1.$ … (6.17)

But $F_{X, Y}(x, \infty)$ gives marginal CDF of the random variable X, i.e.,

$\lim_{y \to \infty} F_{X, Y}(x, y) = F_{X, Y}(x, \infty) = F_X(x).$ … (6.18)

Similarly, $F_{X, Y}(\infty, y)$ gives marginal CDF of the random variable Y, i.e.,

$\lim_{x \to \infty} F_{X, Y}(x, y) = F_{X, Y}(\infty, y) = F_Y(y).$ … (6.19)

Marginal CDF's are just CDF's of univariate random variables which you have studied in Unit 4 and 5. Why the name marginal is added before CDF to get answer of this question refer to (6.49).

(c) $F_{X, Y}$ is increasing or non-decreasing. That is if $x_1, x_2, y_1, y_2 \in \mathbb{R}$ be such that $x_1 \le x_2$ and $y_1 \le y_2$, then $F_{X, Y}(x_1, y_1) \le F_{X, Y}(x_2, y_2)$ … (6.20)

(d) $F_{X, Y}$ is continuous from above, i.e.,

$\lim_{\substack{h \downarrow 0 \\ k \downarrow 0}} F_{X, Y}(x + h, y + k) = F_{X, Y}(x, y) \quad \forall \; x, y \in \mathbb{R}$ … (6.21)

(e) If $x_1, x_2, y_1, y_2 \in \mathbb{R}$ be such that $x_1 \le x_2$ and $y_1 \le y_2$, then (to get more detail of this property refer 6.63 and Fig. 6.6)

$$\mathcal{P}(x_1 < X \le x_2,\ y_1 < Y \le y_2) = F_{X, Y}(x_2, y_2) - F_{X, Y}(x_1, y_2) - F_{X, Y}(x_2, y_1) + F_{X, Y}(x_1, y_1) \quad \dots (6.22)$$

Proofs of these properties are similar to univariate case, so, we are not discussing them.

6.4 BIVARIATE DISCRETE RANDOM VARIABLE

In Sec. 4.6 of Unit 4, we have defined what is a discrete random variable. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. Here instead of one, we will deal with two discrete random variables X and Y simultaneously denoted by (X, Y) and known as discrete bivariate random variable. From the discussion of Sec. 4.6, you know that a discrete random variable either attains finite number of values or at the most countably infinite number of values. Here both X and Y are discrete it means:

- either both X and Y may attain finite number of values or … (6.23)
- both X and Y may attain countably infinite number of values or … (6.24)
- one of them attain finite values and another may attain countably infinite number of values. … (6.25)

Let us consider one example. Suppose in a bag there are 2 red, 4 blue and 5 black balls. Three balls are drawn from this bag randomly. Let X, Y denote the number of red and blue balls, respectively out of the three drawn balls. So, X can attain values 0, 1 and 2 while Y can attain values 0, 1, 2 and 3. The possible values of the random variable X are shown in the first column of Table 6.1 and that of Y are shown in the second row of the same table which is given as follows.

Table 6.1: Values of X and Y together with 12 empty cells where generally probabilities of corresponding values of the bivariate random variable (X, Y) are written

	Values of Y			
Values of X	**0**	**1**	**2**	**3**
0				
1				
2				

In this table, we have 12 empty cells. We will discuss about the entries of these 12 empty cells after defining joint probability mass function of bivariate random variable (X, Y).

In this example, both variables X and Y attain finite number of values. So, the bivariate random variable (X, Y) is a discrete bivariate random variable. Possible values of bivariate random variable (X, Y) in this example are (0, 0), (0, 1), (0, 2), (0, 3), (1, 0), (1, 1), (1, 2), (1, 3), (2, 0), (2, 1), (2, 2), (2, 3). So, the bivariate random variable (X, Y) attains 12 values. These 12 possible values are shown in xy-plane in Fig. 6.2 (a) by solid dots. If we write these 12 values in a set $S_{X,\,Y}$ (say), then $S_{X,\,Y}$ is known as **support of the bivariate random variable** (X, Y), provided if all these 12 points are assigned non zero probability by probability law. But if one or more than one point among these 12 points have assigned 0 probability by the probability law then that point or points will go out from the support of (X, Y). So, support of the bivariate random variable (X, Y) can have at the most 12 points given as follows (to know general definition of support of (X, Y) refer 6.31) … (6.26)

$$S_{X,\,Y} = \begin{Bmatrix} (0,\,0),\,(0,\,1),\,(0,\,2),\,(0,\,3),\,(1,\,0),\,(1,\,1), \\ (1,\,2),\,(1,\,3),\,(2,\,0),\,(2,\,1),\,(2,\,2),\,(2,\,3) \end{Bmatrix} \qquad \ldots (6.27)$$

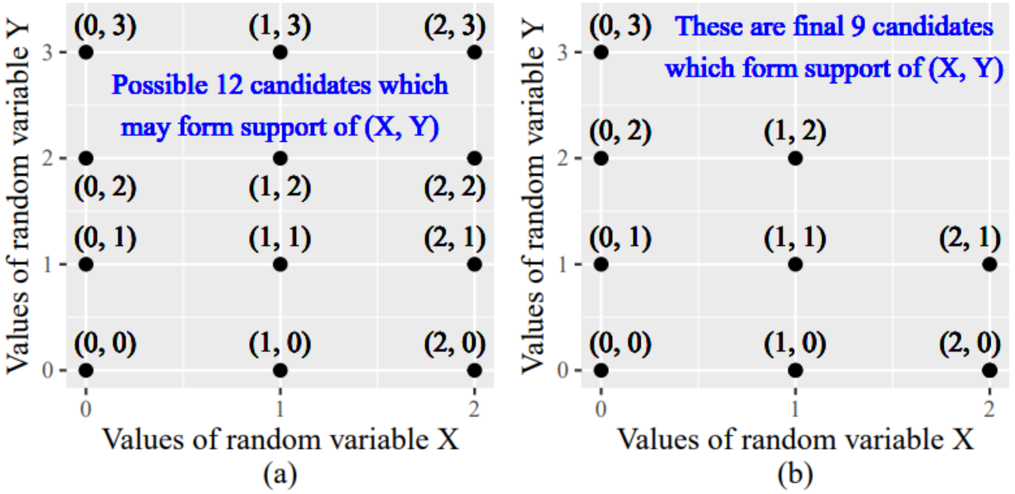


Fig. 6.2: Visualisation of (a) 12 points which are 12 values of the bivariate random variable (X, Y) that are possible candidates to form the support of (X, Y) (b) 9 points which finally form the support of the bivariate random variable (X, Y)

Why we are calling the bivariate variable (X, Y) is a random variable? We are calling it random because it is constituted by random variables X and Y where both X and Y are defined on a sample space of a random experiment. Then the immediate next question that will be arising in your mind is if (X, Y) is a random variable then probabilities should be associated with all the 12 possible values of (X, Y). You are thinking brilliantly and in the right direction. To answer your brilliant question, first, we have to define what is called joint probability mass function.

In univariate case in Unit 4, you studied PMF and CDF of a random variable but in the case of bivariate random variable there are some other concepts which we need to discuss. So, what we need to discuss in this section are listed as follows.

- Joint, Marginal and Conditional Probability Mass Functions
- Joint and Marginal Cumulative Distribution Functions
- Independence of Random Variables

Let us discuss these in three subsections of this section as follows.

6.4.1 Joint, Marginal and Conditional Probability Mass Functions

Joint Probability Mass Function of a Discrete Bivariate Random Variable

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and X, Y be two discrete random variables on Ω, then joint probability mass function (JPMF) of the bivariate random variable (X, Y) is a function which associates unique probability to each value (x, y) of (X, Y) and 0 probability to the values which are not in the support of (X, Y). It is denoted by $p_{X,\,Y}$ and is defined as follows

$$p_{X,\,Y}(x,\,y) = \begin{cases} \mathcal{P}(X = x,\; Y = y), & \text{if } (x,\,y) \text{ is in support of } (X,\,Y) \\ 0, & \text{otherwise} \end{cases} \qquad \ldots (6.28)$$

where $\mathcal{P}(X = x,\; Y = y) = \mathcal{P}(\{\omega \in \Omega : X(\omega) = x,\; Y(\omega) = y\}) \;\; \forall \;\; (x,\,y) \in \mathbb{R}^2$ ... (6.29)

For a function $p_{X,\,Y}$ as defined by (6.28) to become a **valid joint probability mass function** it has to satisfy the following two conditions.

(i) $p_{X,\,Y}(x,\,y) \geq 0 \;\; \forall \; (x,\,y) \in \text{support of } (X,\,Y)$

(ii) $\sum_{(x,\,y) \in S_{X,\,Y}} p_{X,\,Y}(x,\,y) = 1$ where $S_{X,\,Y}$ denotes support of (X, Y). … (6.30)

In the definition of joint PMF of the bivariate random variable (X, Y), we have used the technical word support. We have explained its meaning in (6.26). Now, as promised there let us define it in general as follows.

Support of Bivariate Random Variable (X, Y) is the collection of those points in $\mathbb{R}^2$ which have assigned non zero probability by the probability law, i.e.,

Support of (X, Y) is the set $S_{X,\,Y} = \{(x,\,y) \in \mathbb{R}^2 : p_{X,\,Y}(x,\,y) > 0\}$. … (6.31)

Let us consider the example of 2 red, 4 blue and 5 black balls considered in the beginning of this section. To decide how many of the 12 points will finally qualify to be members of the support of (X, Y), we have to obtain their probabilities using (6.28) and (6.29). Let us evaluate the probabilities for each of the 12 points belong to $S_{X,\,Y}$ given by (6.27) as follows.

$$p_{X,\,Y}(0,\,0) = \mathcal{P}(X = 0,\ Y = 0)$$

$$= \mathcal{P}(\text{getting 0 red ball and 0 blue ball}) \begin{bmatrix}\text{Meaning of 'comma' in} \\ \mathcal{P}(X = 0,\ Y = 0)\text{ is 'and'}\end{bmatrix}$$

$$= \mathcal{P}(\text{All 3 balls are black})$$

$$= \frac{\binom{5}{3}}{\binom{11}{3}} = \frac{\frac{5\times4}{\lfloor 2}}{\frac{11\times10\times9}{\lfloor 3}} = \frac{5\times4}{2}\times\frac{6}{11\times10\times9} = \frac{5\times4}{2}\times\frac{1}{165} = \frac{10}{165} \quad \text{... (6.32)}$$

Similarly, we can obtain other probabilities as follows.

$$p_{X,\,Y}(0,\,1) = \mathcal{P}(X = 0,\ Y = 1) = \mathcal{P}(\text{getting 0 red ball and 1 blue ball})$$

$$= \frac{\binom{4}{1}\binom{5}{2}}{\binom{11}{3}} = 4\times\frac{5\times4}{2}\times\frac{1}{165} = \frac{40}{165} \quad \text{... (6.33)}$$

$$p_{X,\,Y}(0,\,2) = \mathcal{P}(X = 0,\ Y = 2) = \mathcal{P}(\text{getting 0 red ball and 2 blue balls})$$

$$= \frac{\binom{4}{2}\binom{5}{1}}{\binom{11}{3}} = \frac{4\times3}{2}\times5\times\frac{1}{165} = \frac{30}{165} \quad \text{... (6.34)}$$

$$p_{X,\,Y}(0,\,3) = \mathcal{P}(X = 0,\ Y = 3) = \mathcal{P}(\text{getting 0 red ball and 3 blue balls})$$

$$= \frac{\binom{4}{3}}{\binom{11}{3}} = 4\times\frac{1}{165} = \frac{4}{165} \quad \text{... (6.35)}$$

$$p_{X,\,Y}(1,\,0) = \mathcal{P}(X = 1,\ Y = 0) = \mathcal{P}(\text{getting 1 red ball and 0 blue ball})$$

$$= \frac{\binom{2}{1}\binom{5}{2}}{\binom{11}{3}} = 2\times\frac{5\times4}{2}\times\frac{1}{165} = \frac{20}{165} \quad \text{... (6.36)}$$

$$p_{X,\,Y}(1,\,1) = \mathcal{P}(X = 1,\ Y = 1) = \mathcal{P}(\text{getting 1 red ball and 1 blue ball})$$

$$= \frac{\binom{2}{1}\binom{4}{1}\binom{5}{1}}{\binom{11}{3}} = 2\times4\times5\times\frac{1}{165} = \frac{40}{165} \quad \text{... (6.37)}$$

$$p_{X,Y}(1, 2) = \mathcal{P}(X = 1, Y = 2) = \mathcal{P}(\text{getting 1 red ball and 2 blue balls})$$

$$= \frac{\binom{2}{1}\binom{4}{2}}{\binom{11}{3}} = 2 \times \frac{4 \times 3}{2} \times \frac{1}{165} = \frac{12}{165} \quad \ldots (6.38)$$

$$p_{X,Y}(1, 3) = \mathcal{P}(X = 1, Y = 3) = \mathcal{P}(\text{getting 1 red ball and 3 blue balls})$$

$$= 0 \begin{bmatrix} \because \text{Only 3 balls are drawn so we} \\ \text{cannot get 1 red and 3 blue balls} \end{bmatrix} \quad \ldots (6.39)$$

$$p_{X,Y}(2, 0) = \mathcal{P}(X = 2, Y = 0) = \mathcal{P}(\text{getting 2 red balls and 0 blue ball})$$

$$= \frac{\binom{2}{2}\binom{5}{1}}{\binom{11}{3}} = 1 \times 5 \times \frac{1}{165} = \frac{5}{165} \quad \ldots (6.40)$$

$$p_{X,Y}(2, 1) = \mathcal{P}(X = 2, Y = 1) = \mathcal{P}(\text{getting 2 red balls and 1 blue ball})$$

$$= \frac{\binom{2}{2}\binom{4}{1}}{\binom{11}{3}} = 1 \times 4 \times \frac{1}{165} = \frac{4}{165} \quad \ldots (6.41)$$

$$p_{X,Y}(2, 2) = \mathcal{P}(X = 2, Y = 2) = \mathcal{P}(\text{getting 2 red and 2 blue balls})$$

$$= 0 \begin{bmatrix} \because \text{Only 3 balls are drawn so we} \\ \text{cannot get 2 red and 2 blue balls} \end{bmatrix} \quad \ldots (6.42)$$

$$p_{X,Y}(2, 3) = \mathcal{P}(X = 2, Y = 3) = \mathcal{P}(\text{getting 2 red ball and 3 blue balls})$$

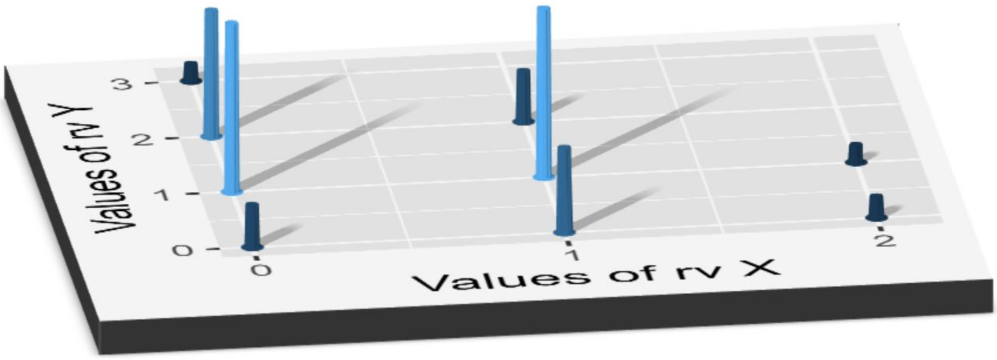
$$= 0 \begin{bmatrix} \because \text{Only 3 balls are drawn so we} \\ \text{cannot get 2 red and 3 blue balls} \end{bmatrix} \quad \ldots (6.43)$$

We have obtained probabilities of all 12 points which were lying in $S_{X,Y}$ given by (6.27). Now, we can complete Table 6.1 by putting these calculated probabilities in its 12 empty cells. Let us call this new table as Table 6.2.

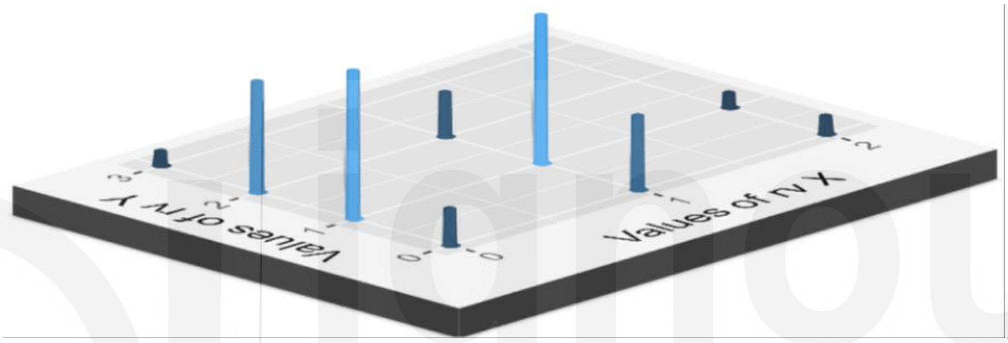
Table 6.2: Probabilities of all the possible combinations of values of X and Y

	Values of Y			
Values of X	**0**	**1**	**2**	**3**
0	$\frac{10}{165}$	$\frac{40}{165}$	$\frac{30}{165}$	$\frac{4}{165}$
1	$\frac{20}{165}$	$\frac{40}{165}$	$\frac{12}{165}$	0
2	$\frac{5}{165}$	$\frac{4}{165}$	0	0

In view of (6.31) three points (1, 3), (2, 2) and (2, 3) will not be part of the support of the bivariate random variable (X, Y) because they have 0 probability refer Table 6.2. So, remaining 9 points form support of the discrete bivariate random variable (X, Y). Joint probability mass function is visualised in Fig. 6.3 (a) and (b) from two different angles. The height of the line segments at each of the 9 points of the support represents probabilities of the corresponding point. … (6.44)



(a)



(b)

Fig. 6.3: Joint probability mass function of bivariate random variable (X, Y) (a) one view (b) another view

Now, let us define support of individual random variables X and Y in terms of support of the bivariate random variable (X, Y). If we denote individual supports of random variables X and Y by S_X and S_Y respectively then S_X and S_Y in terms of support $S_{X,Y}$ of bivariate random variable (X, Y) can be defined as follows.

$$S_X = \{x : (x, y) \in S_{X,Y}\} \text{ and } S_Y = \{y : (x, y) \in S_{X,Y}\}. \quad \ldots (6.45)$$

Note that Table 6.2 have all information regarding probabilities of all possible combinations of values of two random variables X and Y. So, this **tabular form is known as joint probability mass of the bivariate random variable (X, Y)**. The important point to be noted from here is that if Table 6.2 contains all information regarding probabilities of all possible combinations of values of two random variables X and Y then obviously it will contain all information regarding probabilities of individuals random variables X and Y. This is discussed next under the heading marginal probability mass functions.

Marginal Probability Mass Functions

The information about probabilities of individual random variables X and Y can easily be obtained by obtaining rows and columns sums of Table 6.2. After adding one row indicating columns sums and one column indicating rows sums modified table is shown in Table 6.3 as follows. … (6.46)

Table 6.3: Probabilities of all the possible combinations of values of X and Y with rows and columns sums in the bottom and right margin of the table

	Values of Y				
Values of X	**0**	**1**	**2**	**3**	**Rows sums**
0	$\frac{10}{165}$	$\frac{40}{165}$	$\frac{30}{165}$	$\frac{4}{165}$	$\mathbf{\frac{84}{165}}$
1	$\frac{20}{165}$	$\frac{40}{165}$	$\frac{12}{165}$	0	$\mathbf{\frac{72}{165}}$
2	$\frac{5}{165}$	$\frac{4}{165}$	0	0	$\mathbf{\frac{9}{165}}$
Columns sums	$\mathbf{\frac{35}{165}}$	$\mathbf{\frac{84}{165}}$	$\mathbf{\frac{42}{165}}$	$\mathbf{\frac{4}{165}}$	$\mathbf{\frac{165}{165}} = 1$

So, individual probability distribution of random variable X is given by writing probabilities shown in bold in the last column of the Table 6.3 against different values of the random variable X which are shown in the first column of the same Table 6.3. That is, we have to just write the first and the last columns of Table 6.3 in a separate table say Table 6.4 as follows. … (6.47)

Table 6.4: Marginal probability distribution of the random variable X

Values of X	0	1	2	Total
Probabilities ($p_X(x)$)	$\frac{84}{165}$	$\frac{72}{165}$	$\frac{9}{165}$	$\frac{165}{165} = 1$

Similarly, individual probability distribution of random variable Y is given by writing probabilities shown in bold in the last row of the Table 6.3 against different values of the random variable Y which are shown in the second row of the same Table 6.3. That is, we have to just write the second and the last rows of Table 6.3 in a separate table say Table 6.5 as follows. … (6.48)

Table 6.5: Marginal probability distribution of the random variable Y

Values of Y	0	1	2	3	Total
Probabilities ($p_Y(y)$)	$\frac{35}{165}$	$\frac{84}{165}$	$\frac{42}{165}$	$\frac{4}{165}$	$\frac{165}{165} = 1$

Individual probability distributions of random variables X and Y given by Table 6.4 and Table 6.5, respectively are obtained from joint probability mass function of the bivariate random variable (X, Y) so they have the special name called **marginal probability mass functions of X and Y**, respectively. The word marginal is used because probabilities of individual random variables X and Y were lying in the margins (last column and last row) of the Table 6.3. Marginal probability mass functions of random variables X and Y are visualised in Fig. 6.4 (a) and (b) respectively. … (6.49)

With the help of this example, you have understood the idea of marginal probability mass function of individual random variables X and Y. Now, let us define them in general as follows.

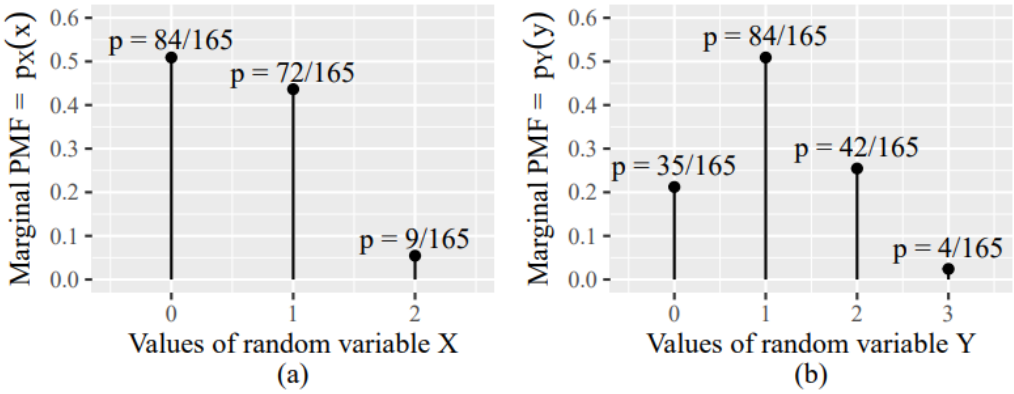


Fig. 6.4: Marginal probability mass functions of random variable (a) X (b) Y

Marginal Probability Mass Functions of X and Y: Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the probability space and X, Y be two discrete random variables having values $x_1, x_2, x_3, \ldots, x_m$ and $y_1, y_2, y_3, \ldots, y_n$ respectively. If joint probability mass function of the bivariate random variable (X, Y) is given by

$$p_{X,Y}(x, y) = \begin{cases} \mathcal{P}(X = x, Y = y), & \text{if } (x, y) \text{ is in support of } (X, Y) \\ 0, & \text{otherwise} \end{cases}$$

then marginal probability mass function of the random variable X is given by

$$\begin{aligned} p_X(x) &= \mathcal{P}(Y = y_1)\mathcal{P}(X = x \mid Y = y_1) + \mathcal{P}(Y = y_2)\mathcal{P}(X = x \mid Y = y_2) + \ldots \\ &\quad + \mathcal{P}(Y = y_n)\mathcal{P}(X = x \mid Y = y_n) \quad [\text{Using total law of probability}] \\ &= \mathcal{P}(X = x, Y = y_1) + \mathcal{P}(X = x, Y = y_2) + \ldots + \mathcal{P}(X = x, Y = y_n) \\ &\quad \left[\because \text{ We know that } \mathcal{P}(E \mid F) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)} \Rightarrow \mathcal{P}(E \cap F) = \mathcal{P}(F)\mathcal{P}(E \mid F)\right] \end{aligned}$$

$$\Rightarrow p_X(x) = \sum_{j=1}^{n} \mathcal{P}(X = x, Y = y_j) \qquad \ldots (6.50)$$

Similarly, marginal probability mass function of the random variable Y is given by

$$\begin{aligned} p_Y(y) &= \mathcal{P}(X = x_1)\mathcal{P}(Y = y \mid X = x_1) + \mathcal{P}(X = x_2)\mathcal{P}(Y = y \mid X = x_2) + \ldots \\ &\quad + \mathcal{P}(X = x_m)\mathcal{P}(Y = y \mid X = x_m) \quad [\text{Using total law of probability}] \\ &= \mathcal{P}(Y = y, X = x_1) + \mathcal{P}(Y = y, X = x_2) + \ldots + \mathcal{P}(Y = y, X = x_m) \end{aligned}$$

$$\Rightarrow p_Y(y) = \sum_{i=1}^{m} \mathcal{P}(X = x_i, Y = y) \qquad \ldots (6.51)$$

So, (6.50) and (6.51) define marginal probability mass functions of X and Y respectively in general.

Now, we define conditional probability mass functions as follows.

Conditional Probability Mass Functions

In Sec. 1.6 of Unit 1, we have learnt to calculate conditional probability of an event E (say) for some given event F (say). Recall that conditional probability of event E given event F is denoted by $\mathcal{P}(E \mid F)$. Also, recall that if events E and F are related to the sample space Ω then sample space for the event

$E|F$ reduces from Ω to F. That is in condition probability conditioning event is the reduced sample space. … (6.52)

Recall one more thing from Unit 4, if X is a random variable defined on the sample space Ω then X = x is an event G (say) where

$$G = \{\omega \in \Omega : X(\omega) = x\} \quad \dots (6.53)$$

Now, you can understand conditional probability mass function easily. Let (X, Y) be a discrete bivariate random variable on the sample space Ω where X and Y can attain values $x_1, x_2, x_3, \dots, x_m$ and $y_1, y_2, y_3, \dots, y_n$, respectively. If $p_{X,Y}(x, y)$ denotes joint PMF of (X, Y) and $p_X(x)$, $p_Y(y)$ be marginal probability mass functions of random variables X and Y respectively, then conditional probability mass function of the random variable X for given $Y = y_j$ is denoted by $p_{X|Y}(x | y_j)$ and is defined by

$$p_{X|Y}(x | y_j) = \frac{p(X = x, Y = y_j)}{p_Y(Y = y_j)}, \; p_Y(y_j) \neq 0, \; x = x_1, x_2, x_3, \dots, x_m \quad \dots (6.54)$$

$$\left[\begin{array}{l} \because \mathcal{P}(E|F) = \dfrac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)}. \text{ You may refer (1.28). Here} \\ E = \{\omega \in \Omega : X(\omega) = x\} \text{ and } F = \{\omega \in \Omega : Y(\omega) = y_j\} \end{array}\right]$$

For a valid conditional probability mass function following two conditions should be satisfied by $p_{X|Y}(x | y_j)$:

(i) $0 \leq p_{X|Y}(x | y_j) \leq 1 \;\; \forall \;\; x = x_1, x_2, x_3, \dots, x_m$ … (6.55)

(ii) $\sum_{i=1}^{m} p_{X|Y}(x_i | y_j) = 1$ … (6.56)

Similarly, conditional probability mass function of the random variable Y for given $X = x_i$ is denoted by $p_{Y|X}(y | x_i)$ and is defined by

$$p_{Y|X}(y | x_i) = \frac{p(X = x_i, Y = y)}{p_X(X = x_i)}, \; p_X(x_i) \neq 0, \; y = y_1, y_2, y_3, \dots, y_n \quad \dots (6.57)$$

For a valid conditional probability mass function following two conditions should be satisfied by $p_{Y|X}(y | x_i)$:

(i) $0 \leq p_{Y|X}(y | x_i) \leq 1 \;\; \forall \;\; y = y_1, y_2, y_3, \dots, y_m$ … (6.58)

(ii) $\sum_{j=1}^{n} p_{Y|X}(y_j | x_i) = 1$ … (6.59)

For example, in the example of 2 red, 4 blue and 5 black balls if $Y = y_j = 1$, then conditional probability mass function of the random variable X for given $Y = y_j = 1$ is given as follows and shown in Fig. 6.5 (a).

$$p_{X|Y}(x | y = 1) = \frac{p(X = x, Y = 1)}{p_Y(Y = 1)}, \; x = 0, 1, 2$$

$$= \begin{cases} 40/84, & x=0 \left[\because p_{X|Y}(X=0 \mid Y=1) = \dfrac{p(X=0,\ Y=1)}{p_Y(Y=1)} = \dfrac{40/165}{84/165} = \dfrac{40}{84}\right] \\ 40/84, & x=1 \left[\because p_{X|Y}(X=1 \mid Y=1) = \dfrac{p(X=1,\ Y=1)}{p_Y(Y=1)} = \dfrac{40/165}{84/165} = \dfrac{40}{84}\right] \\ 4/84, & x=2 \left[\because p_{X|Y}(X=2 \mid Y=1) = \dfrac{p(X=2,\ Y=1)}{p_Y(Y=1)} = \dfrac{4/165}{84/165} = \dfrac{4}{84}\right] \end{cases} \quad (6.60)$$

Similarly, conditional probability mass function of the random variable Y for given $X = x_i = 1$ is given as follows and shown in Fig. 6.5 (b).

$$p_{Y|X}(y \mid x=1) = \frac{p(Y=y,\ X=1)}{p_X(X=1)},\ \ y = 0, 1, 2$$

$$= \begin{cases} 20/72, & y=0 \left[\because p_{Y|X}(Y=0 \mid X=1) = \dfrac{p(Y=0,\ X=1)}{p_X(X=1)} = \dfrac{20/165}{72/165} = \dfrac{20}{72}\right] \\ 40/72, & y=1 \left[\because p_{Y|X}(Y=1 \mid X=1) = \dfrac{p(Y=1,\ X=1)}{p_X(X=1)} = \dfrac{40/165}{72/165} = \dfrac{40}{72}\right] \\ 12/72, & y=2 \left[\because p_{Y|X}(Y=2 \mid X=1) = \dfrac{p(Y=2,\ X=1)}{p_X(X=1)} = \dfrac{12/165}{72/165} = \dfrac{12}{72}\right] \end{cases} \quad (6.61)$$

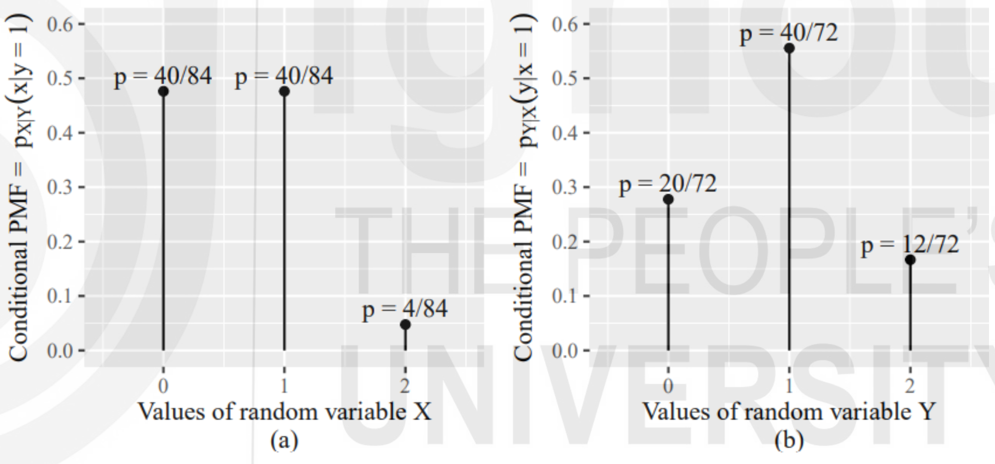


Fig. 6.5: Conditional probability mass functions of random variable (a) X given Y = 1 (b) Y given X = 1

6.4.2 Joint and Marginal Cumulative Distribution Functions

Joint Cumulative Distribution Function

Joint CDF have been already defined refer (6.12).

Keep following important point in mind. Recall that in the case of univariate random variable refer (4.36d), we have

$$\mathcal{P}(a < X \le b) = F_X(b) - F_X(a) \qquad \ldots (6.62)$$

But in bivariate case if $a, b, c, d \in \mathbb{R}$ be such that $a \le b$ and $c \le d$, then probability of the type (6.62) of the joint random variable (X, Y), i.e., $\mathcal{P}(a < X \le b,\ c < Y \le d)$ using joint CDF of random variables X and Y is given by

$$\mathcal{P}\left(a < X \le b,\ c < Y \le d\right) = F_{X,Y}(b, d) - F_{X,Y}(a, d) - F_{X,Y}(b, c) + F_{X,Y}(a, c) \ldots (6.63)$$

Let us visualise rectangular regions corresponding to each term of (6.63) in Fig. 6.6 (a1) to (a5) and specify them as follows.

- **LHS of (6.63)**: The rectangular region corresponding to LHS of (6.63) is shown by rectangle ABCD in Fig. 6.6 (a1).
- **First Term of RHS of (6.63)**: The rectangular region corresponding to the first term of RHS of (6.63) is shown by the rectangle APRT in Fig. 6.6 (a2).
- **Second Term of RHS of (6.63)**: The rectangular region corresponding to the second term of RHS of (6.63) is shown by the rectangle BPRS in Fig. 6.6 (a3).
- **Third Term of RHS of (6.63)**: The rectangular region corresponding to the third term of RHS of (6.63) is shown by the rectangle DQRT in Fig. 6.6 (a4).
- **Fourth Term of RHS of (6.63)**: The rectangular region corresponding to the fourth term of RHS of (6.63) is shown by the rectangle CQRS in Fig. 6.6 (a5).

Due to the importance of the result given by (6.63), we have visualised the idea behind each of its term. We will use this result in solving problems. So, make good understanding of this result and keep it in your mind.

Marginal Cumulative Distribution Functions

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the probability space. Let $x_1, x_2, x_3, \ldots, x_m$ and $y_1, y_2, y_3, \ldots, y_n$ be the supports of random variables X and Y respectively. Let us define events A and $E_i, 1, 2, 3, \ldots, n$ as follows.

$A = \{\omega \in \Omega : X(\omega) \le x\},\ x = x_1, x_2, x_3, \ldots, x_m$ and

$E_j = \{\omega \in \Omega : Y(\omega) = y_j\},\ j = 1, 2, 3, \ldots, n$

Since $y_1, y_2, y_3, \ldots, y_n$ form support of the random variable Y, so $\mathcal{P}\left(E_j\right) > 0,\ \forall\ j = 1, 2, 3, \ldots, n.$

So, using total law of probability, marginal CDF of the random variable X is given by

$$F_X(x) = \mathcal{P}\left(X \le x\right) = \mathcal{P}\left(A\right) = \sum_{j=1}^{n} \mathcal{P}\left(E_j\right)\mathcal{P}\left(A \mid E_j\right)$$

$$= \sum_{j=1}^{n} \mathcal{P}\left(Y = y_j\right)\mathcal{P}\left(X \le x \mid Y = y_j\right)$$

$$= \sum_{j=1}^{n} \mathcal{P}\left(X \le x,\ Y = y_j\right)\left[\because\ \mathcal{P}\left(E \mid F\right) = \frac{\mathcal{P}\left(E \cap F\right)}{\mathcal{P}\left(F\right)} \Rightarrow \mathcal{P}\left(F\right)\mathcal{P}\left(E \mid F\right) = \mathcal{P}\left(E \cap F\right)\right]$$

$$\Rightarrow F_X(x) = F_X(x, \infty) \qquad [\text{By definition of CDF}] \qquad \ldots (6.64)$$

$$\text{or } F_X(x) = \lim_{y \to \infty} F_X(x, y) \qquad [\text{By definition of CDF}] \qquad \ldots (6.65)$$

Similarly, marginal CDF of the random variable Y is given by

$$F_Y(y) = F_{X,Y}(\infty, y) \quad \dots (6.66)$$

$$\text{or } F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y) \quad \dots (6.67)$$

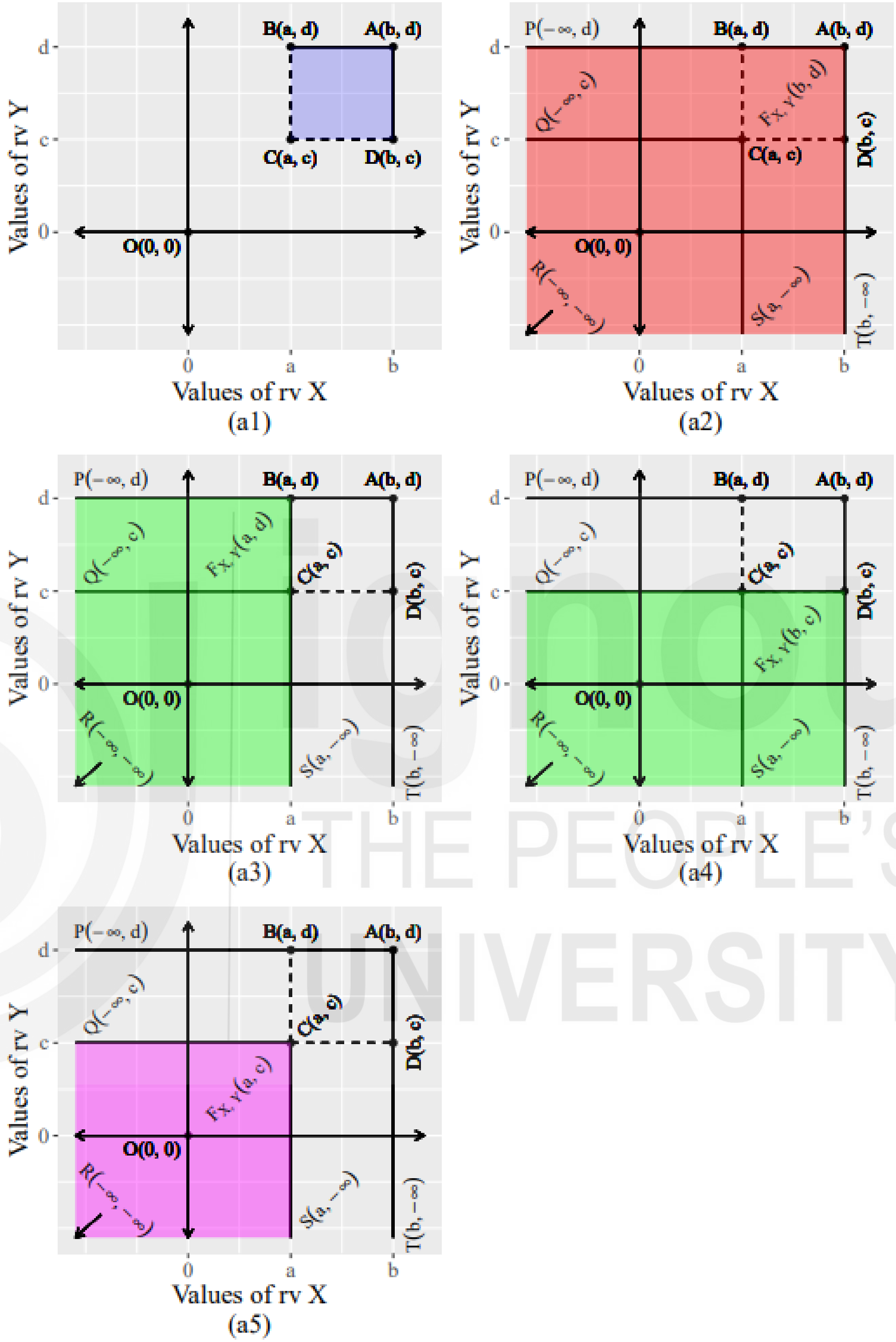


Fig. 6.6: Visualisation of different terms of (6.63) (a) LHS (b) first term of RHS (c) second term of RHS (d) third term of RHS (e) fourth term of RHS

6.4.3 Independence of Random Variables

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the probability space. Let $x_1, x_2, x_3, \dots, x_m$ and $y_1, y_2, y_3, \dots, y_n$ be the supports of random variables X and Y respectively.

Here, we want to discuss independence of two random variables. Recall that in Sec. 1.8 of Unit 1, we have learnt when two or more events are said to independent and what is the condition for independence, you may refer (1.45) and (1.47) to (1.49). So, it will help you in understanding the idea of independence of two random variables if we connect independence of random variables X and Y with the independence of events discussed in Unit 1. Independence of two random variables can be expressed in three equivalent ways mentioned as follows.

- Independence in terms of CDF
- Independence in terms of PMF
- Independence in terms of conditional PMF

Let us discuss these taken one at a time.

- **Independence in Terms of CDF**

As mentioned earlier, we will connect independence of random variables with independence of events. Keeping this in view, let us define events E and F as follows.

$$\begin{aligned} E &= \{\omega \in \Omega : X(\omega) \le x\},\ x = x_1, x_2, x_3, \ldots, x_m \text{ and} \\ F &= \{\omega \in \Omega : Y(\omega) \le y\},\ y = y_1, y_2, y_3, \ldots, y_m \end{aligned} \qquad \ldots (6.68)$$

We know that (refer 1.47) if events E and F are independent under the probability measure $\mathcal{P}$ then

$$\mathcal{P}(E \cap F) = \mathcal{P}(E)\mathcal{P}(F) \qquad \ldots (6.69)$$

Now, if events E and F are independent for all possible values of x and y then

$$\begin{aligned} F_{X,Y}(x, y) &= \mathcal{P}(X \le x,\ Y \le y) && [\text{Using definition of CDF refer (6.10)}] \\ &= \mathcal{P}(E \cap F) && [\text{Using (6.68)}] \\ &= \mathcal{P}(E)\mathcal{P}(F) && [\text{Using (6.69)}] \\ &= \mathcal{P}(X \le x)\mathcal{P}(Y \le y) && [\text{Using (6.68)}] \\ &= F_X(x)F_X(x) && [\text{Using (4.36a)}] \end{aligned}$$

So, random variables X and Y are independent if following holds.

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)\ \ \forall\ \ x \in S_X \text{ and } y \in S_Y \qquad \ldots (6.70)$$

- **Independence in Terms of PMF**

Equation (6.70) expresses condition of independence in terms of CDF. Similarly, we can express condition of independence in terms of PMF instead of CDF if we define events E and F as follows.

$$\begin{aligned} E &= \{\omega \in \Omega : X(\omega) = x\},\ x = x_1, x_2, x_3, \ldots, x_m \text{ and} \\ F &= \{\omega \in \Omega : Y(\omega) = y\},\ y = y_1, y_2, y_3, \ldots, y_m \end{aligned}$$

Following similar steps, we can prove that random variables X and Y will be independent if

$$p_{X,Y}(x, y) = p_X(x)p_X(x)\ \ \forall\ \ x \in S_X \text{ and } y \in S_Y \qquad \ldots (6.71)$$

or $\mathcal{P}(X = x,\ Y = y) = \mathcal{P}(X = x)\mathcal{P}(Y = y)\ \ \forall\ \ x \in S_X$ and $y \in S_Y$ ... (6.72)

- **Independence in Terms of Conditional PMF**

You have studied conditional PMF in this section refer (6.52) to (6.61). Two random variables X and Y are said to be independent if conditional PMF's satisfy the following conditions

$$p_{X|Y}(x\,|\,y) = p_X(x) \qquad \forall\ \ x \in S_X,\ y \in S_Y \text{ and } p_Y(y) > 0 \qquad \text{... (6.73)}$$

$$\text{or}\quad p_{Y|X}(y\,|\,x) = p_Y(y)\ \ \forall\ \ x \in S_X,\ y \in S_Y \text{ and } p_X(x) > 0 \qquad \text{... (6.74)}$$

Now, we discuss one example to apply different ideas discussed in this section.

Example 1: Suppose we have a binary communication channel where support of both input and output random variables X and Y is {0, 1}. Input random variable X attains values 0 and 1 in equally likely fashion. But channel noise is a common problem with a communication channel due to which input 0 may be transmitted to 1 and similarly input 1 may be transmitted to 0. We are given the following channel transition probabilities:

$$\mathcal{P}(Y = 0\,|\,X = 0) = 0.99 \text{ and } \mathcal{P}(Y = 1|\,X = 1) = 0.92. \qquad \text{… (6.75)}$$

On the basis of this information obtain:

(a) joint probability mass function of (X, Y).

(b) marginal PMF's of X and Y.

(c) joint CDF of (X, Y).

(d) marginal CDF's of X and Y.

(e) conditional PMF of Y given X = 1.

(f) are random variables X and Y independent?

Solution: Since input values 0 and 1 are equally likely and X represents input random variable, so, we have

$$\mathcal{P}(X = 0) = 0.5 \text{ and } \mathcal{P}(X = 1) = 0.5. \qquad \text{… (6.76)}$$

Also, if input is 0 then either output will be 0 or 1, so, in conditional space, we should have

$$\mathcal{P}(Y = 0\,|\,X = 0) + \mathcal{P}(Y = 1|\,X = 0) = 1$$

$$\begin{aligned}\Rightarrow \mathcal{P}(Y = 1|\,X = 0) &= 1 - \mathcal{P}(Y = 0\,|\,X = 0)\\ &= 1 - 0.99 \qquad [\text{Using } (6.75)]\\ &= 0.01 \qquad \text{... (6.77)}\end{aligned}$$

Similarly, $\mathcal{P}(Y = 0\,|\,X = 1) = 1 - \mathcal{P}(Y = 1|\,X = 1) = 1 - 0.92 = 0.08$ … (6.78)

Model input values, transition probabilities, output values and probabilities obtained in (6.77) and (6.78) all are shown in Fig. 6.7.

(a) Here random variables X and Y both attain values 0 and 1. Since random variables X and Y both assume finite number of values (2 in this case), so (X, Y) is a discrete bivariate random variable. To obtain joint PMF of the

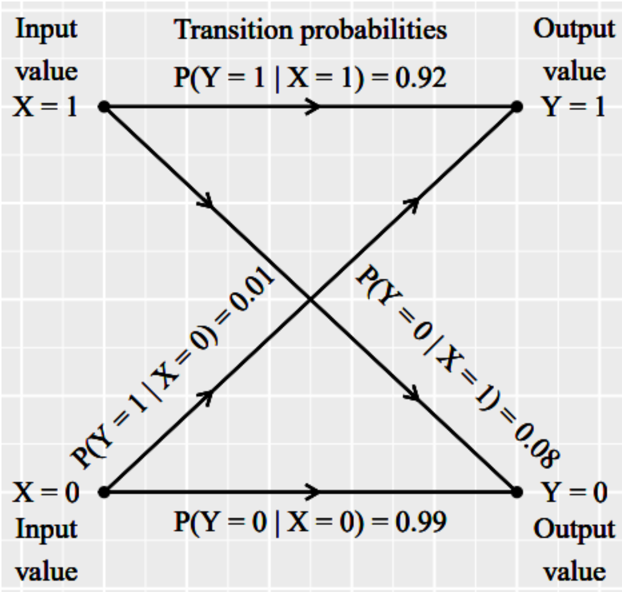


Fig. 6.7: Visualisation of the communication channel with transition probabilities, input values and output values

discrete bivariate random variable (X, Y), we have to fill up four empty cells of the Table 6.6 with corresponding probabilities.

Table 6.6: Number of cells formed by possible values of X and Y

	Values of Y	
Values of X	0	1
0		
1		

Let us obtain probabilities of these four cells as follows.

$$\mathcal{P}(X=0,\ Y=0)=\mathcal{P}(X=0)\mathcal{P}(Y=0\,|\,X=0)$$
$$=0.5\times 0.99 \qquad \text{[Using (6.76) and (6.75)]}$$
$$=0.495$$

Similarly, $\mathcal{P}(X=0,\ Y=1)=\mathcal{P}(X=0)\mathcal{P}(Y=1|\,X=0)$

$$=0.5\times 0.01=0.005 \text{ [Using (6.76) and (6.77)]}$$

$$\mathcal{P}(X=1,\ Y=0)=\mathcal{P}(X=1)\mathcal{P}(Y=0\,|\,X=1)$$
$$=0.5\times 0.08=0.04 \qquad \text{[Using (6.76) and (6.78)]}$$

$$\mathcal{P}(X=1,\ Y=1)=\mathcal{P}(X=1)\mathcal{P}(Y=1|\,X=1)$$
$$=0.5\times 0.92=0.46 \qquad \text{[Using (6.76) and (6.75)]}$$

We have obtained probabilities of the four cells of the Table 6.6 and are shown in the new Table 6.7. The tabular form given by Table 6.7 is known as joint PMF of the random variable (X, Y).

Table 6.7: Joint probability mass function of discrete bivariate random variable (X, Y)

	Values of Y	
Values of X	0	1
0	0.495	0.005
1	0.04	0.46

(b) To obtain marginal PMF's of the random variables X and Y, first, we have to find out sum of each row and column of Table 6.7. Rows and columns sums are shown in Table 6.8.

Table 6.8: Marginal probability mass functions of discrete random variables X and Y as sums of each row and column

	Values of Y		
Values of X	**0**	**1**	**Marginal probabilities**
0	0.495	0.005	**0.5**
1	0.04	0.46	**0.5**
Marginal probabilities	**0.535**	**0.465**	1

So, marginal probability mass function of the random variable X is given by in Table 6.9 as follows.

Table 6.9: Marginal probability mass function of the random variable X

Values of X	0	1	Total
Probabilities ($p_X(x)$)	0.5	0.5	1

Similarly, marginal probability mass function of the random variable Y is given by in Table 6.10 as follows

Table 6.10: Marginal probability mass function of the random variable Y

Values of Y	0	1	Total
Probabilities ($p_Y(y)$)	0.535	0.465	1

(c) For finding joint CDF of (X, Y), we have to compute $F_{X,\,Y}(x,\ y)$ for different values of x and y using (6.10). So, joint CDF of discrete bivariate random variable (X, Y) is given by Table 6.11.

Table 6.11: Joint CDF of discrete bivariate random variable (X, Y)

	Values of Y	
Values of X	**0**	**1**
0	0.495	0.5
1	0.535	1

$$\left[\begin{array}{l} \because F_{X,\,Y}(0,\,0) = \mathcal{P}\left(X \le 0,\ Y \le 0\right) = \mathcal{P}\left(X = 0,\ Y = 0\right) = 0.495 \\ F_{X,\,Y}(0,\,1) = \mathcal{P}\left(X \le 0,\ Y \le 1\right) = \mathcal{P}\left(X = 0,\ Y = 0\right) + \mathcal{P}\left(X = 0,\ Y = 1\right) \\ \qquad\qquad\qquad\qquad\qquad = 0.495 + 0.005 = 0.5,\ \text{etc.} \end{array}\right]$$

(d) Marginal CDF of random variable X is given by combining values of X from the first column and corresponding probabilities from the third column of Table 6.11 and given by Table 6.12.

Table 6.12: Marginal CDF of the random variable X

Values of X	0	1
Probabilities ($p_X(x)$)	0.5	1

Marginal CDF of random variable Y is given by combining values of Y from the first row and corresponding probabilities from the third row of Table 6.11 and given by Table 6.13.

Table 6.13: Marginal CDF of the random variable Y

Values of Y	0	1
Probabilities ($p_Y(y)$)	0.535	1

(e) We can obtain conditional PMF of random variable Y for given X = 1 as follows.

$$p_{Y|X}(y \mid X=1) = \frac{p_{X,\,Y}(X=1,\ Y=y)}{p_X(X=1)},\ \ y = 0, 1$$

$$= \begin{cases} 0.08, & y=0\left[\because\ p_{Y|X}(Y=0 \mid X=1) = \dfrac{p_{X,\,Y}(X=1,\ Y=0)}{p_X(X=1)} = \dfrac{0.04}{0.5} = 0.08\right] \\ 0.92, & y=1\left[\because\ p_{Y|X}(Y=1 \mid X=1) = \dfrac{p_{X,\,Y}(X=1,\ Y=1)}{p_X(X=1)} = \dfrac{0.46}{0.5} = 0.92\right] \end{cases}$$

(f) $\mathcal{P}(X=0,\ Y=0) = 0.495$ ... (6.79)

$\mathcal{P}(X=0)\mathcal{P}(Y=0) = 0.5 \times 0.535 = 0.2675$ ... (6.80)

From (6.79) and (6.80) $\mathcal{P}(X=0,\ Y=0) \neq \mathcal{P}(X=0)\mathcal{P}(Y=0)$

Since (6.72) does not hold so random variables X and Y are not independent.

Now, you can try the following Self-Assessment Question.

SAQ 1

Joint PMF is given and find out other things: Joint PMF of a discrete bivariate random variable (X, Y) is given by

$$p_{X,\,Y}(x,\ y) = \begin{cases} k(2x+3y), & x = 0, 1, 2;\ y = 1, 2, 3, 4 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find value of k.

(b) Find the joint PMF of (X, Y) in formula form and tabular form. Also find marginal PMF's of X and Y.

(c) Find joint CDF of (X, Y).

(d) Find marginal CDF's of X and Y.

(e) Find conditional PMF of X given Y = 3.

(f) Are X and Y independent?

(g) Find $\mathcal{P}(0 < X \le 2, 1 < Y \le 3)$.

6.5 BIVARIATE CONTINUOUS RANDOM VARIABLE

In the previous section, you have studied discrete bivariate random variable and many concepts related to that like joint, marginal and conditional PMF's; joint and marginal CDF's; and finally, independence. In this section, we will study such concepts for continuous bivariate random variable. But in Unit 5, you have already studied univariate continuous random variable. So, from Unit 5, you know how to handle things related to a continuous random variable in continuous world. Combining the knowledge of Unit 5 and previous section of this unit, you can easily understand the corresponding concepts for continuous bivariate random variable. So, we are not going to explain them in detail like discrete bivariate random variable in the previous section. Here, we will define the terms related to continuous bivariate random variable in brief as follows and after that we do one example to explain them.

Joint Cumulative Distribution Function

Joint CDF has been already defined refer (6.12).

Recall that in the case of discrete bivariate random variable refer (6.63), we have

$$\mathcal{P}(a < X \le b,\ c < Y \le d) = F_{X,\,Y}(b,\,d) - F_{X,\,Y}(a,\,d) - F_{X,\,Y}(b,\,c) + F_{X,\,Y}(a,\,c) \quad \ldots (6.81)$$

But like univariate case in continuous world a single point has zero probability, i.e.,

$$\mathcal{P}(X = a,\ Y = b) = 0 \quad \ldots (6.82)$$

In view of (6.82), we can say that all expressions like
$\mathcal{P}(a < X \le b,\ c < Y \le d),\ \mathcal{P}(a \le X \le b,\ c < Y \le d),\ \mathcal{P}(a \le X < b,\ c < Y \le d),$
$\mathcal{P}(a < X < b,\ c < Y \le d),\ \mathcal{P}(a < X \le b,\ c \le Y \le d),\ \mathcal{P}(a < X \le b,\ c \le Y < d)$, etc.
all are equal to $F_{X,\,Y}(b,\,d) - F_{X,\,Y}(a,\,d) - F_{X,\,Y}(b,\,c) + F_{X,\,Y}(a,\,c)$. … (6.83)

Marginal Cumulative Distribution Functions

Marginal CDF of the random variable X is given by

$$F_X(x) = \mathcal{P}(X \le x) = F_X(x,\,\infty) = \lim_{y \to \infty} F_X(x,\,y) \quad \ldots (6.84)$$

Similarly, marginal CDF of the random variable Y is given by

$$F_Y(y) = F_{X,\,Y}(\infty,\,y) = \lim_{x \to \infty} F_{X,\,Y}(x,\,y) \quad \ldots (6.85)$$

Joint Probability Density Function of a Jointly Continuous Bivariate Random Variable

Let (X, Y) be a jointly continuous random variable. The jointly continuous random variable (X, Y) will have a joint probability density function (JPDF) $f_{X,\,Y}$ if there exists a Borel measurable function $f_{X,\,Y} : \mathbb{R} \times \mathbb{R} \to [0,\,\infty]$ such that

$$\mathcal{P}_{X,\,Y}(B) = \mathcal{P}((X,\,Y) \in B) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_B f_{X,\,Y}(x,\,y)\,dx\,dy,\ B \in \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \quad \ldots (6.86)$$

where I_B is indicator function of B

In particular, if $B = (a,\,b] \times (c,\,d]$, then

$$\mathcal{P}_{X,\,Y}((a,\,b] \times (c,\,d]) = \mathcal{P}((X,\,Y) \in (a,\,b] \times (c,\,d]) = \int_c^d \int_a^b f_{X,\,Y}(x,\,y)\,dx\,dy \quad \ldots (6.87)$$

In Secs. 9.4 and 9.5 of Unit 9 of the course MST-011, you have studied to evaluate double integral. You also know that double integral shown in (6.87) gives volume of the region bounded by five planes x = a, x = b, y = c, y = d, z = 0 and one surface given by $z = f_{X, Y}(x, y)$. But in the world of probability theory double integral shown in (6.87) gives the probability $\mathcal{P}\left((X, Y) \in (a, b] \times (c, d]\right)$. Therefore, to be a **valid joint PDF** it must satisfy the following two conditions.

(a) $f_{X, Y}(x, y) \geq 0 \quad \forall \ (x, y) \in \mathbb{R} \times \mathbb{R} \begin{bmatrix} \because \text{Probabilities are always non-negative} \\ \Rightarrow \text{density shoud be} \geq 0 \end{bmatrix}$ … (6.88)

(b) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_B f_{X, Y}(x, y)\, dx\, dy = 1 \begin{bmatrix} \because \text{Sum of all probabilities should be 1 and in} \\ \text{continuous world sum is given by integration} \end{bmatrix}$ … (6.89)

Keeping (6.87) in view, (6.10) can be written as follows.

$$F_{X, Y}(x, y) = \mathcal{P}\left((X, Y) \in (-\infty, x] \times (-\infty, y]\right) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X, Y}(x, y)\, dx\, dy \quad \text{… (6.90)}$$

Marginal Probability Density Functions

If $f_{X, Y}(x, y)$ be the joint PDF of jointly continuous random variable (X, Y) then marginal probability density function of X and Y, respectively, are given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X, Y}(x, y)\ dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X, Y}(x, y)\ dx \quad \text{… (6.91)}$$

Conditional Probability Density Functions

If $f_{X, Y}(x, y)$ be the joint PDF of jointly continuous random variable (X, Y) and $f_X(x)$, $f_Y(y)$ be marginal PDF's of X and Y respectively, then

- Conditional PDF of X given Y is given by

$$f_{X|Y}(x\,|\,y) = \frac{f_{X, Y}(x, y)}{f_Y(y)}, \ f_Y(y) > 0 \ \text{and} \quad \text{… (6.92)}$$

- Conditional PDF of Y given X is given by

$$f_{Y|X}(y\,|\,x) = \frac{f_{X, Y}(x, y)}{f_X(x)}, \ f_X(x) > 0 \quad \text{… (6.93)}$$

Independence of Random Variables

Like discrete bivariate case, independence of two jointly continuous random variables X and Y can be expressed in three ways.

- Independence in terms of CDF
- Independence in terms of PDF
- Independence in terms of conditional PDF

Let us discuss these taken one at a time.

- **Independence in Terms of CDF**

Jointly continuous random variables X and Y are independent if the following holds.

$$F_{X,\,Y}(x,\,y) = F_X(x)F_Y(y) \;\; \forall \;\; x,\, y \in \mathbb{R} \qquad \ldots (6.94)$$

- **Independence in Terms of PDF**

Jointly continuous random variables X and Y are independent if the following holds.

$$f_{X,\,Y}(x,\,y) = f_X(x)\,f_Y(y) \;\; \forall \;\; x,\, y \in \mathbb{R} \qquad \ldots (6.95)$$

- **Independence in Terms of Conditional PDF**

Jointly continuous random variables X and Y are independent if the following holds.

$$f_{X|Y}\left(x\,|\,y\right) = f_X\left(x\right) \qquad \forall \;\; x,\, y \in \mathbb{R} \;\text{ and }\; f_Y\left(y\right) > 0 \qquad \ldots (6.96)$$

$$\text{or } \; f_{Y|X}\left(y\,|\,x\right) = f_Y\left(y\right) \quad \forall \;\; x,\, y \in \mathbb{R} \;\text{ and }\; f_X(x) > 0 \qquad \ldots (6.97)$$

Now, we discuss one example to apply different ideas discussed in this section.

Example 2: Joint PDF of jointly continuous random variable (X, Y) is given by

$$f_{X,\,Y}(x,\,y) = \begin{cases} k(4-x)(5-y), & 0 \le x \le 4,\; 0 \le y \le 5 \\ 0, & \text{otherwise} \end{cases} \qquad \ldots (6.98)$$

Obtain:

(a) value of k.

(b) marginal PDF's of X and Y.

(c) joint CDF of (X, Y).

(d) marginal CDF's of X and Y.

(e) conditional PDF of X given Y.

(f) are random variables X and Y independent?

Solution: (a) We can find the value of k using (6.89) k as follows.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_B f_{X,\,Y}(x,\,y)\,dx\,dy = 1 \Rightarrow \int_0^5\int_0^4 k(4-x)(5-y)\,dx\,dy = 1$$

$$\Rightarrow \int_0^5 k(5-y)\left[4x - \frac{x^2}{2}\right]_{x=0}^{x=4} dy = 1 \Rightarrow \int_0^5 k(5-y)\left[16-8\right]dy = 1$$

$$\Rightarrow 8k\left[5y - \frac{y^2}{2}\right]_{y=0}^{y=5} = 1 \Rightarrow 8k\left[25 - \frac{25}{2}\right] = 1 \Rightarrow 100k = 1 \Rightarrow k = \frac{1}{100} \qquad \ldots (6.99)$$

(b) We can find marginal PDF of X using (6.91) as follows.

$$f_X\left(x\right) = \int_{-\infty}^{\infty} f_{X,\,Y}(x,\,y)\;dy = \frac{4-x}{100}\int_0^5 (5-y)\,dy = \frac{4-x}{100}\left[5y - \frac{y^2}{2}\right]_{y=0}^{y=5}$$

$$= \frac{4-x}{100}\left[25 - \frac{25}{2}\right] = \frac{4-x}{8},\;\; 0 \le x \le 4 \qquad \ldots (6.100)$$

Similarly, marginal PDF of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx = \frac{5-y}{100}\int_0^4 (4-x)\,dx = \frac{5-y}{100}\left[4x - \frac{x^2}{2}\right]_{x=0}^{x=4}$$

$$= \frac{5-y}{100}[16-8] = \frac{10-2y}{25},\ 0 \le y \le 5 \qquad \text{... (6.101)}$$

(c) Using (6.90), joint CDF is given by

$$F_{X,Y}(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(x, y)\,dx\,dy \qquad \text{... (6.102)}$$

To evaluate (6.102), we generally consider five cases I, II, III, IV and V as shown in Fig. 6.8 (a) for the general case where $a \le x \le b$ and $c \le y \le d$.

In the present case, we have a = 0, b = 4, c = 0, d = 5. So, we have to consider the following five cases.

Case I: x < 0 or y < 0

In this case in view of (6.98) and (6.102), we have $F_{X,Y}(x, y) = 0$.

Case II: $x \ge 4$ and $y \ge 5$

In this case, we have $F_{X,Y}(x, y) = 1$ $[\because$ Sum of all probability is $1]$

Case III: $0 \le x \le 4$ and $0 \le y \le 5$

$$F_{X,Y}(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(x, y)\,dx\,dy = \int_0^y\int_0^x \frac{(5-y)(4-x)}{100}\,dx\,dy$$

$$= \int_0^y \frac{(5-y)}{100}\left[4x - \frac{x^2}{2}\right]_0^x dy = \int_0^y \frac{(5-y)}{100}\left[4x - \frac{x^2}{2}\right] dy$$

$$= \frac{1}{100}\left[4x - \frac{x^2}{2}\right]\left[5y - \frac{y^2}{2}\right]_0^y = \frac{(8x - x^2)(10y - y^2)}{400} \qquad \text{... (6.103)}$$

Case IV: $0 \le x \le 4$ and $y \ge 5$

$$F_{X,Y}(x, y) = F_{X,Y}(x, 5)$$

$$= \frac{(8x - x^2)(10\times 5 - 5^2)}{400} \quad [\text{Putting } y = 5 \text{ in } (6.103)]$$

$$= \frac{8x - x^2}{16} \qquad \text{... (6.104)}$$

Case V: $x \ge 4$ and $0 \le y \le 5$

$$F_{X,Y}(x, y) = F_{X,Y}(4, y)$$

$$= \frac{(8\times 4 - 4^2)(10y - y^2)}{400} \quad [\text{Putting } x = 4 \text{ in } (6.103)]$$

$$= \frac{10y - y^2}{25} \qquad \text{... (6.105)}$$

Hence, on combining all cases CDF of (X, Y) is given by

$$F_{X,Y}(x,y)=\begin{cases}0, & \text{if } x<0 \text{ or } y<0\\ \dfrac{(8x-x^2)(10y-y^2)}{400}, & 0\le x\le 4,\ 0\le y\le 5\\ \dfrac{8x-x^2}{16}, & 0\le x\le 4,\ y\ge 5\\ \dfrac{10y-y^2}{25}, & x\ge 4,\ 0\le y\le 5\\ 1, & \text{if } x\ge 4 \text{ and } y\ge 5\end{cases} \qquad \ldots (6.106)$$

(d) Combining (6.84) and (6.104), marginal CDF of X is given by

$$F_X(x)=\begin{cases}0, & \text{if } x<0\\ \dfrac{8x-x^2}{16}, & 0\le x\le 4\\ 1, & \text{if } x\ge 4\end{cases}$$

Similarly, combining (6.85) and (6.105) marginal CDF of Y is given by

$$F_Y(y)=\begin{cases}0, & \text{if } y<0\\ \dfrac{10y-y^2}{25}, & 0\le y\le 5\\ 1, & \text{if } y\ge 5\end{cases}$$

(e) Using (6.92), conditional density of X given Y is given by

$$f_{X|Y}(x\,|\,y)=\frac{f_{X,Y}(x,y)}{f_Y(y)},\ f_Y(y)>0$$

$$=\frac{\dfrac{(8x-x^2)(10y-y^2)}{400}}{\dfrac{10y-y^2}{25}} \qquad [\text{Using (6.98), (6.99) and (6.101)}]$$

$$=\frac{8x-x^2}{16},\ 0\le x\le 4$$

(f) Since $F_{X,Y}(x,y)=F_X(x)F_Y(y)\ \ \forall\ x,y\in\mathbb{R}$

Hence, random variables X and Y are independent.

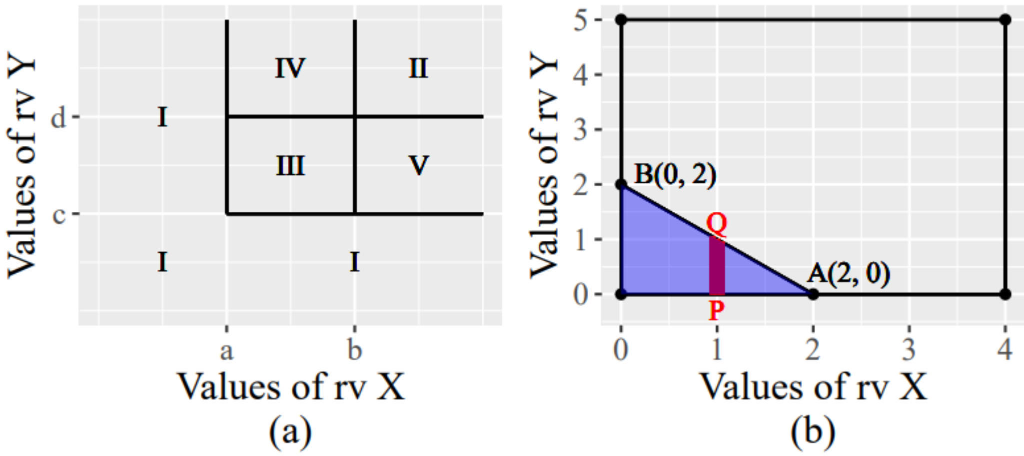


Fig. 6.8: Visualisation of the (a) five regions corresponding to five cases in part (c) of Example 2 (b) vertical strip PQ to obtain limits of y in terms of x refer solution of SAQ 3

In this example, joint PDF was given and we obtained joint CDF using the integral given by (6.90). But you know that differentiation and integration are reverse operators so if joint CDF is given then joint PDF can be obtained by partially differentiating CDF with respect to x and y, i.e.,

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y}\left(F_{X,Y}(x, y)\right) \text{ or } f_{X,Y}(x, y) = \frac{\partial^2}{\partial y \partial x}\left(F_{X,Y}(x, y)\right) \quad \text{... (6.107)}$$

$$\left[\begin{array}{l}\text{Here we are assuming that } \frac{\partial^2}{\partial x \partial y}\left(F_{X,Y}(x, y)\right) = \frac{\partial^2}{\partial y \partial x}\left(F_{X,Y}(x, y)\right)\text{. But in} \\ \text{general it does not hold. In fact, it holds only when first order partial} \\ \text{derivatives } \frac{\partial}{\partial x}\left(F_{X,Y}(x, y)\right) \text{ and } \frac{\partial}{\partial x}\left(F_{X,Y}(x, y)\right) \text{ are continuous functions}\end{array}\right]$$

You will study partial derivative in the course MST-022 in detail. Let us briefly explain it through an example as follows. … (6.107a)

$\frac{\partial^2}{\partial x \partial y}\left(F(x,y)\right)$ means first differentiate $F(x,y)$ partially w.r.t. y and then the resulting function w.r.t. x. When we differentiate a function partially w.r.t. one variable, then the other variable is treated as constant … (6.107b)

For example, Let $F(x,y) = x^5y^3 + x^{13}y^7$

If we differentiate it partially w.r.t. y, we have

$$\frac{\partial}{\partial y}\left(F(x,y)\right) = x^5\left(3y^2\right) + x^{13}\left(7y^6\right) = 3x^5y^2 + 7x^{13}y^6 \quad \begin{bmatrix}\text{Here, x is treated} \\ \text{as constant}\end{bmatrix}$$

If we now partially differentiate this resulting expression w.r.t. x, we have

$$\frac{\partial^2}{\partial x \partial y}\left(F(x,y)\right) = 3\left(5x^4\right)y^2 + 7\left(13x^{12}\right)y^6 = 15x^4y^2 + 91x^{12}y^6 \begin{bmatrix}\text{Here, y is treated} \\ \text{as constant}\end{bmatrix}$$

If we first partially differentiate w.r.t. x and then w.r.t. y, we get

$$\frac{\partial}{\partial y}\left(F(x,y)\right) = \left(5x^4\right)y^3 + 13x^{12}y^7 = 5x^4y^3 + 13x^{12}y^7 \quad \begin{bmatrix}\text{Here, y is treated} \\ \text{as constant}\end{bmatrix}$$

$$\frac{\partial^2}{\partial y \partial x}\left(F(x,y)\right) = 5x^4\left(3y^2\right) + 13x^{12}\left(7y^6\right) = 15x^4y^2 + 91x^{12}y^6 \begin{bmatrix}\text{Here, y is treated} \\ \text{as constant}\end{bmatrix}$$

Note that here both way we get the same result.

Now, you can try the following two Self-Assessment Questions.

SAQ 2

For given Joint CDF Obtain Joint PDF: In Example 2 joint PDF was given and you obtained joint CDF given by (6.106). Here assume that joint CDF is given to you. Find joint PDF of the jointly continuous random variable (X, Y).

SAQ 3

For given CDF obtain PDF and probabilities: Consider the same jointly continuous random variable (X, Y) discussed in Example 2. Find $\mathcal{P}\left(X + Y \le 2\right)$.

6.6 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- **Bivariate Random Variable**: Let $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ be two random variables. The ordered pair random variable (X, Y) which associates a unique ordered pair of real numbers to each member of the sample space (Ω) is called a **bivariate random variable** on Ω.

- If both X and Y are discrete random variables then the ordered pair random variable (X, Y) is called a **discrete bivariate random variable**. If both X and Y are **jointly** continuous random variables then the ordered pair random variable (X, Y) is called a **continuous bivariate random variable**.

- **CDF of (X, Y) is defined by**

  $$F_{\mathcal{P}_{X,Y}}(x, y) = \mathcal{P}_{X,Y}\left((-\infty, x] \times (-\infty, y]\right) = \mathcal{P}\left(X \le x,\ Y \le y\right),\ (x, y) \in \mathbb{R}^2$$

- **Properties of CDF are listed as follows**:

  - $F_{X,Y}(-\infty, -\infty) = \lim\limits_{\substack{x \to -\infty \\ y \to -\infty}} F_{X,Y}(x, y) = 0$

  - $F_{X,Y}(-\infty, y) = \lim\limits_{x \to -\infty} F_{X,Y}(x, y) = 0$

  - $F_{X,Y}(x, -\infty) = \lim\limits_{y \to -\infty} F_{X,Y}(x, y) = 0$

  - $F_{X,Y}(\infty, \infty) = \lim\limits_{\substack{x \to \infty \\ y \to \infty}} F_{X,Y}(x, y) = 1$

  - Marginal CDF of X is $\lim\limits_{y \to \infty} F_{X,Y}(x, y) = F_{X,Y}(x, \infty) = F_X(x)$.

  - Marginal CDF of Y is $\lim\limits_{x \to \infty} F_{X,Y}(x, y) = F_{X,Y}(\infty, y) = F_Y(y)$.

- $F_{X,Y}$ is **increasing or non-decreasing**. That is if $x_1, x_2, y_1, y_2 \in \mathbb{R}$ be such that $x_1 \le x_2$ and $y_1 \le y_2$, then $F_{X,Y}(x_1, y_1) \le F_{X,Y}(x_2, y_2)$.

- $F_{X,Y}$ is continuous from above, i.e.,

  $$\lim_{\substack{h \downarrow 0 \\ k \downarrow 0}} F_{X,Y}(x + h, y + k) = F_{X,Y}(x, y) \quad \forall\ x, y \in \mathbb{R}.$$

- If $x_1, x_2, y_1, y_2 \in \mathbb{R}$ be such that $x_1 \le x_2$ and $y_1 \le y_2$, then (to get more detail of this property refer 6.63 and Fig. 6.6)

  $$\begin{aligned}\mathcal{P}\left(x_1 < X \le x_2,\ y_1 < Y \le y_2\right) &= F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) \\ &\quad - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1)\end{aligned}$$

- **Joint PMF**: $p_{X,Y}(x, y) = \begin{cases} \mathcal{P}\left(X = x, Y = y\right), & \text{if } (x, y) \text{ is in support of } (X, Y) \\ 0, & \text{otherwise} \end{cases}$

- Marginal probability mass function of the random variable X is given by

  $$p_X(x) = \sum_{j=1}^{n} \mathcal{P}\left(X = x, Y = y_j\right)$$

Similarly, marginal PMF of Y is given by $p_Y(y) = \sum_{i=1}^{m} \mathcal{P}\left(X = x_i,\ Y = y\right)$

- Conditional PMF of X given Y is

$$p_{X|Y}(x \mid y_j) = \frac{p(X = x,\ Y = y_j)}{p_Y(Y = y_j)},\ p_Y(y_j) \neq 0,\ x = x_1, x_2, x_3, \ldots, x_m$$

Similarly, conditional PMF of Y given X is

$$p_{Y|X}(y \mid x_i) = \frac{p(X = x_i,\ Y = y)}{p_X(X = x_i)},\ p_X(x_i) \neq 0,\ y = y_1, y_2, y_3, \ldots, y_n$$

- Two random variables X and Y are said to be independent if
  - For both discrete and continuous case

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)\ \forall\ x \in S_X \text{ and } y \in S_Y$$

  - For discrete case

$$p_{X,Y}(x, y) = p_X(x)p_X(x)\quad \forall\ x \in S_X \text{ and } y \in S_Y$$

$$\mathcal{P}\left(X = x,\ Y = y\right) = \mathcal{P}\left(X = x\right)\mathcal{P}\left(Y = y\right)\ \forall\ x \in S_X \text{ and } y \in S_Y$$
$$p_{X|Y}(x \mid y) = p_X(x)\ \forall\ x \in S_X, y \in S_Y \text{ and } p_Y(y) > 0$$
$$p_{Y|X}(y \mid x) = p_Y(y)\ \forall\ x \in S_X, y \in S_Y \text{ and } p_X(x) > 0$$

  - For continuous case

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)\quad \forall\ x, y \in \mathbb{R}$$

$$f_{X|Y}\left(x \mid y\right) = f_X\left(x\right)\qquad \forall\ x, y \in \mathbb{R} \text{ and } f_Y\left(y\right) > 0$$
$$f_{Y|X}\left(y \mid x\right) = f_Y\left(y\right)\qquad \forall\ x, y \in \mathbb{R} \text{ and } f_X(x) > 0$$

- Probability in terms of joint density

$$\mathcal{P}_{X,Y}\left(B\right) = \mathcal{P}\left((X, Y) \in B\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_B f_{X,Y}(x, y)\,dx\,dy,\ B \in \mathcal{B}\left(\mathbb{R}\right) \times \mathcal{B}\left(\mathbb{R}\right)$$

where I_B is indicator function of B

- CDF in continuous case

$$F_{X,Y}\left(x, y\right) = \mathcal{P}\left((X, Y) \in (-\infty, x] \times (-\infty, y]\right) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(x, y)\,dx\,dy$$

- Marginal in continuous case

$$f_X\left(x\right) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dy \quad \text{and} \quad f_Y\left(y\right) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dx$$

- Conditional in continuous case
  - Conditional PDF of X given Y is given by

$$f_{X|Y}\left(x \mid y\right) = \frac{f_{X,Y}(x, y)}{f_Y\left(y\right)},\ f_Y\left(y\right) > 0 \text{ and}$$

  - Conditional PDF of Y given X is given by

$$f_{Y|X}\left(y \mid x\right) = \frac{f_{X,Y}(x, y)}{f_X\left(x\right)},\ f_X\left(x\right) > 0$$

6.7 TERMINAL QUESTIONS

1. Joint PMF of a discrete bivariate random variable (X, Y) is shown in Table 6.14. Find joint and marginal CDF's.

Table 6.14: Joint PMF of discrete bivariate random variable (X, Y)

	Values of Y					
Values of X	**1**	**2**	**3**	**4**	**5**	**6**
1	0.03	0.02	0.05	0.02	0.07	0.04
2	0.01	0.06	0.03	0.08	0.01	0.02
3	0.03	0.04	0.07	0.02	0.05	0.06
4	0.01	0.03	0.01	0.04	0.01	0.02
5	0.04	0.05	0.02	0.02	0.03	0.01

2. Some entries of the CDF of a discrete bivariate random variable (X, Y) have lost. Available entries are given in Table 6.15 as follows.

Table 6.15: Joint CDF of discrete bivariate random variable (X, Y)

	Values of Y					
Values of X	**1**	**2**	**3**	**4**	**5**	**6**
1	NA	0.05	NA	0.12	0.19	NA
2	0.04	NA	0.20	NA	0.38	NA
3	NA	0.19	NA	0.46	NA	0.71
4	0.08	0.23	0.39	NA	0.69	NA
5	NA	0.32	NA	0.68	NA	1

Find the required probability in each of the following parts.

(a) $\mathcal{P}(2 < X \le 4,\ 3 < Y \le 5)$ (b) $\mathcal{P}(3 < X \le 5,\ 2 < Y \le 6)$

(c) $\mathcal{P}(1 < X \le 4,\ 2 < Y \le 5)$

6.8 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. (a) We know that sum of all probabilities is 1. So, we must have

$$\sum_{x=0}^{2}\sum_{y=1}^{4} p_{X,\,Y}(x,\ y) = 1 \Rightarrow k\sum_{x=0}^{2}\sum_{y=1}^{4}(2x+3y) = 1 \Rightarrow k\sum_{x=0}^{2}\left(\sum_{y=1}^{4}2x + \sum_{y=1}^{4}3y\right) = 1$$

$$\Rightarrow k\sum_{x=0}^{2}\left(4(2x) + 3(1+2+3+4)\right) = 1 \Rightarrow k\sum_{x=0}^{2}\left(8x+30\right) = 1$$

$$\Rightarrow k\left(\sum_{x=0}^{2}8x + \sum_{x=0}^{2}30\right) = 1 \Rightarrow k\left(8(0+1+2)+3(30)\right) = 1 \Rightarrow k = \frac{1}{114}$$

(b) Joint PMF of the random variable (X, Y) in formula form is given by

$$p_{X,\,Y}(x,\ y) = \begin{cases} \dfrac{(2x+3y)}{114}, & x = 0,\ 1,\ 2;\ y = 1,\ 2,\ 3,\ 4 \\ 0, & \text{otherwise} \end{cases}$$

Joint PMF in tabular form with marginal sums is given by Table 6.16.

Table 6.16: Joint PMF of discrete bivariate random variable (X, Y)

	Values of Y				
Values of X	**1**	**2**	**3**	**4**	**Marginal sum**
0	$\frac{3}{114}$	$\frac{6}{114}$	$\frac{9}{114}$	$\frac{12}{114}$	$\mathbf{\frac{30}{114}}$
1	$\frac{5}{114}$	$\frac{8}{114}$	$\frac{11}{114}$	$\frac{14}{114}$	$\mathbf{\frac{38}{114}}$
2	$\frac{7}{114}$	$\frac{10}{114}$	$\frac{13}{114}$	$\frac{16}{114}$	$\mathbf{\frac{46}{114}}$
Marginal sum	$\mathbf{\frac{15}{114}}$	$\mathbf{\frac{24}{114}}$	$\mathbf{\frac{33}{114}}$	$\mathbf{\frac{42}{114}}$	$\mathbf{\frac{114}{114}} = 1$

Using marginal sums from Table 6.16, marginal PMF of X is given in Table 6.17.

Table 6.17: Marginal PMF of the random variable X

Values of X	0	1	2	Total
Probabilities ($p_X(x)$)	$\frac{30}{114}$	$\frac{38}{114}$	$\frac{46}{114}$	$\frac{114}{114} = 1$

Similarly, marginal PMF of Y is given in Table 6.18.

Table 6.18: Marginal PMF of the random variable Y

Values of Y	1	2	3	4	Total
Probabilities ($p_Y(y)$)	$\frac{15}{114}$	$\frac{24}{114}$	$\frac{33}{114}$	$\frac{42}{114}$	$\frac{114}{114} = 1$

(c) Joint CDF of the random variable (X, Y) is given by Table 6.19, where probability in each cell is obtained using (6.10) repeatedly.

Table 6.19: Joint CDF of discrete bivariate random variable (X, Y)

	Values of Y			
Values of X	**1**	**2**	**3**	**4**
0	$\frac{3}{114}$	$\frac{9}{114}$	$\frac{18}{114}$	$\frac{30}{114}$
1	$\frac{8}{114}$	$\frac{22}{114}$	$\frac{42}{114}$	$\frac{68}{114}$
2	$\frac{15}{114}$	$\frac{39}{114}$	$\frac{72}{114}$	$\frac{114}{114} = 1$

(d) Marginal CDF of the random variable X is given by just writing first and fifth columns of Table 6.19 together in a new Table 6.20.

Table 6.20: Marginal CDF of discrete random variable X

Values of X	0	1	2
$F_X(x)$	$\frac{30}{114}$	$\frac{68}{114}$	1

Similarly, marginal CDF of the random variable Y is given by just writing second and fifth rows of Table 6.19 together in a new Table 6.21.

Table 6.21: Marginal CDF of discrete random variable Y

Values of X	1	2	3	4
$F_X(x)$	$\frac{15}{114}$	$\frac{39}{114}$	$\frac{72}{114}$	1

(e) We can find conditional PMF of the random variable X given Y = 3 as follows.

$$p_{X|Y}(x \mid Y=3) = \frac{p(X=x,\ Y=3)}{p_Y(Y=3)},\ \ x = 0, 1, 2$$

$$= \begin{cases} 9/33, \quad x=0 \left[\because p_{X|Y}(X=0 \mid Y=3) = \dfrac{p(X=0,\ Y=3)}{p_Y(Y=3)} = \dfrac{9/114}{33/114} = \dfrac{9}{33}\right] \\ 11/33,\ x=1 \left[\because p_{X|Y}(X=1 \mid Y=3) = \dfrac{p(X=1,\ Y=3)}{p_Y(Y=3)} = \dfrac{11/114}{33/114} = \dfrac{11}{33}\right] \\ 13/33,\ x=2 \left[\because p_{X|Y}(X=2 \mid Y=3) = \dfrac{p(X=2,\ Y=3)}{p_Y(Y=3)} = \dfrac{13/114}{33/114} = \dfrac{13}{33}\right] \end{cases}$$

(f) From Table 6.16, we have

$$\mathcal{P}(X=0,\ Y=1) = \frac{3}{114} \qquad \text{... (6.108)}$$

$$\mathcal{P}(X=0)\mathcal{P}(Y=1) = \frac{30}{114} \times \frac{15}{114} = \frac{75}{114 \times 19} \qquad \text{... (6.109)}$$

From (6.108) and (6.109) it is clear that
$\mathcal{P}(X=0,\ Y=1) \neq \mathcal{P}(X=0)\mathcal{P}(Y=1)$

Since (6.72) does not hold so random variables X and Y are not independent.

(g) Using (6.63), we have

$$\mathcal{P}(a < X \le b,\ c < Y \le d) = F_{X,Y}(b, d) - F_{X,Y}(a, d) - F_{X,Y}(b, c) + F_{X,Y}(a, c)$$

In our case a = 0, b = 2, c = 1, d = 3, so, we have

$$\mathcal{P}(0 < X \le 2,\ 1 < Y \le 3) = F_{X,Y}(2, 3) - F_{X,Y}(0, 3) - F_{X,Y}(2, 1) + F_{X,Y}(0, 1)$$

$$= \frac{72}{114} - \frac{18}{114} - \frac{15}{114} + \frac{3}{114} = \frac{42}{114} \left[\begin{matrix}\text{Using probabilities} \\ \text{from cells of Table 6.19}\end{matrix}\right]$$

2. Given CDF of the jointly continuous random variable (X, Y) is

$$F_{X,Y}(x, y) = \begin{cases} 0, & \text{if } x < 0 \text{ or } y < 0 \\ \dfrac{(8x - x^2)(10y - y^2)}{400}, & 0 \le x \le 4,\ 0 \le y \le 5 \\ \dfrac{8x - x^2}{16}, & 0 \le x \le 4,\ y \ge 5 \\ \dfrac{10y - y^2}{25}, & x \ge 4,\ 0 \le y \le 5 \\ 1, & \text{if } x \ge 4 \text{ and } y \ge 5 \end{cases} \qquad \text{… (6.110)}$$

In view of (6.107), to obtain joint PDF of the jointly continuous random variable (X, Y), we have to differentiate (6.110) partially first w. r. t. y and then w. r. t. x or first w. r. t. x and then w. r. t. y. So, let us first differentiate (6.110) partially w. r. t y., we have

$$\frac{\partial}{\partial y}\left(F_{X,Y}(x, y)\right) = \begin{cases} 0, & \text{if } x < 0 \text{ or } y < 0 \\ \dfrac{\left(8x - x^2\right)\left(10 - 2y\right)}{400}, & 0 \le x \le 4,\ 0 \le y \le 5 \\ 0, & 0 \le x \le 4,\ y \ge 5 \\ \dfrac{10 - 2y}{25}, & x \ge 4,\ 0 \le y \le 5 \\ 0, & \text{if } x \ge 4 \text{ and } y \ge 5 \end{cases} \quad \ldots (6.111)$$

Now, differentiating (6.111) partially w. r. t. x, we get

$$\frac{\partial^2}{\partial x \partial y}\left(F_{X,Y}(x, y)\right) = \begin{cases} 0, & \text{if } x < 0 \text{ or } y < 0 \\ \dfrac{\left(8 - 2x\right)\left(10 - 2y\right)}{400}, & 0 \le x \le 4,\ 0 \le y \le 5 \\ 0, & 0 \le x \le 4,\ y \ge 5 \\ 0, & x \ge 4,\ 0 \le y \le 5 \\ 0, & \text{if } x \ge 4 \text{ and } y \ge 5 \end{cases}$$

or $$\frac{\partial^2}{\partial x \partial y}\left(F_{X,Y}(x, y)\right) = \begin{cases} \dfrac{\left(4 - x\right)\left(5 - y\right)}{100}, & 0 \le x \le 4,\ 0 \le y \le 5 \\ 0, & \text{otherwise} \end{cases}$$

Hence, joint PDF of jointly continuous random variable (X, Y) is given by

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y}\left(F_{X,Y}(x, y)\right) = \begin{cases} \dfrac{\left(4 - x\right)\left(5 - y\right)}{100}, & 0 \le x \le 4,\ 0 \le y \le 5 \\ 0, & \text{otherwise} \end{cases}$$

This matches with (6.98) as expected where k = 1/100 refer (6.99).

3. Before obtaining required probability first, we have to look at the region formed by $X + Y \le 2$ which is shown in Fig. 6.8 (b) in blue colour. The region of integration is triangular in shape. So, limit of only one variable will be constant and limits of other variable will be in terms of another variable. Such things we have discussed in detail in Unit 9 of the course MST-011. Let us keep limits of random variable X as constant and limit of random variable Y in terms of X. To do so, we have to draw a vertical strip PQ (say). Strip PQ starts from y = 0 and ends at

$$y = 2 - x \left[\begin{array}{l} \because \text{Equation of line AB using intercept form is } \dfrac{x}{2} + \dfrac{y}{2} = 1 \\ \text{You may refer (6.18) in Unit 6 of the course MST-011} \end{array}\right]$$

Now, required probability is given by

$$\mathcal{P}\left(X + Y \le 2\right) = \int_0^2 \int_0^{2-x} f_{X,Y}(x, y)\,dy\,dx = \int_0^2 \int_0^{2-x} \frac{(4 - x)(5 - y)}{100}\,dy\,dx$$

$$= \int_0^2 \frac{(4 - x)}{100}\left[5y - \frac{y^2}{2}\right]_0^{2-x} dx = \int_0^2 \frac{(4 - x)}{100}\left[10 - 5x - \frac{(2 - x)^2}{2}\right] dx$$

$$= \int_0^2 \frac{(4 - x)}{100}\left[\frac{20 - 10x - 4 - x^2 + 4x}{2}\right] dx = \int_0^2 \frac{(4 - x)}{100}\left[\frac{16 - 6x - x^2}{2}\right] dx$$

$$= \int_0^2 \left[\frac{64 - 24x - 4x^2 - 16x + 6x^2 + x^3}{200} \right] dx = \int_0^2 \left[\frac{x^3 + 2x^2 - 40x + 64}{200} \right] dx$$

$$= \frac{1}{200} \left[\frac{x^4}{4} + \frac{2x^3}{3} - 20x^2 + 64x \right]_0^2 = \frac{1}{200} \left[4 + \frac{16}{3} - 80 + 128 \right] = \frac{43}{150}$$

Terminal Questions

1. Using (6.10) repeatedly on values of Table 6.14 joint CDF of the discrete bivariate random variable (X, Y) is given by Table 6.22 as follows.

Table 6.22: Joint CDF of discrete bivariate random variable (X, Y)

	Values of Y					
Values of X	**1**	**2**	**3**	**4**	**5**	**6**
1	0.03	0.05	0.10	0.12	0.19	0.23
2	0.04	0.12	0.20	0.30	0.38	0.44
3	0.07	0.19	0.34	0.46	0.59	0.71
4	0.08	0.23	0.39	0.55	0.69	0.83
5	0.12	0.32	0.50	0.68	0.85	1

Marginal CDF of random variable X is given by combining values of X from the first column and corresponding probabilities from the seventh column of Table 6.22 and given by Table 6.23.

Table 6.23: Marginal CDF of the random variable X

Values of X	1	2	3	4	5
Probabilities ($p_X(x)$)	0.23	0.44	0.71	0.83	1

Marginal CDF of random variable Y is given by combining values of Y from the first row and corresponding probabilities from the sixth row of Table 6.22 and given by Table 6.23.

Table 6.23: Marginal CDF of the random variable Y

Values of Y	1	2	3	4	5	6
Probabilities ($p_Y(y)$)	0.12	0.32	0.50	0.68	0.85	1

2. Using (6.63), we have

$$\mathcal{P}\left(a < X \le b,\ c < Y \le d\right) = F_{X,Y}(b, d) - F_{X,Y}(a, d) - F_{X,Y}(b, c) + F_{X,Y}(a, c)$$

(a) In our case a = 2, b = 4, c = 3, d = 5, so, we have

$$\mathcal{P}\left(2 < X \le 4,\ 3 < Y \le 5\right) = F_{X,Y}(4, 5) - F_{X,Y}(2, 5) - F_{X,Y}(4, 3) + F_{X,Y}(2, 3)$$
$$= 0.69 - 0.38 - 0.39 + 0.20 = 0.12$$

(b) Similarly, in this case a = 3, b = 5, c = 2, d = 6, so, we have

$$\mathcal{P}\left(3 < X \le 5,\ 2 < Y \le 6\right) = F_{X,Y}(5, 6) - F_{X,Y}(3, 6) - F_{X,Y}(5, 2) + F_{X,Y}(3, 2)$$
$$= 1 - 0.71 - 0.32 + 0.19 = 0.16$$

(c) In this case a = 1, b = 4, c = 2, d = 5, so, we have

$$\mathcal{P}\left(1 < X \le 4,\ 2 < Y \le 5\right) = F_{X,Y}(4, 5) - F_{X,Y}(1, 5) - F_{X,Y}(4, 2) + F_{X,Y}(1, 2)$$
$$= 0.69 - 0.19 - 0.23 + 0.05 = 0.32$$