Pairs of Random Variables

- In this chapter, we consider experiments that produce a collection of random variables, $X_1, X_2, \ldots, X_n$, where n can be any integer.

- For most of this chapter, we study $n = 2$ random variables: X and Y. A pair of random variables is enough to show the important concepts and useful problem-solving techniques. Moreover, the definitions and theorems we introduce for X and Y generalize to n random variables. These generalized definitions appear near the end of this chapter in Section 5.10.

- We also note that a pair of random variables X and Y is the same as the two-dimensional vector $\begin{bmatrix} X & Y \end{bmatrix}'$. Similarly, the random variables $X_1, \ldots, X_n$ can be written as the n dimensional vector $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}'$. Since the components of $\mathbf{X}$ are random variables, $\mathbf{X}$ is called a *random vector*. Thus this chapter begins our study of random vectors.

- This subject is continued in Chapter 8, which uses techniques of linear algebra to develop further the properties of random vectors.

Pairs of Random Variables: Definitions

- We begin here with the definition of $F_{X,Y}(x, y)$, the *joint cumulative distribution function* of two random variables, a generalization of the CDF introduced in Section 3.4 and again in Section 4.2.

- The joint CDF is a complete probability model for any experiment that produces two random variables. However, it not very useful for analyzing practical experiments.

- More useful models are $P_{X,Y}(x, y)$, the *joint probability mass function* for two discrete random variables, presented in Sections 5.2 and 5.3, and $f_{X,Y}(x, y)$, the *joint probability density function* of two continuous random variables, presented in Sections 5.4 and 5.5.

- Section 5.7 considers functions of two random variables and expectations, respectively.

- We extend the definition of independent events to define independent random variables.

- The subject of Section 5.9 is the special case in which X and Y are Gaussian.

Example 5.1

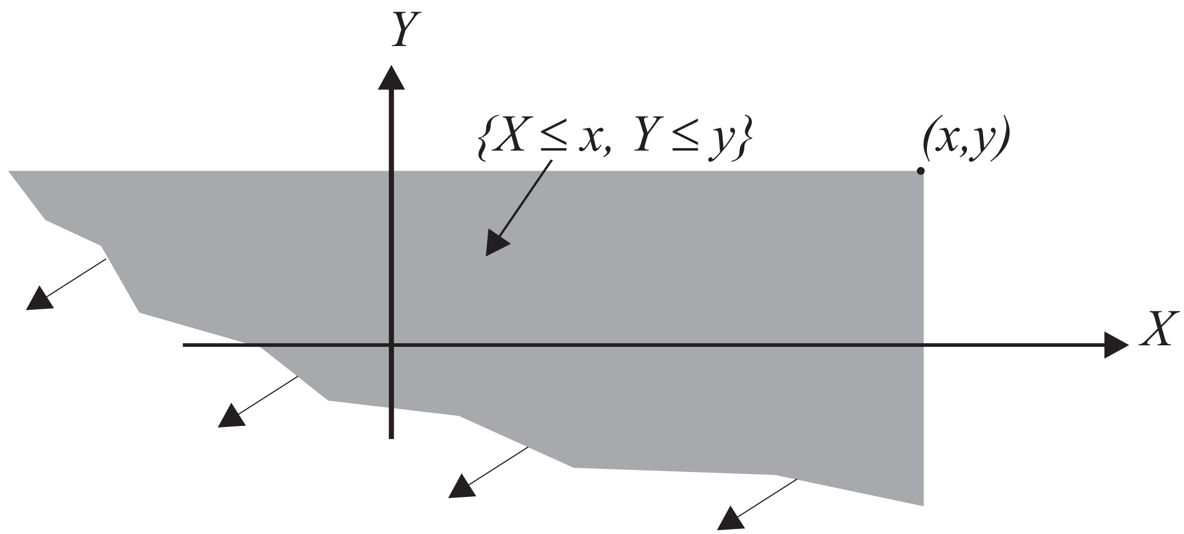
We would like to measure random variable X, but we instead observe

$$Y = X + Z. \tag{5.1}$$

The noise Z prevents us from perfectly observing X. In some settings, Z is an interfering signal. In the simplest setting, Z is just noise inside the circuitry of your measurement device that is unrelated to X. In this case, it is appropriate to assume that the signal and noise are independent; that is, the events $X = x$ and $Z = z$ are independent. This simple model produces three random variables, X, Y and Z, but any pair completely specifies the remaining random variable. Thus we will see that a probability model for the pair (X, Z) or for the pair (X, Y) will be sufficient to analyze experiments related to this system.

Joint Cumulative Distribution Function

Figure 5.1



The area of the (X, Y) plane corresponding to the joint cumulative distribution function $F_{X,Y}(x, y)$.

Definition 5.1 Joint Cumulative Distribution Function (CDF)

The joint cumulative distribution function *of random variables* X *and* Y *is*

$$F_{X,Y}(x,y) = \mathsf{P}\left[X \leq x, Y \leq y\right].$$

Theorem 5.1

For any pair of random variables, X, Y,

(a) $0 \leq F_{X,Y}(x, y) \leq 1$,
(b) $F_{X,Y}(\infty, \infty) = 1$,
(c) $F_X(x) = F_{X,Y}(x, \infty)$,
(d) $F_Y(y) = F_{X,Y}(\infty, y)$,
(e) $F_{X,Y}(x, -\infty) = 0$,
(f) $F_{X,Y}(-\infty, y) = 0$,
(g) If $x \leq x_1$ and $y \leq y_1$, then

$$F_{X,Y}(x, y) \leq F_{X,Y}(x_1, y_1)$$

Example 5.2 Problem

X years is the age of children entering first grade in a school. Y years is the age of children entering second grade. The joint CDF of X and Y is

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 5, \\ 0 & y < 6, \\ (x-5)(y-6) & 5 \leq x < 6, 6 \leq y < 7, \\ y-6 & x \geq 6, 6 \leq y < 7, \\ x-5 & 5 \leq x < 6, y \geq 7, \\ 1 & \text{otherwise.} \end{cases} \tag{5.3}$$

Find $F_X(x)$ and $F_Y(y)$.

Example 5.2 Solution

Using Theorem 5.1(b) and Theorem 5.1(c), we find

$$F_X(x) = \begin{cases} 0 & x < 5, \\ x - 5 & 5 \leq x < 6, \\ 1 & x \geq 6, \end{cases} \qquad F_Y(y) = \begin{cases} 0 & y < 6, \\ y - 6 & 6 \leq y < 7, \\ 1 & y \geq 7. \end{cases} \tag{5.4}$$

Referring to Theorem 4.6, we see from Equation (5.4) that X is a continuous uniform $(5, 6)$ random variable and Y is a continuous uniform $(6, 7)$ random variable.

Theorem 5.2

$$\mathsf{P}\left[x_1 < X \leq x_2, y_1 < Y \leq y_2\right] = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1).$$

Quiz 5.1

Express the following extreme values of the joint CDF $F_{X,Y}(x, y)$ as numbers or in terms of the CDFs $F_X(x)$ and $F_Y(y)$.

(a) $F_{X,Y}(-\infty, 2)$

(b) $F_{X,Y}(\infty, \infty)$

(c) $F_{X,Y}(\infty, y)$

(d) $F_{X,Y}(\infty, -\infty)$

Quiz 5.1 Solution

Each value of the joint CDF can be found by considering the corresponding probability.

(a) $F_{X,Y}(-\infty, 2) = \text{P}[X \leq -\infty, Y \leq 2] \leq \text{P}[X \leq -\infty] = 0$ since X cannot take on the value $-\infty$.

(b) $F_{X,Y}(\infty, \infty) = \text{P}[X \leq \infty, Y \leq \infty] = 1$.
This result is given in Theorem 5.1.

(c) $F_{X,Y}(\infty, y) = \text{P}[X \leq \infty, Y \leq y] = \text{P}[Y \leq y] = F_Y(y)$.

(d) $F_{X,Y}(\infty, -\infty) = \text{P}[X \leq \infty, Y \leq -\infty] = \text{P}[Y \leq -\infty] = 0$ since Y cannot take on the value $-\infty$.

Section 5.2

Joint Probability Mass Function

Definition 5.2 Joint Probability Mass Function (PMF)

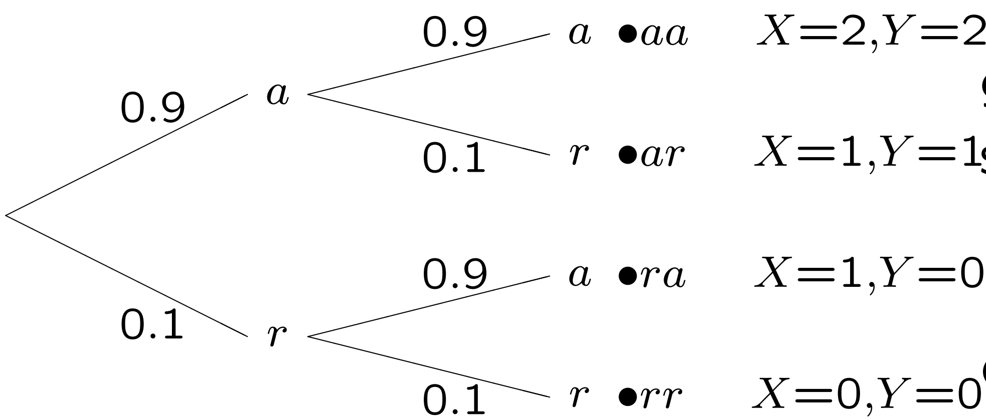
The joint probability mass function *of discrete random variables* X *and* Y *is*

$$P_{X,Y}(x, y) = \mathsf{P}\left[X = x, Y = y\right].$$

Example 5.3 Problem

Test two integrated circuits one after the other. On each test, the possible outcomes are a (accept) and r (reject). Assume that all circuits are acceptable with probability 0.9 and that the outcomes of successive tests are independent. Count the number of acceptable circuits X and count the number of successful tests Y before you observe the first reject. (If both tests are successful, let $Y = 2$.) Draw a tree diagram for the experiment and find the joint PMF $P_{X,Y}(x, y)$.

Example 5.3 Solution



The experiment has the tree diagram shown to the left. The sample space of the experiment is

$$S = \{aa, ar, ra, rr\}. \tag{5.7}$$

Observing the tree diagram, we compute

$$\mathrm{P}[aa] = 0.81, \qquad \mathrm{P}[ar] = 0.09, \tag{5.8}$$

$$\mathrm{P}[ra] = 0.09, \qquad \mathrm{P}[rr] = 0.01. \tag{5.9}$$

Each outcome specifies a pair of values X and Y. Let $g(s)$ be the function that transforms each outcome s in the sample space S into the pair of random variables (X, Y). Then

$$g(aa) = (2, 2), \quad g(ar) = (1, 1), \quad g(ra) = (1, 0), \quad g(rr) = (0, 0). \tag{5.10}$$

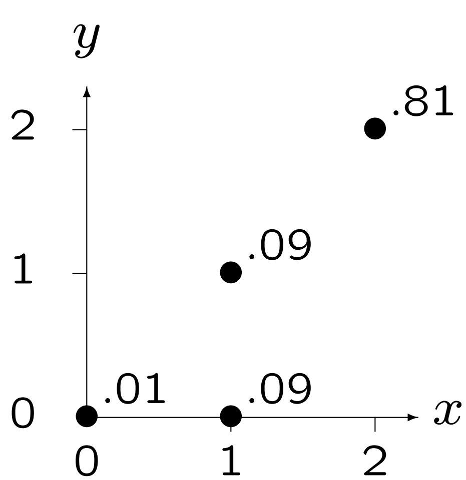
[Continued]

Example 5.3 Solution (Continued 2)

For each pair of values x, y, $P_{X,Y}(x,y)$ is the sum of the probabilities of the outcomes for which $X = x$ and $Y = y$. For example, $P_{X,Y}(1,1) = \mathrm{P}[ar]$.

$P_{X,Y}(x,y)$	$y = 0$	$y = 1$	$y = 2$
$x = 0$	0.01	0	0
$x = 1$	0.09	0.09	0
$x = 2$	0	0	0.81

The joint PMF can be represented by the table on left, or, as shown below, as a set of labeled points in the x, y plane where each point is a possible value of the pair (x, y), or as a simple list:



$$P_{X,Y}(x,y) = \begin{cases} 0.81 & x = 2, y = 2, \\ 0.09 & x = 1, y = 1, \\ 0.09 & x = 1, y = 0, \\ 0.01 & x = 0, y = 0. \\ 0 & \text{otherwise} \end{cases}$$

Theorem 5.3

For discrete random variables X and Y and any set B in the X, Y plane, the probability of the event $\{(X, Y) \in B\}$ is

$$\mathsf{P}[B] = \sum_{(x,y)\in B} P_{X,Y}(x, y).$$

Example 5.4 Problem

Continuing Example 5.3, find the probability of the event B that X, the number of acceptable circuits, equals Y, the number of tests before observing the first failure.

Example 5.4 Solution

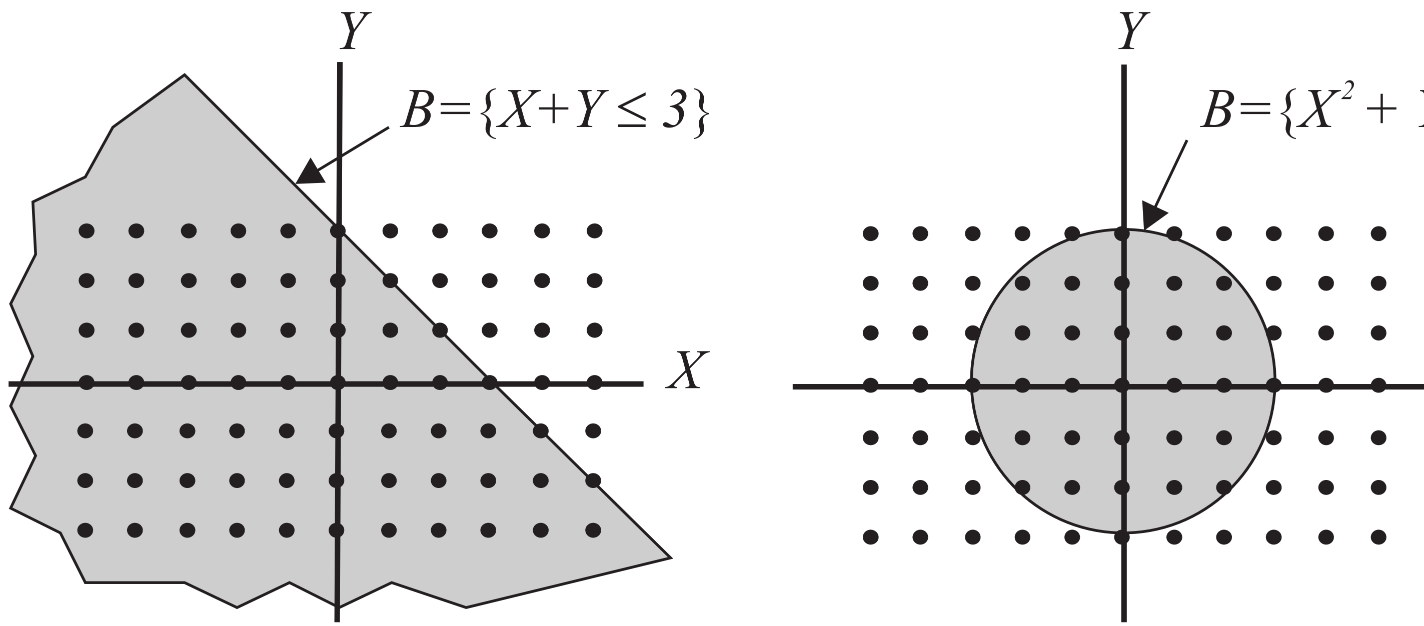
Mathematically, B is the event $\{X = Y\}$. The elements of B with nonzero probability are

$$B \cap S_{X,Y} = \{(0,0), (1,1), (2,2)\}. \tag{5.12}$$

Therefore,

$$\begin{aligned} \mathsf{P}[B] &= P_{X,Y}(0,0) + P_{X,Y}(1,1) + P_{X,Y}(2,2) \\ &= 0.01 + 0.09 + 0.81 = 0.91. \end{aligned} \tag{5.13}$$

Figure 5.2



Subsets B of the (X, Y) plane. Points $(X, Y) \in S_{X,Y}$ are marked by bullets.

Quiz 5.2

The joint PMF $P_{Q,G}(q,g)$ for random variables Q and G is given in the following table:

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

Calculate the following probabilities:

(a) $\text{P}[Q=0]$

(b) $\text{P}[Q=G]$

(c) $\text{P}[G>1]$

(d) $\text{P}[G>Q]$

Quiz 5.2 Solution

From the joint PMF of Q and G given in the table, we can calculate the requested probabilities by summing the PMF over those values of Q and G that correspond to the event.

(a) The probability that $Q = 0$ is

$$\begin{aligned} \mathrm{P}[Q=0] &= P_{Q,G}(0,0) + P_{Q,G}(0,1) + P_{Q,G}(0,2) + P_{Q,G}(0,3) \\ &= 0.06 + 0.18 + 0.24 + 0.12 = 0.6. \end{aligned} \tag{1}$$

(b) The probability that $Q = G$ is

$$\mathrm{P}[Q=G] = P_{Q,G}(0,0) + P_{Q,G}(1,1) = 0.18. \tag{2}$$

(c) The probability that $G > 1$ is

$$\begin{aligned} \mathrm{P}[G>1] &= \sum_{g=2}^{3} \sum_{q=0}^{1} P_{Q,G}(q,g) \\ &= 0.24 + 0.16 + 0.12 + 0.08 = 0.6. \end{aligned} \tag{3}$$

(d) The probability that $G > Q$ is

$$\begin{aligned} \mathrm{P}[G>Q] &= \sum_{q=0}^{1} \sum_{g=q+1}^{3} P_{Q,G}(q,g) \\ &= 0.18 + 0.24 + 0.12 + 0.16 + 0.08 = 0.78. \end{aligned} \tag{4}$$

Section 5.3

Marginal PMF

Theorem 5.4

For discrete random variables X and Y with joint PMF $P_{X,Y}(x,y)$,

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x,y), \qquad P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x,y).$$

Example 5.5 Problem

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81

In Example 5.3, we found that random variables X and Y have the joint PMF shown in this table. Find the marginal PMFs for the random variables X and Y.

Example 5.5 Solution

We note that both X and Y have range $\{0, 1, 2\}$. Theorem 5.4 gives

$$P_X(0) = \sum_{y=0}^{2} P_{X,Y}(0, y) = 0.01 \qquad P_X(1) = \sum_{y=0}^{2} P_{X,Y}(1, y) = 0.18 \tag{5.14}$$

$$P_X(2) = \sum_{y=0}^{2} P_{X,Y}(2, y) = 0.81 \qquad P_X(x) = 0 \quad x \neq 0, 1, 2 \tag{5.15}$$

Referring to the table representation of $P_{X,Y}(x, y)$, we observe that each value of $P_X(x)$ is the result of adding all the entries in one row of the table. Similarly, the formula for the PMF of Y in Theorem 5.4, $P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x, y)$, is the sum of all the entries in one column of the table.

[Continued]

Example 5.5 Solution (Continued 2)

We display $P_X(x)$ and $P_Y(y)$ by rewriting the table and placing the row sums and column sums in the margins.

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	$P_X(x)$
$x=0$	0.01	0	0	0.01
$x=1$	0.09	0.09	0	0.18
$x=2$	0	0	0.81	0.81
$P_Y(y)$	0.10	0.09	0.81	

Thus the column in the right margin shows $P_X(x)$ and the row in the bottom margin shows $P_Y(y)$. Note that the sum of all the entries in the bottom margin is 1 and so is the sum of all the entries in the right margin. This is simply a verification of Theorem 3.1(b), which states that the PMF of any random variable must sum to 1.

Quiz 5.3

The probability mass function $P_{H,B}(h,b)$ for the two random variables H and B is given in the following table. Find the marginal PMFs $P_H(h)$ and $P_B(b)$.

$P_{H,B}(h,b)$	$b=0$	$b=2$	$b=4$
$h=-1$	0	0.4	0.2
$h=0$	0.1	0	0.1
$h=1$	0.1	0.1	0

(5.16)

Quiz 5.3 Solution

By Theorem 5.4, the marginal PMF of H is

$$P_H(h) = \sum_{b=0,2,4} P_{H,B}(h,b)\,. \tag{1}$$

For each value of h, this corresponds to calculating the row sum across the table of the joint PMF. Similarly, the marginal PMF of B is

$$P_B(b) = \sum_{h=-1}^{1} P_{H,B}(h,b)\,. \tag{2}$$

For each value of b, this corresponds to the column sum down the table of the joint PMF. The easiest way to calculate these marginal PMFs is to simply sum each row and column:

$P_{H,B}(h,b)$	$b=0$	$b=2$	$b=4$	$P_H(h)$
$h=-1$	0	0.4	0.2	0.6
$h=0$	0.1	0	0.1	0.2
$h=1$	0.1	0.1	0	0.2
$P_B(b)$	0.2	0.5	0.3	

Joint Probability Density Function

Definition 5.3 Joint Probability Density Function (PDF)

The joint PDF of the continuous random variables X and Y is a function $f_{X,Y}(x, y)$ with the property

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du.$$

Theorem 5.5

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x\, \partial y}$$

Example 5.6 Problem

Use the joint CDF for childrens' ages X and Y given in Example 5.2 to derive the joint PDF presented in Equation (5.5).

Example 5.6 Solution

Referring to Equation (5.3) for the joint CDF $F_{X,Y}(x, y)$, we must evaluate the partial derivative $\partial^2 F_{X,Y}(x, y)/\partial x\,\partial y$ for each of the six regions specified in Equation (5.3). However, $\partial^2 F_{X,Y}(x, y)/\partial x\,\partial y$ is nonzero only if $F_{X.Y}(x.y)$ is a function of both x and y. In this example, only the region $\{5 \leq x < 6, 6 \leq y < 7\}$ meets this requirement. Over this region,

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x\,\partial y}[(x-5)(y-6)] = \frac{\partial}{\partial x}[x-5]\frac{\partial}{\partial y}[y-6] = 1. \qquad (5.18)$$

Over all other regions, the joint PDF $f_{X,Y}(x, y)$ is zero.

Theorem 5.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.2):

(a) $f_{X,Y}(x, y) \geq 0$ for all (x, y),

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$.

Theorem 5.7

The probability that the continuous random variables (X, Y) are in A is

$$\mathrm{P}[A] = \iint_A f_{X,Y}(x, y)\, dx\, dy.$$

Example 5.7 Problem

Random variables X and Y have joint PDF

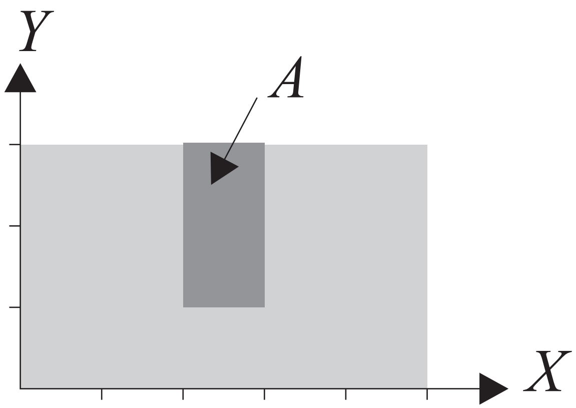
$$f_{X,Y}(x,y) = \begin{cases} c & 0 \leq x \leq 5, 0 \leq y \leq 3, \\ 0 & \text{otherwise.} \end{cases} \tag{5.19}$$

Find the constant c and $\text{P}[A] = \text{P}[2 \leq X < 3, 1 \leq Y < 3]$.

Example 5.7 Solution

The large rectangle in the diagram is the area of nonzero probability. Theorem 5.6 states that the integral of the joint PDF over this rectangle is 1:

$$1 = \int_0^5 \int_0^3 c\, dy\, dx = 15c. \qquad (5.20)$$

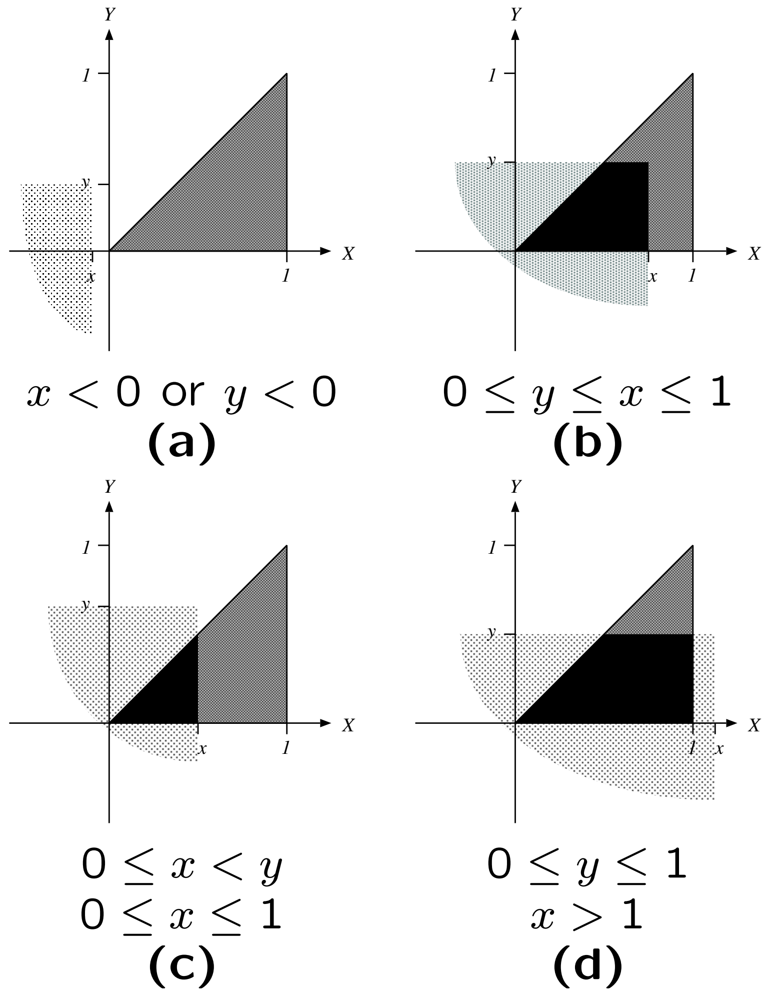


Therefore, $c = 1/15$. The small dark rectangle in the diagram is the event $A = \{2 \le X < 3, 1 \le Y < 3\}$. P$[A]$ is the integral of the PDF over this rectangle, which is

$$\mathrm{P}[A] = \int_2^3 \int_1^3 \frac{1}{15}\, dv\, du = 2/15. \qquad (5.21)$$

This probability model is an example of a pair of random variables uniformly distributed over a rectangle in the X, Y plane.

Figure 5.3

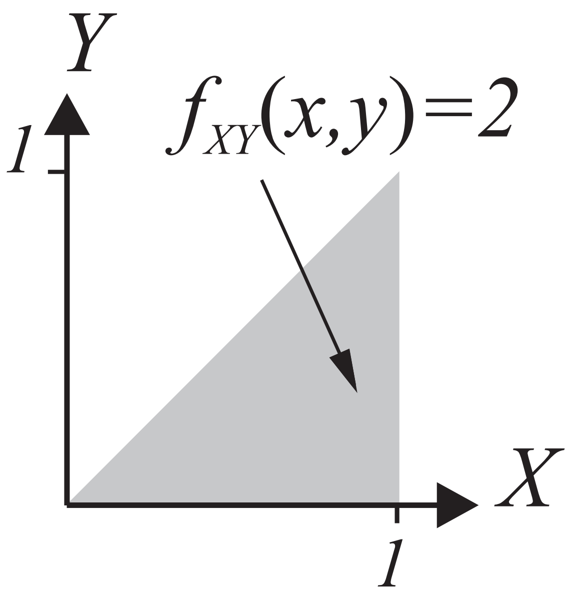


(Case (e) covering the whole triangle is omitted.)

Five cases for the CDF $F_{X,Y}(x, y)$ of Example 5.8.

Example 5.8 Problem

Find the joint CDF $F_{X,Y}(x, y)$ when X and Y have joint PDF



$$f_{X,Y}(x, y) = \begin{cases} 2 & 0 \leq y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.22}$$

Example 5.8 Solution

We can derive the joint CDF using Definition 5.3 in which we integrate the joint PDF $f_{X,Y}(x,y)$ over the area shown in Figure 5.1. To perform the integration it is extremely useful to draw a diagram that clearly shows the area with nonzero probability and then to use the diagram to derive the limits of the integral in Definition 5.3.

The difficulty with this integral is that the nature of the region of integration depends critically on x and y. In this apparently simple example, there are five cases to consider! The five cases are shown in Figure 5.3. First, we note that with $x < 0$ or $y < 0$, the triangle is completely outside the region of integration, as shown in Figure 5.3a. Thus we have $F_{X,Y}(x,y) = 0$ if either $x < 0$ or $y < 0$. Another simple case arises when $x \geq 1$ and $y \geq 1$. In this case, we see in Figure 5.3e that the triangle is completely inside the region of integration, and we infer from Theorem 5.6 that $F_{X,Y}(x,y) = 1$. The other cases we must consider are more complicated. In each case, since $f_{X,Y}(x,y) = 2$ over the triangular region, the value of the integral is two times the indicated area. When (x,y) is inside the area of nonzero probability (Figure 5.3b), the integral is

$$F_{X,Y}(x,y) = \int_0^y \int_v^x 2\,du\,dv = 2xy - y^2 \qquad \text{(Figure 5.3b).} \tag{5.23}$$

[Continued]

Example 5.8 Solution (Continued 2)

In Figure 5.3c, (x, y) is above the triangle, and the integral is

$$F_{X,Y}(x,y) = \int_0^x \int_v^x 2\,du\,dv = x^2 \qquad \text{(Figure 5.3c).} \tag{5.24}$$

The remaining situation to consider is shown in Figure 5.3d, when (x, y) is to the right of the triangle of nonzero probability, in which case the integral is

$$F_{X,Y}(x,y) = \int_0^y \int_v^1 2\,du\,dv = 2y - y^2 \qquad \text{(Figure 5.3d)} \tag{5.25}$$

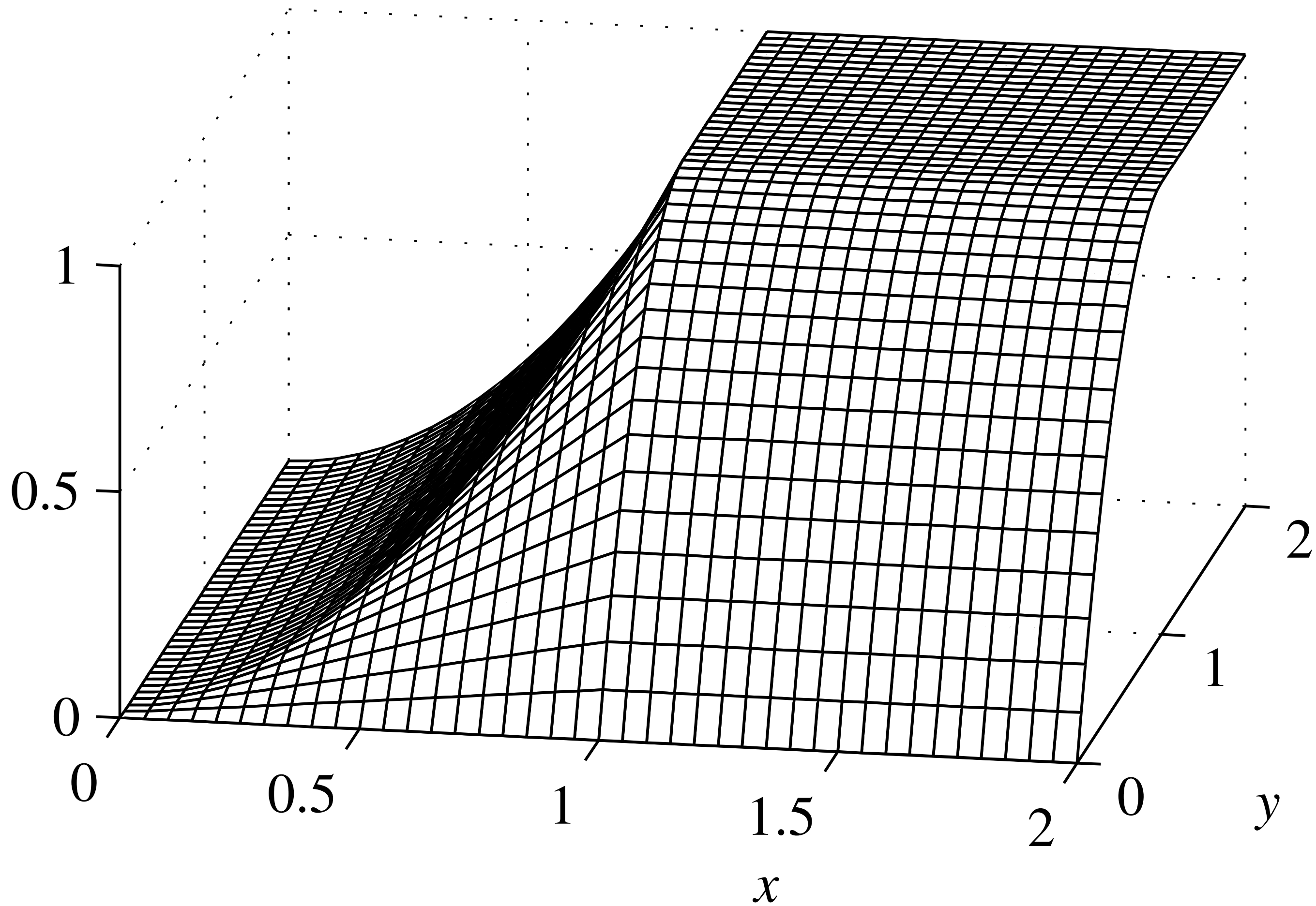
The resulting CDF, corresponding to the five cases of Figure 5.3, is

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 & \textbf{(a)}, \\ 2xy - y^2 & 0 \le y \le x \le 1 & \textbf{(b)}, \\ x^2 & 0 \le x < y, 0 \le x \le 1 & \textbf{(c)}, \\ 2y - y^2 & 0 \le y \le 1, x > 1 & \textbf{(d)}, \\ 1 & x > 1, y > 1 & \textbf{(e)}. \end{cases} \tag{5.26}$$

In Figure 5.4, the surface plot of $F_{X,Y}(x, y)$ shows that cases (a) through (e) correspond to contours on the "hill" that is $F_{X,Y}(x, y)$. In terms of visualizing the random variables, the surface plot of $F_{X,Y}(x, y)$ is less instructive than the simple triangle characterizing the PDF $f_{X,Y}(x, y)$.

Because the PDF in this example is $f_{X,Y}(x, y) = 2$ over $(x, y) \in S_{X,Y}$, each probability is just two times the area of the region shown in one of the diagrams (either a triangle or a trapezoid). You may want to apply some high school geometry to verify that the results obtained from the integrals are indeed twice the areas of the regions indicated. The approach taken in our solution, integrating over $S_{X,Y}$ to obtain the CDF, works for any PDF.

Figure 5.4



A graph of the joint CDF $F_{X,Y}(x, y)$ of Example 5.8.

Example 5.9 Problem

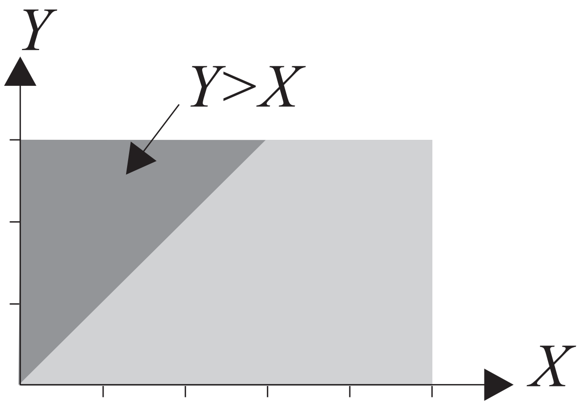
As in Example 5.7, random variables X and Y have joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 1/15 & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{5.27}$$

What is $P[A] = P[Y > X]$?

Example 5.9 Solution

Applying Theorem 5.7, we integrate $f_{X,Y}(x,y)$ over the part of the X,Y plane satisfying $Y > X$. In this case,



$$P[A] = \int_0^3 \left(\int_x^3 \frac{1}{15} \right) dy\,dx \tag{5.28}$$

$$= \int_0^3 \frac{3-x}{15}\,dx = -\left.\frac{(3-x)^2}{30}\right|_0^3 = \frac{3}{10}. \tag{5.29}$$

Quiz 5.4

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x,y) = \begin{cases} cxy & 0 \le x \le 1, 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{5.30}$$

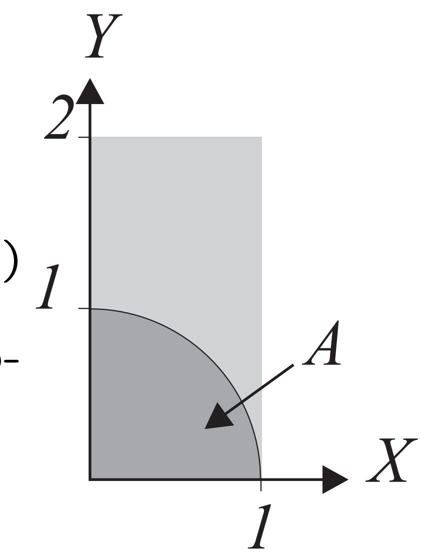
Find the constant c. What is the probability of the event $A = X^2 + Y^2 \le 1$?

Quiz 5.4 Solution

To find the constant c, we apply

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy$$
$$= \int_0^2 \int_0^1 cxy\, dx\, dy = c\int_0^2 y \left(\left.\frac{x^2}{2}\right|_0^1\right) dy = \frac{c}{2}\int_0^2 y\, dy = \left.\frac{cy^2}{4}\right|_0^2 = c. \tag{1}$$

Thus $c = 1$.



To calculate $\mathrm{P}[A]$, we write

$$\mathrm{P}[A] = \iint_A f_{X,Y}(x,y)\, dx\, dy \tag{2}$$

To integrate over A, we convert to polar coordinates using the substitutions $x = r\cos\theta$, $y = r\sin\theta$ and $dx\, dy = r\, dr\, d\theta$.

This yields

$$\mathrm{P}[A] = \int_0^{\pi/2}\int_0^1 r^2 \sin\theta\cos\theta\, r\, dr\, d\theta$$
$$= \left(\int_0^1 r^3\, dr\right)\int_0^{\pi/2} \sin\theta\cos\theta\, d\theta = \left(r^4/4\Big|_0^1\right)\left(\left.\frac{\sin^2\theta}{2}\right|_0^{\pi/2}\right) = \frac{1}{8}. \tag{3}$$

Marginal PDF

Theorem 5.8

If X and Y are random variables with joint PDF $f_{X,Y}(x,y)$,

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

Proof: Theorem 5.8

From the definition of the joint PDF, we can write

$$F_X(x) = \mathsf{P}\left[X \leq x\right] = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{X,Y}(u, y)\ dy \right) du. \tag{5.31}$$

Taking the derivative of both sides with respect to x (which involves differentiating an integral with variable limits), we obtain

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dy$$

. A similar argument holds for $f_Y(y)$.

Example 5.10 Problem

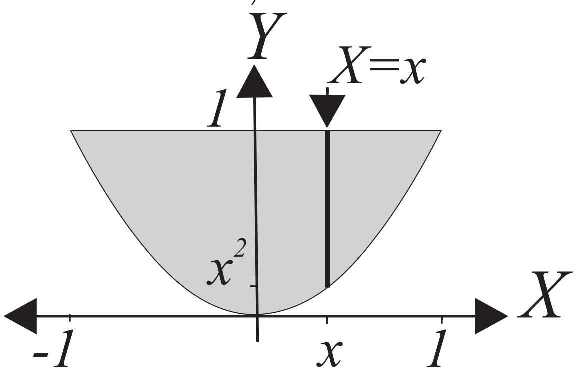
The joint PDF of X and Y is

$$f_{X,Y}(x,y) = \begin{cases} 5y/4 & -1 \leq x \leq 1, x^2 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.32}$$

Find the marginal PDFs $f_X(x)$ and $f_Y(y)$.

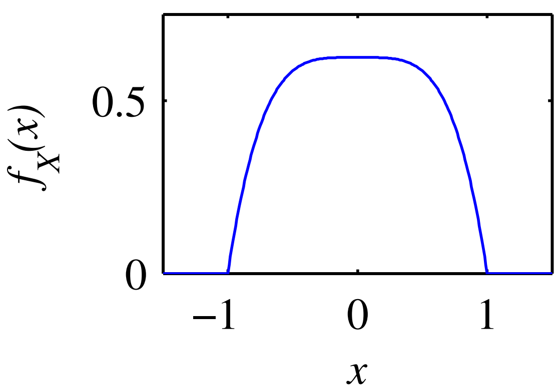
Example 5.10 Solution

We use Theorem 5.8 to find the marginal PDF $f_X(x)$. In the figure that accompanies Equation (5.33) below, the gray bowl-shaped region depicts those values of X and Y for which $f_{X,Y}(x,y) > 0$. When $x < -1$ or when $x > 1$, $f_{X,Y}(x,y) = 0$, and therefore $f_X(x) = 0$. For $-1 \le x \le 1$,



$$f_X(x) = \int_{x^2}^{1} \frac{5y}{4}\, dy = \frac{5(1-x^4)}{8}. \tag{5.33}$$

The complete expression for the marginal PDF of X is

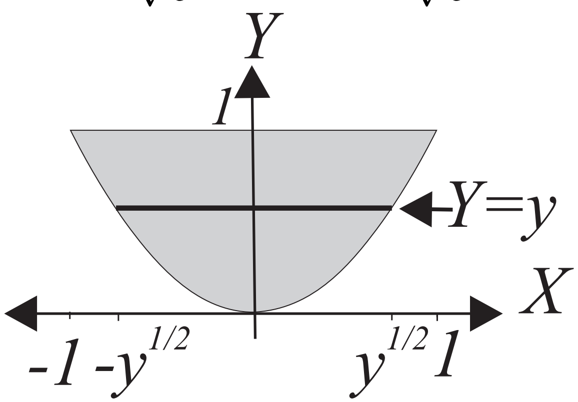


$$f_X(x) = \begin{cases} 5(1-x^4)/8 & -1 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.34}$$

[Continued]

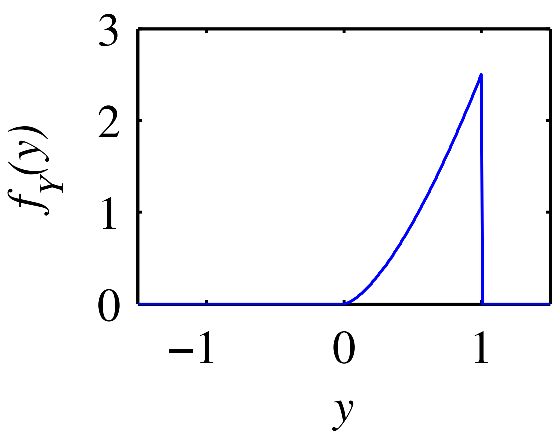
Example 5.10 Solution (Continued 2)

For the marginal PDF of Y, we note that for $y < 0$ or $y > 1$, $f_Y(y) = 0$. For $0 \le y \le 1$, we integrate over the horizontal bar marked $Y = y$. The boundaries of the bar are $x = -\sqrt{y}$ and $x = \sqrt{y}$. Therefore, for $0 \le y \le 1$,



$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{5y}{4}\,dx = \frac{5y}{4}x\Big|_{x=-\sqrt{y}}^{x=\sqrt{y}} = 5y^{3/2}/2. \qquad (5.35)$$

The complete marginal PDF of Y is



$$f_Y(y) = \begin{cases} (5/2)y^{3/2} & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases} \qquad (5.36)$$

Quiz 5.5

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x,y) = \begin{cases} 6(x+y^2)/5 & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.37}$$

Find $f_X(x)$ and $f_Y(y)$, the marginal PDFs of X and Y.

Quiz 5.5 Solution

By Theorem 5.8, the marginal PDF of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy \tag{1}$$

Note that $f_X(x) = 0$ for $x < 0$ or $x > 1$. For $0 \le x \le 1$,

$$f_X(x) = \frac{6}{5}\int_0^1 (x+y^2)\,dy = \frac{6}{5}\left(xy + y^3/3\right)\Big|_{y=0}^{y=1} = \frac{6x+2}{5} \tag{2}$$

The complete expression for the PDF of X is

$$f_X(x) = \begin{cases} (6x+2)/5 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

By the same method we obtain the marginal PDF for Y. For $0 \le y \le 1$,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy \\ &= \frac{6}{5}\int_0^1 (x+y^2)\,dx = \frac{6}{5}\left(\frac{x^2}{2} + xy^2\right)\Bigg|_{x=0}^{x=1} = \frac{6y^2+3}{5}. \end{aligned} \tag{4}$$

Since $f_Y(y) = 0$ for $y < 0$ or $y > 1$, the complete expression for the PDF of Y is

$$f_Y(y) = \begin{cases} (3+6y^2)/5 & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Section 5.6

Independent Random Variables

Definition 5.4 Independent Random Variables

Random variables X *and* Y *are* independent *if and only if*

$$\text{Discrete:} \qquad P_{X,Y}(x,y) = P_X(x)P_Y(y)$$

$$\text{Continuous:} \quad f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

Example 5.11 Problem

Are the childrens' ages X and Y in Example 5.2 independent?

Example 5.11 Solution

In Example 5.2, we derived the CDFs $F_X(x)$ and $F_Y(y)$, which showed that X is uniform $(5,6)$ and Y is uniform $(6,7)$. Thus X and Y have marginal PDFs

$$f_X(x) = \begin{cases} 1 & 5 \leq x < 6, \\ 0 & \text{otherwise}, \end{cases} \qquad f_Y(y) = \begin{cases} 1 & 6 \leq x < 7, \\ 0 & \text{otherwise}. \end{cases} \tag{5.38}$$

Referring to Equation (5.5), we observe that $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. Thus X and Y are independent.

Example 5.12 Problem

$$f_{X,Y}(x,y) = \begin{cases} 4xy & 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent?

Example 5.12 Solution

The marginal PDFs of X and Y are

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1, \\ 0 & \text{otherwise}, \end{cases} \qquad f_Y(y) = \begin{cases} 2y & 0 \leq y \leq 1, \\ 0 & \text{otherwise}. \end{cases} \tag{5.39}$$

It is easily verified that $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ for all pairs (x,y), and so we conclude that X and Y are independent.

Example 5.13 Problem

$$f_{U,V}(u,v) = \begin{cases} 24uv & u \geq 0, v \geq 0, u+v \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.40}$$

Are U and V independent?

Example 5.13 Solution

Since $f_{U,V}(u,v)$ looks similar in form to $f_{X,Y}(x,y)$ in the previous example, we might suppose that U and V can also be factored into marginal PDFs $f_U(u)$ and $f_V(v)$. However, this is not the case. Owing to the triangular shape of the region of nonzero probability, the marginal PDFs are

$$f_U(u) = \begin{cases} 12u(1-u)^2 & 0 \leq u \leq 1, \\ 0 & \text{otherwise,} \end{cases} \quad f_V(v) = \begin{cases} 12v(1-v)^2 & 0 \leq v \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, U and V are not independent. Learning U changes our knowledge of V. For example, learning $U = 1/2$ informs us that $\mathrm{P}[V \leq 1/2] = 1$.

Example 5.14 Problem

Consider again the noisy observation model of Example 5.1. Suppose X is a Gaussian $(0, \sigma_X)$ information signal sent by a radio transmitter and $Y = X + Z$ is the output of a low-noise amplifier attached to the antenna of a radio receiver. The noise Z is a Gaussian $(0, \sigma_Z)$ random variable that is generated within the receiver. What is the joint PDF $f_{X,Z}(x, z)$?

Example 5.14 Solution

From the information given, we know that X and Z have PDFs

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-x^2/2\sigma_X^2}, \qquad f_Z(z) = \frac{1}{\sqrt{2\pi\sigma_Z^2}} e^{-z^2/2\sigma_Z^2}. \tag{5.41}$$

The signal X depends on the information being transmitted by the sender and the noise Z depends on electrons bouncing around in the receiver circuitry. As there is no reason for these to be related, we model X and Z as independent. Thus, the joint PDF is

$$f_{X,Z}(x,z) = f_X(x)\, f_Z(z) = \frac{1}{2\pi\sqrt{\sigma_X^2\sigma_Z^2}} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_X^2}+\frac{z^2}{\sigma_Z^2}\right)}. \tag{5.42}$$

Quiz 5.6(A)

Random variables X and Y in Example 5.3 and random variables Q and G in Quiz 5.2 have joint PMFs:

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

(a) Are X and Y independent?

(b) Are Q and G independent?

Quiz 5.6(A) Solution

(a) For random variables X and Y from Example 5.3, we observe that $P_Y(1) = 0.09$ and $P_X(0) = 0.01$. However,

$$P_{X,Y}(0,1) = 0 \neq P_X(0)\, P_Y(1) \tag{1}$$

Since we have found a pair x, y such that $P_{X,Y}(x,y) \neq P_X(x)P_Y(y)$, we can conclude that X and Y are dependent. Note that whenever $P_{X,Y}(x,y) = 0$, independence requires that either $P_X(x) = 0$ or $P_Y(y) = 0$.

(b) For random variables Q and G from Quiz 5.2, it is not obvious whether they are independent. Unlike X and Y in part (a), there are no obvious pairs q, g that fail the independence requirement. In this case, we calculate the marginal PMFs from the table of the joint PMF $P_{Q,G}(q,g)$ in Quiz 5.2. In transposed form, this table is

$P_{Q,G}(q,g)$	$q=0$	$q=1$	$P_G(g)$
$g=0$	0.06	0.04	0.10
$g=1$	0.18	0.12	0.30
$g=2$	0.24	0.16	0.40
$g=3$	0.12	0.08	0.20
$P_Q(q)$	0.60	0.40	

Careful study of the table will verify that $P_{Q,G}(q,g) = P_Q(q)P_G(g)$ for every pair q, g. Hence Q and G are independent.

Quiz 5.6(B)

Random variables X_1 and X_2 are independentand identically distributed with probability density function

$$f_X(x) = \begin{cases} x/2 & 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{5.43}$$

What is the joint PDF $f_{X_1,X_2}(x_1, x_2)$?

Quiz 5.6(B) Solution

Since X_1 and X_2 are identical, $f_{X_1}(x) = f_{X_2}(x) = f_X(x)$. Since X_1 and X_2 are independent,

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2) = \begin{cases} \frac{x_1}{2} \cdot \frac{x_2}{2} & 0 \leq x_1, x_2 \leq 2, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Section 5.7

Expected Value of a Function of Two Random Variables

Theorem 5.9

For random variables X and Y, the expected value of $W = g(X, Y)$ is

Discrete: $$\mathrm{E}[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y}(x, y)$$

Continuous: $$\mathrm{E}[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\, dx\, dy.$$

Theorem 5.10

$$\mathsf{E}\left[a_1 g_1(X,Y) + \cdots + a_n g_n(X,Y)\right] = a_1\,\mathsf{E}\left[g_1(X,Y)\right] + \cdots + a_n\,\mathsf{E}\left[g_n(X,Y)\right].$$

Proof: Theorem 5.10

Let $g(X,Y) = a_1 g_1(X,Y) + \cdots + a_n g_n(X,Y)$. For discrete random variables X, Y, Theorem 5.9 states

$$\mathsf{E}[g(X,Y)] = \sum_{x \in S_X} \sum_{y \in S_Y} \left(a_1 g_1(x,y) + \cdots + a_n g_n(x,y)\right) P_{X,Y}(x,y). \tag{5.44}$$

We can break the double summation into n weighted double summations:

$$\mathsf{E}[g(X,Y)] = a_1 \sum_{x \in S_X} \sum_{y \in S_Y} g_1(x,y) P_{X,Y}(x,y) + \cdots + a_n \sum_{x \in S_X} \sum_{y \in S_Y} g_n(x,y) P_{X,Y}(x,y).$$

By Theorem 5.9, the ith double summation on the right side is $\mathsf{E}[g_i(X,Y)]$; thus,

$$\mathsf{E}[g(X,Y)] = a_1 \mathsf{E}[g_1(X,Y)] + \cdots + a_n \mathsf{E}[g_n(X,Y)]. \tag{5.45}$$

For continuous random variables, Theorem 5.9 says

$$\mathsf{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(a_1 g_1(x,y) + \cdots + a_n g_n(x,y)\right) f_{X,Y}(x,y)\, dx\, dy. \tag{5.46}$$

To complete the proof, we express this integral as the sum of n integrals and recognize that each of the new integrals is a weighted expected value, $a_i \mathsf{E}[g_i(X,Y)]$.

Theorem 5.11

For any two random variables X and Y,

$$\mathsf{E}[X+Y] = \mathsf{E}[X] + \mathsf{E}[Y].$$

Theorem 5.12

The variance of the sum of two random variables is

$$\mathsf{Var}\left[X+Y\right]=\mathsf{Var}\left[X\right]+\mathsf{Var}\left[Y\right]+2\,\mathsf{E}\left[(X-\mu_X)(Y-\mu_Y)\right].$$

Proof: Theorem 5.12

Since $\mathsf{E}[X+Y] = \mu_X + \mu_Y$,

$$\begin{aligned}
\mathsf{Var}[X+Y] &= \mathsf{E}\left[(X+Y-(\mu_X+\mu_Y))^2\right] \\
&= \mathsf{E}\left[((X-\mu_X)+(Y-\mu_Y))^2\right] \\
&= \mathsf{E}\left[(X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2\right]. \quad (5.47)
\end{aligned}$$

We observe that each of the three terms in the preceding expected values is a function of X and Y. Therefore, Theorem 5.10 implies

$$\mathsf{Var}[X+Y] = \mathsf{E}\left[(X-\mu_X)^2\right] + 2\,\mathsf{E}\left[(X-\mu_X)(Y-\mu_Y)\right] + \mathsf{E}\left[(Y-\mu_Y)^2\right]. \quad (5.48)$$

The first and last terms are, respectively, $\mathsf{Var}[X]$ and $\mathsf{Var}[Y]$.

Example 5.15 Problem

A company website has three pages. They require 750 kilobytes, 1500 kilobytes, and 2500 kilobytes for transmission. The transmission speed can be 5 Mb/s for external requests or 10 Mb/s for internal requests. Requests arrive randomly from inside and outside the company independently of page length, which is also random. The probability models for transmision speed, R, and page length, L, are:

$$P_R(r) = \begin{cases} 0.4 & r = 5, \\ 0.6 & r = 10, \\ 0 & \text{otherwise}, \end{cases} \qquad P_L(l) = \begin{cases} 0.3 & l = 750, \\ 0.5 & l = 1500, \\ 0.2 & l = 2500, \\ 0 & \text{otherwise}. \end{cases} \tag{5.49}$$

Write an expression for the transmission time $g(R, L)$ seconds. Derive the expected transmission time $\mathrm{E}[g(R, L)]$. Does $\mathrm{E}[(g(R, L)] = g(\mathrm{E}[R], \mathrm{E}[L])$?

Example 5.15 Solution

The transmission time T seconds is the the page length (in kb) divided by the transmission speed (in kb/s), or $T = 8L/1000R$. Because R and L are independent, $P_{R,L}(r,l) = P_R(r)P_L(l)$ and

$$\begin{aligned}
\mathrm{E}[g(R,L)] &= \sum_r \sum_l P_R(r)\, P_L(l)\, \frac{8l}{1000r} \\
&= \frac{8}{1000} \left(\sum_r \frac{P_R(r)}{r} \right) \left(\sum_l P_L(l)\, l \right) \\
&= \frac{8}{1000} \left(\frac{0.4}{5} + \frac{0.6}{10} \right) (0.3(750) + 0.5(1500) + 0.2(2500)) \\
&= 1.652 \text{ s.}
\end{aligned} \tag{5.50}$$

By comparison, $\mathrm{E}[R] = \sum_r r P_R(r) = 8$ Mb/s and $\mathrm{E}[L] = \sum_l l P_L(l) = 1475$ kilobytes. This implies

$$g(\mathrm{E}[R], \mathrm{E}[L]) = \frac{8\,\mathrm{E}[L]}{1000\,\mathrm{E}[R]} = 1.475 \text{ s} \neq \mathrm{E}[g(R,L)]. \tag{5.51}$$

Section 5.8

Covariance, Correlation and Independence

Definition 5.5 Covariance

The covariance *of two random variables* X *and* Y *is*

$$\mathsf{Cov}\left[X, Y\right] = \mathsf{E}\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right].$$

Example 5.16

Suppose we perform an experiment in which we measure X and Y in centimeters (for example the height of two sisters). However, if we change units and measure height in meters, we will perform the same experiment except we observe $\hat{X} = X/100$ and $\hat{Y} = Y/100$. In this case, $\hat{X}$ and $\hat{Y}$ have expected values $\mu_{\hat{X}} = \mu_X/100$ m, $\mu_{\hat{Y}} = \mu_Y/100$ m and

$$\begin{aligned}\text{Cov}\left[\hat{X}, \hat{Y}\right] &= \text{E}\left[(\hat{X} - \mu_{\hat{X}})(\hat{Y} - \mu_{\hat{Y}})\right] \\ &= \frac{\text{E}\left[(X - \mu_X)(Y - \mu_Y)\right]}{10,000} = \frac{\text{Cov}\left[X, Y\right]}{10,000} \text{ m}^2. \end{aligned} \tag{5.53}$$

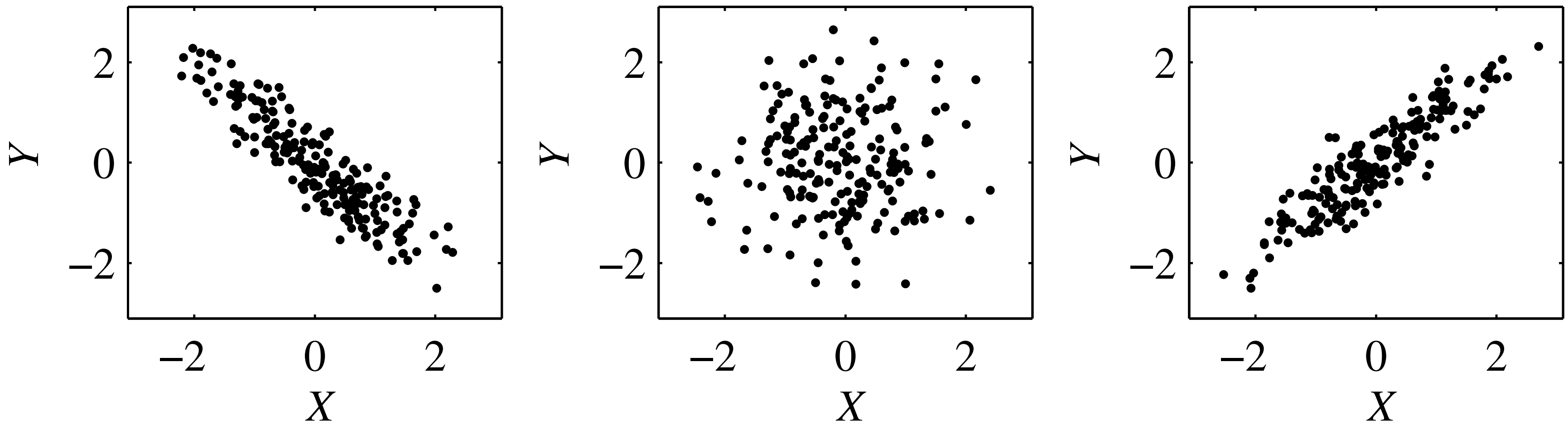
Changing the unit of measurement from cm^2 to m^2 reduces the covariance by a factor of $10,000$. However, the tendency of $X - \mu_X$ and $Y - \mu_Y$ to have the same sign is the same as the tendency of $\hat{X} - \mu_{\hat{X}}$ and $\hat{Y} - \mu_{\hat{Y}}$ to have the same sign. (Both are an indication of how likely it is that a girl is taller than average if her sister is taller than average).

Definition 5.6 Correlation Coefficient

The correlation coefficient *of two random variables* X *and* Y *is*

$$\rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X]\,\text{Var}[Y]}} = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y}.$$

Figure 5.5



(a) $\rho_{X,Y} = -0.9$ **(b)** $\rho_{X,Y} = 0$ **(c)** $\rho_{X,Y} = 0.9$

Each graph has 200 samples, each marked by a dot, of the random variable pair (X, Y) such that $\mathsf{E}[X] = \mathsf{E}[Y] = 0$, $\mathsf{Var}[X] = \mathsf{Var}[Y] = 1$.

Theorem 5.13

If $\hat{X} = aX + b$ and $\hat{Y} = cY + d$, then

(a) $\rho_{\hat{X},\hat{Y}} = \rho_{X,Y}$,

(b) $\text{Cov}[\hat{X}, \hat{Y}] = ac\,\text{Cov}[X, Y]$.

Theorem 5.14

$$-1 \leq \rho_{X,Y} \leq 1.$$

Proof: Theorem 5.14

Let σ_X^2 and σ_Y^2 denote the variances of X and Y, and for a constant a, let $W = X - aY$. Then,

$$\mathsf{Var}[W] = \mathsf{E}\left[(X - aY)^2\right] - (\mathsf{E}\,[X - aY])^2\,. \tag{5.54}$$

Since $\mathsf{E}[X - aY] = \mu_X - a\mu_Y$, expanding the squares yields

$$\begin{aligned} \mathsf{Var}[W] &= \mathsf{E}\left[X^2 - 2aXY + a^2Y^2\right] - \left(\mu_X^2 - 2a\mu_X\mu_Y + a^2\mu_Y^2\right) \\ &= \mathsf{Var}[X] - 2a\,\mathsf{Cov}\,[X,Y] + a^2\,\mathsf{Var}[Y]. \end{aligned} \tag{5.55}$$

Since $\mathsf{Var}[W] \geq 0$ for any a, we have $2a\,\mathsf{Cov}[X,Y] \leq \mathsf{Var}[X] + a^2\,\mathsf{Var}[Y]$. Choosing $a = \sigma_X/\sigma_Y$ yields $\mathsf{Cov}[X,Y] \leq \sigma_Y\sigma_X$, which implies $\rho_{X,Y} \leq 1$. Choosing $a = -\sigma_X/\sigma_Y$ yields $\mathsf{Cov}[X,Y] \geq -\sigma_Y\sigma_X$, which implies $\rho_{X,Y} \geq -1$.

Theorem 5.15

If X and Y are random variables such that $Y = aX + b$,

$$\rho_{X,Y} = \begin{cases} -1 & a < 0, \\ 0 & a = 0, \\ 1 & a > 0. \end{cases}$$

5.8 Comment: Examples of Correlation

Some examples of positive, negative, and zero correlation coefficients include:

- X is a student's height. Y is the same student's weight. $0 < \rho_{X,Y} < 1$.
- X is the distance of a cellular phone from the nearest base station. Y is the power of the received signal at the cellular phone. $-1 < \rho_{X,Y} < 0$.
- X is the temperature of a resistor measured in degrees Celsius. Y is the temperature of the same resistor measured in Kelvins. $\rho_{X,Y} = 1$.
- X is the gain of an electrical circuit measured in decibels. Y is the attenuation, measured in decibels, of the same circuit. $\rho_{X,Y} = -1$.
- X is the telephone number of a cellular phone. Y is the Social Security number of the phone's owner. $\rho_{X,Y} = 0$.

Definition 5.7 Correlation

The correlation *of* X *and* Y *is* $r_{X,Y} = \mathrm{E}[XY]$

Theorem 5.16

(a) $\text{Cov}[X, Y] = r_{X,Y} - \mu_X \mu_Y$.

(b) $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\,\text{Cov}[X, Y]$.

(c) If $X = Y$, $\text{Cov}[X, Y] = \text{Var}[X] = \text{Var}[Y]$ and $r_{X,Y} = \text{E}[X^2] = \text{E}[Y^2]$.

Proof: Theorem 5.16

Cross-multiplying inside the expected value of Definition 5.5 yields

$$\text{Cov}\,[X,Y] = \text{E}\,[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y]\,. \tag{5.56}$$

Since the expected value of the sum equals the sum of the expected values,

$$\text{Cov}\,[X,Y] = \text{E}\,[XY] - \text{E}\,[\mu_X Y] - \text{E}\,[\mu_Y X] + \text{E}\,[\mu_Y \mu_X]\,. \tag{5.57}$$

Note that in the expression $\text{E}[\mu_Y X]$, μ_Y is a constant. Referring to Theorem 3.12, we set $a = \mu_Y$ and $b = 0$ to obtain $\text{E}[\mu_Y X] = \mu_Y \text{E}[X] = \mu_Y \mu_X$. The same reasoning demonstrates that $\text{E}[\mu_X Y] = \mu_X \text{E}[Y] = \mu_X \mu_Y$. Therefore,

$$\text{Cov}\,[X,Y] = \text{E}\,[XY] - \mu_X \mu_Y - \mu_Y \mu_X + \mu_Y \mu_X = r_{X,Y} - \mu_X \mu_Y. \tag{5.58}$$

The other relationships follow directly from the definitions and Theorem 5.12.

Example 5.17 Problem

For the integrated circuits tests in Example 5.3, we found in Example 5.5 that the probability model for X and Y is given by the following matrix.

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	$P_X(x)$
$x=0$	0.01	0	0	0.01
$x=1$	0.09	0.09	0	0.18
$x=2$	0	0	0.81	0.81
$P_Y(y)$	0.10	0.09	0.81	

Find $r_{X,Y}$ and $\text{Cov}[X,Y]$.

Example 5.17 Solution

By Definition 5.7,

$$r_{X,Y} = \mathrm{E}[XY] = \sum_{x=0}^{2} \sum_{y=0}^{2} xy P_{X,Y}(x,y) \tag{5.59}$$

$$= (1)(1)0.09 + (2)(2)0.81 = 3.33. \tag{5.60}$$

To use Theorem 5.16(a) to find the covariance, we find

$$\mathrm{E}[X] = (1)(0.18) + (2)(0.81) = 1.80,$$
$$\mathrm{E}[Y] = (1)(0.09) + (2)(0.81) = 1.71. \tag{5.61}$$

Therefore, by Theorem 5.16(a), $\mathrm{Cov}[X,Y] = 3.33 - (1.80)(1.71) = 0.252$.

Definition 5.8 Orthogonal Random Variables

Random variables X *and* Y *are* orthogonal *if* $r_{X,Y} = 0$.

Definition 5.9 Uncorrelated Random Variables

Random variables X and Y are uncorrelated *if* $\text{Cov}[X, Y] = 0$.

Theorem 5.17

For independent random variables X and Y,

(a) $\mathsf{E}[g(X)h(Y)] = \mathsf{E}[g(X)]\,\mathsf{E}[h(Y)]$,

(b) $r_{X,Y} = \mathsf{E}[XY] = \mathsf{E}[X]\,\mathsf{E}[Y]$,

(c) $\mathsf{Cov}[X,Y] = \rho_{X,Y} = 0$,

(d) $\mathsf{Var}[X+Y] = \mathsf{Var}[X] + \mathsf{Var}[Y]$,

Proof: Theorem 5.17

We present the proof for discrete random variables. By replacing PMFs and sums with PDFs and integrals we arrive at essentially the same proof for continuous random variables. Since $P_{X,Y}(x,y) = P_X(x)P_Y(y)$,

$$\begin{aligned} \mathsf{E}\left[g(X)h(Y)\right] &= \sum_{x\in S_X}\sum_{y\in S_Y} g(x)h(y)P_X(x)\,P_Y(y) \\ &= \left(\sum_{x\in S_X} g(x)P_X(x)\right)\left(\sum_{y\in S_Y} h(y)P_Y(y)\right) = \mathsf{E}\left[g(X)\right]\mathsf{E}\left[h(Y)\right]. \qquad (5.62) \end{aligned}$$

If $g(X) = X$, and $h(Y) = Y$, this equation implies $r_{X,Y} = \mathsf{E}[XY] = \mathsf{E}[X]\,\mathsf{E}[Y]$. This equation and Theorem 5.16(a) imply $\mathrm{Cov}[X,Y] = 0$. As a result, Theorem 5.16(b) implies $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$. Furthermore, $\rho_{X,Y} = \mathrm{Cov}[X,Y]/(\sigma_X\sigma_Y) = 0$.

Example 5.18 Problem

For the noisy observation $Y = X + Z$ of Example 5.1, find the covariances $\text{Cov}[X, Z]$ and $\text{Cov}[X, Y]$ and the correlation coefficients $\rho_{X,Z}$ and $\rho_{X,Y}$.

Example 5.18 Solution

We recall from Example 5.1 that the signal X is Gaussian $(0, \sigma_X)$, that the noise Z is Gaussian $(0, \sigma_Z)$, and that X and Z are independent. We know from Theorem 5.17(c) that independence of X and Z implies

$$\text{Cov}[X, Z] = \rho_{X,Z} = 0. \tag{5.63}$$

In addition, by Theorem 5.17(d),

$$\text{Var}[Y] = \text{Var}[X] + \text{Var}[Z] = \sigma_X^2 + \sigma_Z^2. \tag{5.64}$$

Since $\text{E}[X] = \text{E}[Z] = 0$, Theorem 5.11 tells us that $\text{E}[Y] = \text{E}[X] + \text{E}[Z] = 0$ and Theorem 5.17)(b) says that $\text{E}[XZ] = \text{E}[X]\,\text{E}[Z] = 0$. This permits us to write

$$\begin{aligned} \text{Cov}[X, Y] &= \text{E}[XY] = \text{E}[X(X+Z)] \\ &= \text{E}\left[X^2 + XZ\right] = \text{E}\left[X^2\right] + \text{E}[XZ] = \text{E}\left[X^2\right] = \sigma_X^2. \end{aligned} \tag{5.65}$$

This implies

$$\rho_{X,Y} = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]\,\text{Var}[Y]}} = \frac{\sigma_X^2}{\sqrt{\sigma_X^2(\sigma_X^2 + \sigma_Z^2)}} = \sqrt{\frac{\sigma_X^2/\sigma_Z^2}{1 + \sigma_X^2/\sigma_Z^2}}. \tag{5.66}$$

Quiz 5.8(A)

Random variables L and T have joint PMF

$P_{L,T}(l,t)$	$t = 40\,\text{sec}$	$t = 60\,\text{sec}$
$l = 1$ page	0.15	0.1
$l = 2$ pages	0.30	0.2
$l = 3$ pages	0.15	0.1.

Find the following quantities.

(a) E[L] and Var[L]

(b) E[T] and Var[T]

(c) The covariance Cov[L, T]

(d) The correlation coefficient $\rho_{L,T}$

Quiz 5.8(A) Solution

It is helpful to first make a table that includes the marginal PMFs.

$P_{L,T}(l,t)$	$t = 40$	$t = 60$	$P_L(l)$
$l = 1$	0.15	0.1	0.25
$l = 2$	0.3	0.2	0.5
$l = 3$	0.15	0.1	0.25
$P_T(t)$	0.6	0.4	

(a) The expected value of L is

$$E[L] = 1(0.25) + 2(0.5) + 3(0.25) = 2. \quad (1)$$

Since the second moment of L is

$$E\left[L^2\right] = 1^2(0.25) + 2^2(0.5) + 3^2(0.25) = 4.5, \quad (2)$$

the variance of L is

$$\text{Var}[L] = E\left[L^2\right] - (E[L])^2 = 0.5. \quad (3)$$

(b) The expected value of T is

$$E[T] = 40(0.6) + 60(0.4) = 48. \quad (4)$$

The second moment of T is

$$E\left[T^2\right] = 40^2(0.6) + 60^2(0.4) = 2400. \quad (5)$$

[Continued]

Quiz 5.8(A) Solution (Continued 2)

Thus

$$\text{Var}[T] = \mathsf{E}\left[T^2\right] - (\mathsf{E}[T])^2 = 96. \tag{6}$$

(c) First we need to find

$$\begin{aligned}\mathsf{E}[LT] &= \sum_{t=40,60} \sum_{l=1}^{3} lt P_{LT}(lt) \\ &= 1(40)(0.15) + 2(40)(0.3) + 3(40)(0.15) \\ &\quad + 1(60)(0.1) + 2(60)(0.2) + 3(60)(0.1) \\ &= 96.\end{aligned} \tag{7}$$

The covariance of L and T is

$$\text{Cov}[L,T] = \mathsf{E}[LT] - \mathsf{E}[L]\,\mathsf{E}[T] = 96 - 2(48) = 0. \tag{8}$$

(d) Since $\text{Cov}[L,T] = 0$, the correlation coefficient is $\rho_{L,T} = 0$.

Quiz 5.8(B)

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x,y) = \begin{cases} xy & 0 \leq x \leq 1, 0 \leq y \leq 2, \\ 0 & \text{otherwise.} \end{cases} \tag{5.67}$$

Find the following quantities.

(a) $\mathrm{E}[X]$ and $\mathrm{Var}[X]$

(b) $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$

(c) The covariance $\mathrm{Cov}[X,Y]$

(d) The correlation coefficient $\rho_{X,Y}$

Quiz 5.8(B) Solution

As in the discrete case, the calculations become easier if we first calculate the marginal PDFs $f_X(x)$ and $f_Y(y)$. For $0 \leq x \leq 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy = \int_0^2 xy\,dy = \frac{1}{2}xy^2\Big|_{y=0}^{y=2} = 2x. \tag{1}$$

Similarly, for $0 \leq y \leq 2$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dx = \int_0^2 xy\,dx = \frac{1}{2}x^2y\Big|_{x=0}^{x=1} = \frac{y}{2}. \tag{2}$$

The complete expressions for the marginal PDFs are

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1, \\ 0 & \text{otherwise,} \end{cases} \qquad f_Y(y) = \begin{cases} y/2 & 0 \leq y \leq 2, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

From the marginal PDFs, it is straightforward to calculate the various expectations.
(a) The first and second moments of X are

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\ dx = \int_0^1 2x^2\,dx = \frac{2}{3}. \tag{4}$$

$$\mathrm{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\ dx = \int_0^1 2x^3\,dx = \frac{1}{2}. \tag{5}$$

[Continued]

Quiz 5.8(B) Solution (Continued 2)

The variance of X is

$$\text{Var}[X] = \text{E}[X^2] - (\text{E}[X])^2 = \frac{1}{18}.$$

(a) The first and second moments of Y are

$$\text{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_0^2 \frac{1}{2} y^2\, dy = \frac{4}{3}, \tag{6}$$

$$\text{E}\left[Y^2\right] = \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy = \int_0^2 \frac{1}{2} y^3\, dy = 2. \tag{7}$$

The variance of Y is

$$\text{Var}[Y] = \text{E}\left[Y^2\right] - (\text{E}[Y])^2 = 2 - \frac{16}{9} = \frac{2}{9}. \tag{8}$$

(b) We start by finding

$$\begin{aligned} \text{E}[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\, dx, dy \\ &= \int_0^1 \int_0^2 x^2 y^2\, dx, dy = \left.\frac{x^3}{3}\right|_0^1 \left.\frac{y^3}{3}\right|_0^2 = \frac{8}{9}. \end{aligned} \tag{9}$$

The covariance of X and Y is then

$$\text{Cov}[X,Y] = \text{E}[XY] - \text{E}[X]\,\text{E}[Y] = \frac{8}{9} - \frac{2}{3} \cdot \frac{4}{3} = 0. \tag{10}$$

(c) Since $\text{Cov}[X,Y] = 0$, the correlation coefficient is $\rho_{X,Y} = 0$.

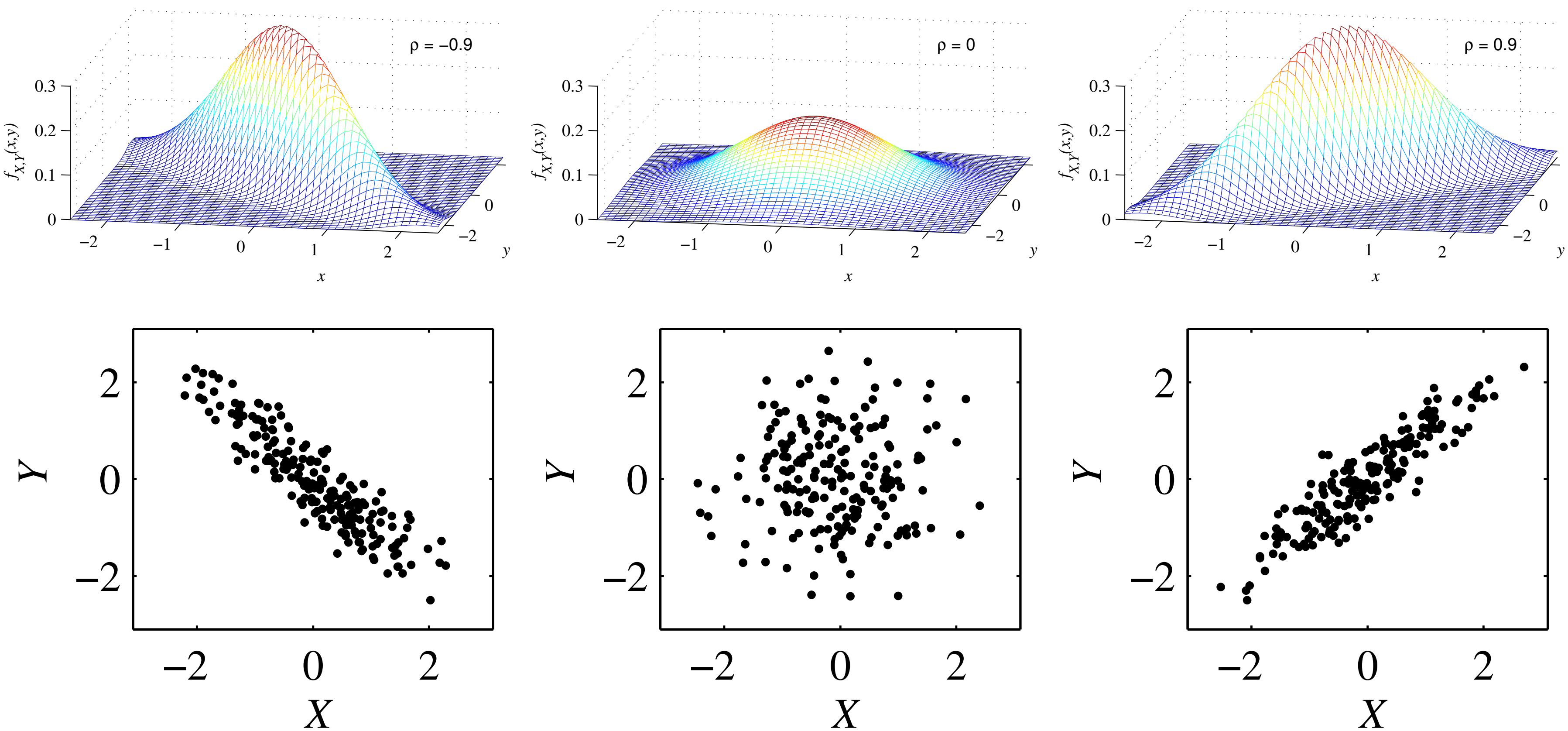
Bivariate Gaussian Random Variables

Definition 5.10 Bivariate Gaussian Random Variables

Random variables X *and* Y *have a* bivariate Gaussian PDF *with parameters* μ_X, μ_Y, $\sigma_X > 0$, $\sigma_Y > 0$, *and* $\rho_{X,Y}$ *satisfying* $-1 < \rho_{X,Y} < 1$ *if*

$$f_{X,Y}(x,y) = \frac{\exp\left[-\dfrac{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{2\rho_{X,Y}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}{2(1-\rho_{X,Y}^2)}\right]}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}},$$

Figure 5.6



The Joint Gaussian PDF $f_{X,Y}(x, y)$ for $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$, and three values of $\rho_{X,Y} = \rho$. Next to each PDF, we plot 200 sample pairs (X, Y) generated with that PDF.

Theorem 5.18

If X and Y are the bivariate Gaussian random variables in Definition 5.10, X is the Gaussian (μ_X, σ_X) random variable and Y is the Gaussian (μ_Y, σ_Y) random variable:

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/2\sigma_X^2}, \qquad f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-(y-\mu_Y)^2/2\sigma_Y^2}.$$

Proof: Theorem 5.18

Integrating $f_{X,Y}(x,y)$ in Equation (5.69) over all y, we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$
$$= \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/2\sigma_X^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\tilde{\sigma}_Y\sqrt{2\pi}} e^{-(y-\tilde{\mu}_Y(x))^2/2\tilde{\sigma}_Y^2}\, dy}_{1} \tag{5.70}$$

The integral above the bracket equals 1 because it is the integral of a Gaussian PDF. The remainder of the formula is the PDF of the Gaussian (μ_X, σ_X) random variable. The same reasoning with the roles of X and Y reversed leads to the formula for $f_Y(y)$.

Theorem 5.19

Bivariate Gaussian random variables X and Y in Definition 5.10 have correlation coefficient $\rho_{X,Y}$.

Theorem 5.20

Bivariate Gaussian random variables X and Y are uncorrelated if and only if they are independent.

Theorem 5.21

If X and Y are bivariate Gaussian random variables with PDF given by Definition 5.10, and W_1 and W_2 are given by the linearly independent equations

$$W_1 = a_1 X + b_1 Y, \qquad W_2 = a_2 X + b_2 Y,$$

then W_1 and W_2 are bivariate Gaussian random variables such that

$$\begin{aligned} \mathrm{E}[W_i] &= a_i \mu_X + b_i \mu_Y, & i = 1, 2, \\ \mathrm{Var}[W_i] &= a_i^2 \sigma_X^2 + b_i^2 \sigma_Y^2 + 2 a_i b_i \rho_{X,Y} \sigma_X \sigma_Y, & i = 1, 2, \\ \mathrm{Cov}[W_1, W_2] &= a_1 a_2 \sigma_X^2 + b_1 b_2 \sigma_Y^2 + (a_1 b_2 + a_2 b_1) \rho_{X,Y} \sigma_X \sigma_Y. \end{aligned}$$

Example 5.19 Problem

For the noisy observation in Example 5.14, find the PDF of $Y = X + Z$.

Example 5.19 Solution

Since X is Gaussian $(0, \sigma_X)$ and Z is Gaussian $(0, \sigma_Z)$ and X and Z are independent, X and Z are jointly Gaussian. It follows from Theorem 5.21 that Y is Gaussian with $\mathrm{E}[Y] = \mathrm{E}[X] + \mathrm{E}[Z] = 0$ and variance $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2$. The PDF of Y is

$$f_Y(y) = \frac{1}{\sqrt{2\pi(\sigma_X^2 + \sigma_Z^2)}} e^{-y^2/2(\sigma_X^2+\sigma_Z^2)}. \tag{5.71}$$

Example 5.20 Problem

Continuing Example 5.19, find the joint PDF of X and Y when $\sigma_X = 4$ and $\sigma_Z = 3$.

Example 5.20 Solution

From Theorem 5.21, we know that X and Y are bivariate Gaussian. We also know that $\mu_X = \mu_Y = 0$ and that Y has variance $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2 = 25$. Substituting $\sigma_X = 4$ and $\sigma_Z = 3$ in the formula for the correlation coefficient derived in Example 5.18, we have

$$\rho_{X,Y} = \sqrt{\frac{\sigma_X^2/\sigma_Z^2}{1+\sigma_X^2/\sigma_Z^2}} = \frac{4}{5}. \tag{5.72}$$

Applying these parameters to Definition 5.10, we obtain

$$f_{X,Y}(x,y) = \frac{1}{24\pi} e^{-\left(25x^2/16 - 2xy + y^2\right)/18}. \tag{5.73}$$

Quiz 5.9

Let X and Y be jointly Gaussian $(0, 1)$ random variables with correlation coefficient $1/2$. What is the joint PDF of X and Y?

Quiz 5.9 Solution

This problem just requires identifying the various parameters in Definition 5.10. Specifically, from the problem statement, we know $\rho = 1/2$ and

$$\mu_X = 0, \qquad \mu_Y = 0,$$
$$\sigma_X = 1, \qquad \sigma_Y = 1.$$

Applying these facts to Definition 5.10, we have

$$f_{X,Y}(x,y) = \frac{e^{-2(x^2 - xy + y^2)/3}}{\sqrt{3\pi^2}}. \tag{1}$$

Section 5.10

Multivariate Probability Models

Definition 5.11 Multivariate Joint CDF

The joint CDF *of* $X_1, \ldots, X_n$ *is*

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathsf{P}\left[X_1 \le x_1, \ldots, X_n \le x_n\right].$$

Definition 5.12 Multivariate Joint PMF

The joint PMF *of the discrete random variables* $X_1, \ldots, X_n$ *is*

$$P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathsf{P}\left[X_1 = x_1, \ldots, X_n = x_n\right].$$

Definition 5.13 Multivariate Joint PDF

The joint PDF *of the continuous random variables* $X_1, \ldots, X_n$ *is the function*

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \frac{\partial^n F_{X_1,\ldots,X_n}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}.$$

Theorem 5.22

If $X_1, \ldots, X_n$ are discrete random variables with joint PMF $P_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$,

(a) $P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) \geq 0$,

(b) $\displaystyle\sum_{x_1 \in S_{X_1}} \cdots \sum_{x_n \in S_{X_n}} P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = 1$.

Theorem 5.23

If $X_1, \ldots, X_n$ have joint PDF $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$,

(a) $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) \geq 0$,

(b) $F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1,\ldots,X_n}(u_1, \ldots, u_n)\, du_1 \cdots du_n$,

(c) $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n = 1$.

Theorem 5.24

The probability of an event A expressed in terms of the random variables $X_1, \ldots, X_n$ is

Discrete: $\mathrm{P}[A] = \sum_{(x_1,\ldots,x_n)\in A} P_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$

Continuous: $\mathrm{P}[A] = \int \cdots_A \int f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n.$

Example 5.21 Problem

Consider a set of n independent trials in which there are r possible outcomes $s_1, \ldots, s_r$ for each trial. In each trial, $\mathrm{P}[s_i] = p_i$. Let N_i equal the number of times that outcome s_i occurs over n trials. What is the joint PMF of $N_1, \ldots, N_r$?

Example 5.21 Solution

The solution to this problem appears in Theorem 2.9 and is repeated here:

$$P_{N_1,\ldots,N_r}(n_1,\ldots,n_r) = \binom{n}{n_1,\ldots,n_r} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}. \tag{5.74}$$

Theorem 5.25

For a joint PMF $P_{W,X,Y,Z}(w,x,y,z)$ of discrete random variables W, X, Y, Z, some marginal PMFs are

$$P_{X,Y,Z}(x,y,z) = \sum_{w \in S_W} P_{W,X,Y,Z}(w,x,y,z)\,,$$

$$P_{W,Z}(w,z) = \sum_{x \in S_X} \sum_{y \in S_Y} P_{W,X,Y,Z}(w,x,y,z)\,,$$

Theorem 5.26

For a joint PDF $f_{W,X,Y,Z}(w,x,y,z)$ of continuous random variables W, X, Y, Z, some marginal PDFs are

$$f_{W,X,Y}(w,x,y) = \int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)\ dz,$$

$$f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)\ dw\,dy\,dz.$$

Example 5.22 Problem

As in Quiz 5.10, the random variables $Y_1, \ldots, Y_4$ have the joint PDF

$$f_{Y_1,\ldots,Y_4}(y_1, \ldots, y_4) = \begin{cases} 4 & 0 \le y_1 \le y_2 \le 1, 0 \le y_3 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.75}$$

Find the marginal PDFs $f_{Y_1,Y_4}(y_1, y_4)$, $f_{Y_2,Y_3}(y_2, y_3)$, and $f_{Y_3}(y_3)$.

Example 5.22 Solution

$$f_{Y_1,Y_4}(y_1,y_4) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{Y_1,\ldots,Y_4}(y_1,\ldots,y_4)\;dy_2\,dy_3. \tag{5.76}$$

In the foregoing integral, the hard part is identifying the correct limits. These limits will depend on y_1 and y_4. For $0 \le y_1 \le 1$ and $0 \le y_4 \le 1$,

$$f_{Y_1,Y_4}(y_1,y_4) = \int_{y_1}^{1}\int_{0}^{y_4} 4\,dy_3\,dy_2 = 4(1-y_1)y_4. \tag{5.77}$$

The complete expression for $f_{Y_1,Y_4}(y_1,y_4)$ is

$$f_{Y_1,Y_4}(y_1,y_4) = \begin{cases} 4(1-y_1)y_4 & 0 \le y_1 \le 1, 0 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.78}$$

Similarly, for $0 \le y_2 \le 1$ and $0 \le y_3 \le 1$,

$$f_{Y_2,Y_3}(y_2,y_3) = \int_{0}^{y_2}\int_{y_3}^{1} 4\,dy_4\,dy_1 = 4y_2(1-y_3). \tag{5.79}$$

[Continued]

Example 5.22 Solution (Continued 2)

The complete expression for $f_{Y_2,Y_3}(y_2, y_3)$ is

$$f_{Y_2,Y_3}(y_2, y_3) = \begin{cases} 4y_2(1-y_3) & 0 \le y_2 \le 1, 0 \le y_3 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.80}$$

Lastly, for $0 \le y_3 \le 1$,

$$f_{Y_3}(y_3) = \int_{-\infty}^{\infty} f_{Y_2,Y_3}(y_2, y_3)\, dy_2 = \int_0^1 4y_2(1-y_3)\, dy_2 = 2(1-y_3). \tag{5.81}$$

The complete expression is

$$f_{Y_3}(y_3) = \begin{cases} 2(1-y_3) & 0 \le y_3 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.82}$$

Definition 5.14 N Independent Random Variables

Random variables $X_1, \ldots, X_n$ *are* independent *if for all* $x_1, \ldots, x_n$,

Discrete: $P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P_{X_1}(x_1)P_{X_2}(x_2)\cdots P_{X_N}(x_n)$

Continuous: $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$.

Definition 5.15 Independent and Identically Distributed (iid)

$X_1, \ldots, X_n$ *are* independent and identically distributed (iid) *if*

Discrete: $P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P_X(x_1)P_X(x_2)\cdots P_X(x_n)$

Continuous: $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_X(x_1)f_X(x_2)\cdots f_X(x_n).$

Example 5.23 Problem

The random variables $X_1, \ldots, X_n$ have the joint PDF

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \begin{cases} 1 & 0 \le x_i \le 1, i = 1, \ldots, n, \\ 0 & \text{otherwise.} \end{cases} \tag{5.83}$$

Let A denote the event that $\max_i X_i \le 1/2$. Find $P[A]$.

Example 5.23 Solution

We can solve this problem by applying Theorem 5.24:

$$\begin{aligned}\mathsf{P}[A] = \mathsf{P}\left[\max_i X_i \le 1/2\right] &= \mathsf{P}[X_1 \le 1/2, \ldots, X_n \le 1/2] \\ &= \int_0^{1/2} \cdots \int_0^{1/2} 1\, dx_1 \cdots dx_n = \frac{1}{2^n}. \qquad (5.84)\end{aligned}$$

As n grows, the probability that the maximum is less than 1/2 rapidly goes to 0.

We note that inspection of the joint PDF reveals that $X_1, \ldots, X_4$ are iid continuous uniform $(0, 1)$ random variables. The integration in Equation (5.84) is easy because independence implies

$$\begin{aligned}\mathsf{P}[A] &= \mathsf{P}[X_1 \le 1/2, \ldots, X_n \le 1/2] \\ &= \mathsf{P}[X_1 \le 1/2] \times \cdots \times \mathsf{P}[X_n \le 1/2] = (1/2)^n. \qquad (5.85)\end{aligned}$$

Quiz 5.10

The random variables $Y_1, \ldots, Y_4$ have the joint PDF

$$f_{Y_1,\ldots,Y_4}(y_1, \ldots, y_4) = \begin{cases} 4 & 0 \le y_1 \le y_2 \le 1, 0 \le y_3 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.86}$$

Let C denote the event that $\max_i Y_i \le 1/2$. Find $P[C]$.

Quiz 5.10 Solution

We find P$[C]$ by integrating the joint PDF over the region of interest. Specifically,

$$\begin{aligned}
\mathrm{P}[C] &= \int_0^{\frac{1}{2}} dy_2 \int_0^{y_2} dy_1 \int_0^{\frac{1}{2}} dy_4 \int_0^{y_4} 4 dy_3 \\
&= 4 \left(\int_0^{\frac{1}{2}} y_2 \, dy_2 \right) \left(\int_0^{\frac{1}{2}} y_4 \, dy_4 \right) \\
&= 4 \left(\frac{1}{2} y_2^2 \Big|_0^{\frac{1}{2}} \right) \left(\frac{1}{2} y_4^2 \Big|_0^{\frac{1}{2}} \right) = 4 \left(\frac{1}{8} \right)^2 = \frac{1}{16}.
\end{aligned} \tag{1}$$

Section 5.11

Matlab

Sample Space Grids

- We start with the case when X and Y are finite random variables with ranges

$$S_X = \{x_1, \ldots, x_n\}, \qquad S_Y = \{y_1, \ldots, y_m\}. \tag{5.87}$$

  In this case, we can take advantage of Matlab techniques for surface plots of $g(x, y)$ over the x, y plane.

- In Matlab, we represent S_X and S_Y by the n element vector `sx` and m element vector `sy`.

- The function `[SX,SY]=ndgrid(sx,sy)` produces the pair of $n \times m$ matrices,

$$\mathtt{SX} = \begin{bmatrix} x_1 & \cdots & x_1 \\ \vdots & & \vdots \\ x_n & \cdots & x_n \end{bmatrix}, \qquad \mathtt{SY} = \begin{bmatrix} y_1 & \cdots & y_m \\ \vdots & & \vdots \\ y_1 & \cdots & y_m \end{bmatrix}. \tag{5.88}$$

  We refer to matrices `SX` and `SY` as a *sample space grid* because they are a grid representation of the joint sample space

$$S_{X,Y} = \{(x, y) | x \in S_X, y \in S_Y\}. \tag{5.89}$$

  That is, `[SX(i,j) SY(i,j)]` is the pair (x_i, y_j).

Probabilities on Grids

- To complete the probability model, for X and Y, in Matlab, we employ the $n \times m$ matrix `PXY` such that `PXY(i,j)` $= P_{X,Y}(x_i, y_j)$.

- To make sure that probabilities have been generated properly, we note that

  ```
  [SX(:)  SY(:)  PXY(:)]
  ```

  is a matrix whose rows list all possible pairs x_i, y_j and corresponding probabilities $P_{X,Y}(x_i, y_j)$.

- Given a function $g(x, y)$ that operates on the elements of vectors `x` and `y`, the advantage of this grid approach is that the Matlab function `g(SX,SY)` will calculate $g(x, y)$ for each $x \in S_X$ and $y \in S_Y$.

- In particular, `g(SX,SY)` produces an $n \times m$ matrix with i, jth element $g(x_i, y_j)$.

Example 5.24 Problem

An Internet photo developer website prints compressed photo images. Each image file contains a variable-sized image of $X \times Y$ pixels described by the joint PMF

$P_{X,Y}(x,y)$	$y = 400$	$y = 800$	$y = 1200$
$x = 800$	0.2	0.05	0.1
$x = 1200$	0.05	0.2	0.1
$x = 1600$	0	0.1	0.2.

(5.90)

For random variables X, Y, write a script `imagepmf.m` that defines the sample space grid matrices `SX`, `SY`, and `PXY`.

Example 5.24 Solution

In the script `imagepmf.m`, the matrix `SX` has $\begin{bmatrix}800 & 1200 & 1600\end{bmatrix}'$ for each column and `SY` has $\begin{bmatrix}400 & 800 & 1200\end{bmatrix}$ for each row.

After running `imagepmf.m`, we can inspect the variables:

```
%imagepmf.m
PXY=[0.2  0.05 0.1; ...
     0.05 0.2  0.1; ...
     0    0.1  0.2];
[SX,SY]=ndgrid([800 1200 1600],...
    [400 800 1200]);
```

```
>> imagepmf; SX
SX =
     800    800   800
    1200   1200  1200
    1600   1600  1600
>> SY
SY =
     400    800  1200
     400    800  1200
     400    800  1200
```

Example 5.25 Problem

At 24 bits (3 bytes) per pixel, a 10:1 image compression factor yields image files with $B = 0.3XY$ bytes. Find the expected value $\mathrm{E}[B]$ and the PMF $P_B(b)$.

Example 5.25 Solution

```
%imagesize.m
imagepmf;
SB=0.3*(SX.*SY);
eb=sum(sum(SB.*PXY))
sb=unique(SB)'
pb=finitepmf(SB,PXY,sb)'
```

The script `imagesize.m` produces the expected value as `eb`, and produces the PMF, which is represented by the vectors `sb` and `pb`. The 3×3 matrix `SB` has i,jth element $g(x_i, y_j) = 0.3x_iy_j$. The calculation of `eb` is simply a Matlab implementation of Theorem 5.9. Since some elements of `SB` are identical, `sb=unique(SB)` extracts the unique elements. Although `SB` and `PXY` are both 3×3 matrices, each is stored internally by Matlab as a 9-element vector. Hence, we can pass `SB` and `PXY` to the `finitepmf()` function, which was designed to handle a finite random variable described by a pair of column vectors. Figure 5.7 shows one result of running the program `imagesize`. The vectors `sb` and `pb` comprise $P_B(b)$. For example, $P_B(288000) = 0.3$.

Figure 5.7

```
>> imagesize
eb =
      319200
sb =
    96000   144000   192000   288000   384000   432000   576000
pb =
   0.2000   0.0500   0.0500   0.3000   0.1000   0.1000   0.2000
```

Output resulting from `imagesize.m` in Example 5.25.

Example 5.26 Problem

Write a function `xy=imagerv(m)` that generates m sample pairs of the image size random variables X, Y of Example 5.25.

Example 5.26 Solution

The function `imagerv` uses the `imagesize.m` script to define the matrices `SX`, `SY`, and `PXY`. It then calls the `finiterv.m` function. Here is the code `imagerv.m` and a sample run:

```
function xy = imagerv(m);
imagepmf;
S=[SX(:) SY(:)];
xy=finiterv(S,PXY(:),m);
```

```
>> xy=imagerv(3)
xy =
         800          400
        1200          800
        1600          800
```

Example 5.27 Problem

Given a list `xy` of sample pairs of random variables X, Y with Matlab range grids `SX` and `SY`, write a Matlab function

```
fxy=freqxy(xy,SX,SY)
```

that calculates the relative frequency of every pair x, y. The output `fxy` should correspond to the matrix `[SX(:) SY(:) PXY(:)]`.

Example 5.27 Solution

```
function fxy = freqxy(xy,SX,SY)
xy=[xy; SX(:) SY(:)];
[U,I,J]=unique(xy,'rows');
N=hist(J,1:max(J))-1;
N=N/sum(N);
fxy=[U N(:)];
fxy=sortrows(fxy,[2 1 3]);
```

The matrix `[SX(:) SY(:)]` in `freqxy` has rows that list all possible pairs x, y. We append this matrix to `xy` to ensure that the new `xy` has every possible pair x, y. Next, the `unique` function copies all unique rows of `xy` to the matrix `U` and also provides the vector `J` that indexes the rows of `xy` in `U`; that is, `xy=U(J)`. In addition, the number of occurrences of `j` in `J` indicates the number of occurrences in `xy` of row `j` in `U`. Thus we use the `hist` function on `J` to calculate the relative frequencies. We include the correction factor `-1` because we had appended `[SX(:) SY(:)]` to `xy` at the start. Lastly, we reorder the rows of `fxy` because the output of `unique` produces the rows of `U` in a different order from `[SX(:) SY(:) PXY(:)]`.

Example 5.28 Problem

Generate $m = 10,000$ samples of random variables X, Y of Example 5.25. Calculate the relative frequencies and use `stem3` to graph them.

Example 5.28 Solution

The script `imagestem.m` generates the following relative frequency stem plot.

```
%imagestem.m
imagepmf;
xy=imagerv(10000);
fxy=freqxy(xy,SX,SY);
stem3(fxy(:,1),...
    fxy(:,2),fxy(:,3));
xlabel('\it x');
ylabel('\it y');
```