Chapter 4
Pairs of Random Variables

Chien-Kang Huang (黃乾綱)
臺大工科海洋系副教授

Outline

- Chapter 2, 3 analyze experiments in which outcome is one number
  - Chapter 2 considers only discrete random variable
  - Chapter 3 considers only continuous random variable
- Chapter 4, 5 analyze experiments in which outcome is a collection of numbers.
  - Chapter 4 analyzes experiments that produce **2** random variables.
  - Chapter 5 analyzes experiments that produce n random variables.
- We find that many formulas come in pairs.
  - One formula contains **sums** for **discrete** random variable
  - One formula contains **integrals** for **continuous** random variable

Joint CDF $F_{X,Y}(x,y)$ → joint PMF $P_{X,Y}(x,y)$ and joint PDF $f_{X,Y}(x,y)$

- Joint CDF $F_{X,Y}(x,y)$
  - The joint CDF is a complete probability model for any experiment that produces two random variables.
  - However, it is not very useful for analyzing practical experiments.
- Joint PMF $P_{X,Y}(x,y)$
  - More useful model for two discrete random variables.
- Joint PDF $f_{X,Y}(x,y)$
  - More useful model for two continuous random variables.

- Pairs of random variables appear in a wide variety of practical situations.

Examples

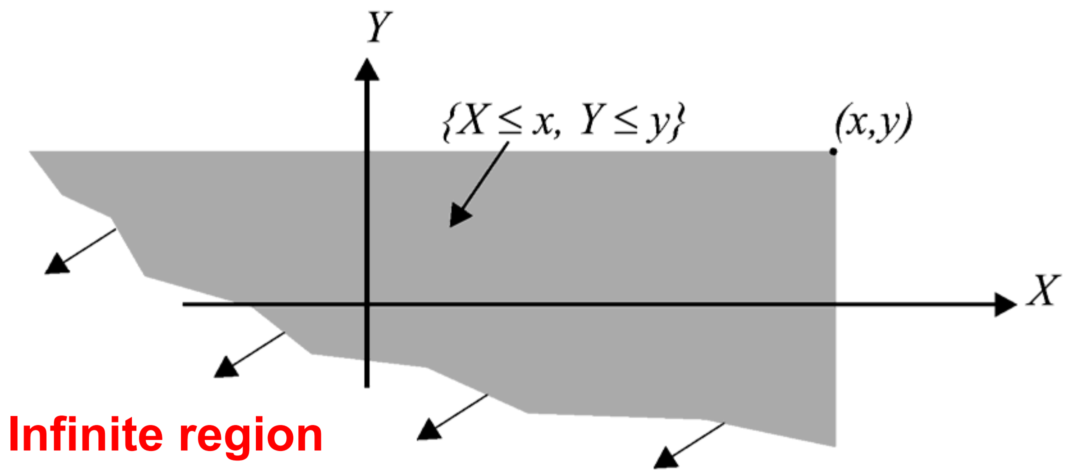
- X: the **signal** emitted by a radio transmitter
  Y: the corresponding **signal** that eventually arrives at a receiver.
- We observe Y, but we really want to know X. Noise and distortion prevent us from observing X directly and we use the probability model $f_{X,Y}(x, y)$ to estimate X.
- X: distance of the telephone from the base station
  Y: the strength of the signal at a cellular telephone base station receiver.

4.1 Joint Cumulative Distribution Function

Number of random variables	Outcomes	Events
One random variables	Points on a line.	Points or intervals on the line.
Two random variables	Points in a plane.	Points, line, segments, projection or areas in the plane.

- The CDF $F_X(x)$ is the probability of the interval to the left of x
- The joint CDF $F_{X,Y}(x,y)$ is the probability of the area in the plane below and to the left of (x,y). (reference Figure 4.1)

Figure 4.1



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

Definition 4.1 Joint Cumulative Distribution Function (CDF)

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

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

- The joint CDF is a complete probability model.
- Event $\{X \leq x\}$ suggests that Y can have any value so long as the condition on X is met.

$$F_X(x) = P[X \leq x] = P[X \leq x, y < \infty] = \lim_{y \to \infty} F_{X,Y}(x, y) = F_{X,Y}(x, \infty)$$

Theorem 4.1

For any pair of random variables, X, Y,

(a) $0 \leq F_{X,Y}(x, y) \leq 1$,

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

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

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

(e) If $x \leq x_1$ and $y \leq y_1$, then $F_{X,Y}(x, y) \leq F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = 1$.

Quiz 4.1

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 $F_Y(y)$.

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

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

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

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

Quiz 4.1 Solution

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

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

(2) $F_{X,Y}(\infty, \infty) = P[X \leq \infty, Y \leq \infty] = 1$. This result is given in Theorem 4.1.

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

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

4.2 Joint Probability Mass Function

Joint Probability Mass Function

Definition 4.2 (PMF)

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

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

- $S_{X,Y}$: the set of possible values of the pair (X, Y)

$$S_{X,Y} = \{(x, y) \mid P_{X,Y}(x, y) > 0\}$$

- $\{X = x, Y = y\}$ is an event in an experiment.

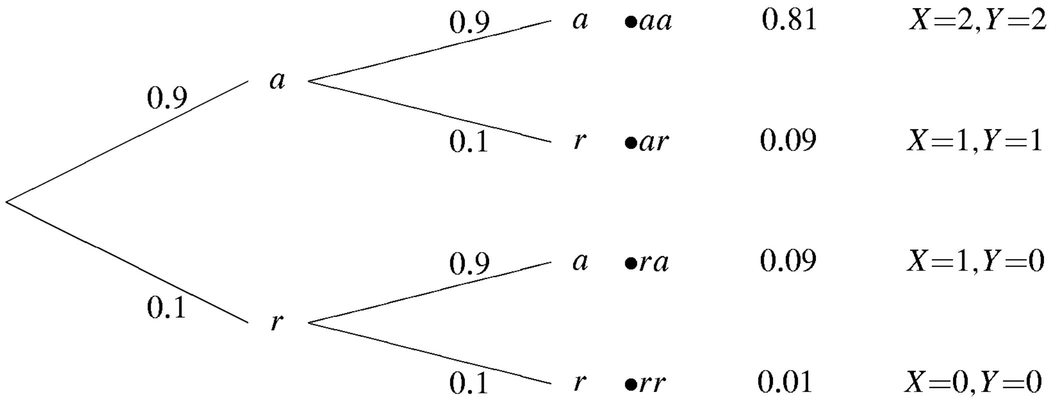
Example 4.1 Problem

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 of X and Y.

- There are various ways to represent a joint PMF.
  - A **list**, a **matrix**, and a **graph**

Example 4.1 Solution

The experiment has the following tree diagram.



The sample space of the experiment is

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

Observing the tree diagram, we compute

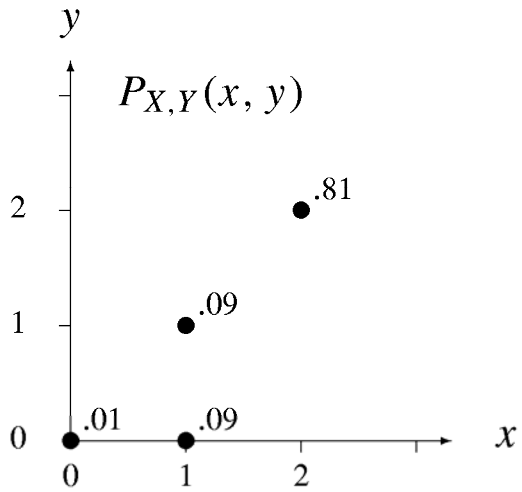
$$P[aa] = 0.81, \qquad P[ar] = P[ra] = 0.09, \qquad P[rr] = 0.01.$$

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 [Continued]

Example 4.1 Solution (continued)

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

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, $P_{X,Y}(1,1) = P[ar]$. The joint PMF can be given as a set of labeled points in the x, y plane where each point is a possible value (probability > 0) of the pair (x, y), or as a simple list:



graph

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

list

[Continued]

Example 4.1 Solution (continued)

A third representation of $P_{X,Y}(x, y)$ is the matrix:

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

matrix

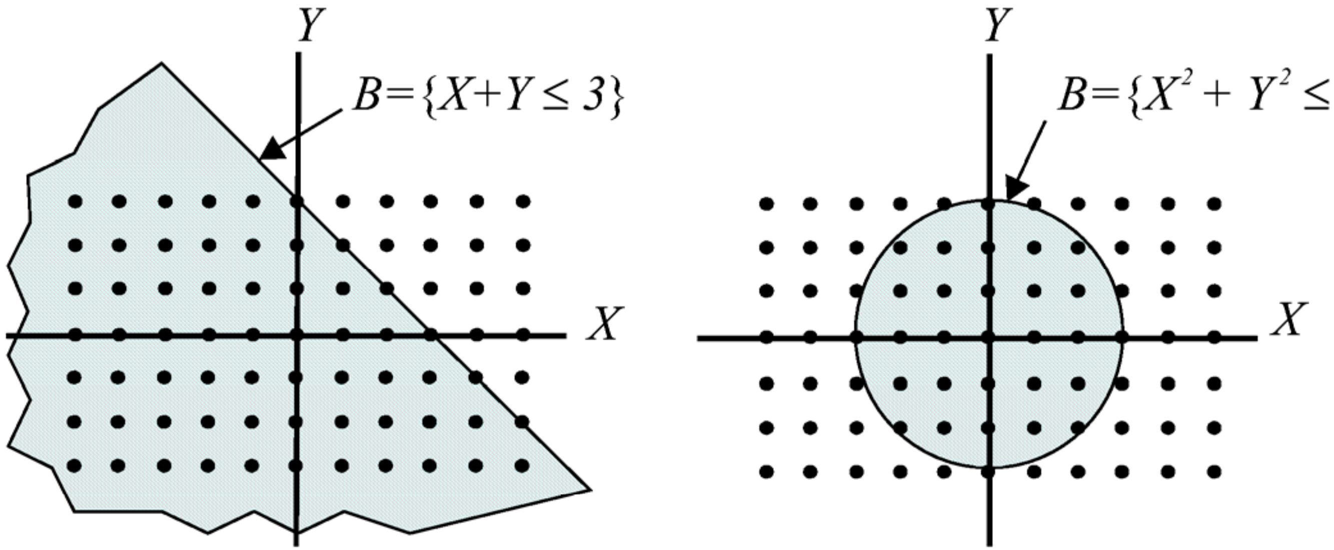
Comments of Axiom

- Axiom 2
  - Note that all the probabilities add up to 1. The reflects the second axiom of probability (section 1.3) that states $P[S] = 1$. Using the notation of random variables, we write this as

$$\sum_{x \in S_X} \sum_{y \in S_Y} P_{X,Y}(x, y) = 1$$

- Axiom 1
  - $P_{X,Y}(x, y) \geq 0$ for all pairs x, y
- Axiom 3
  - Union of disjoint events.

Figure 4.2 Two examples of events.



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

- When $(X, Y) \in B$, we say the event B occurs.
- We write $P[B]$ as a shorthand for $P[(X, Y) \in B]$.

Theorem 4.2

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

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

Example 4.2 Problem

Continuing Example 4.1, 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 4.2 Solution

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

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

Therefore,

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

Quiz 4.2

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:

(1) $P[Q = 0]$

(2) $P[Q = G]$

(3) $P[G > 1]$

(4) $P[G > Q]$

Quiz 4.2 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.

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

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

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

$$P[Q=G] = P_{Q,G}(0,0) + P_{Q,G}(1,1) = 0.18$$

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

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

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

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

4.3 Marginal PMF

- In an experiment that produces two random variables X and Y, it is always possible to consider one of the random variables, Y, and ignore the other one, X.
- If we have already analyzed the experiment to derive the joint PMF $P_{X,Y}(x,y)$, it would be convenient to derive $P_Y(y)$ from $P_{X,Y}(x,y)$ without reexamining the details of the experiment.
- $\{Y = y\}$ is an event, so that $P_Y(y) = P[Y = y]$ is the probability of an event.
- According to Theorem 4.2, it implies We can find $P_Y(y)$ by summing $P_{X,Y}(x,y)$ over all points in $S_{X,Y}$ with the property $Y = y$. In the sum, y is a constant, and each term corresponds to a value of $x \in S_X$.

Theorem 4.3

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

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

- When a random variable X is part of an experiment that produces two random variables, we sometimes refer to its PMF as a **marginal probability mass function**.
- This terminology comes from the matrix representation of the joint PMF. By adding rows and columns and writing the results in the margins, we obtain the marginal PMFs of X and Y.

Example 4.3 Problem

In Example 4.1, we found the joint PMF of X and Y to be

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

Find the marginal PMFs for the random variables X and Y.

Example 4.3 Solution

To find $P_X(x)$, we note that both X and Y have range $\{0, 1, 2\}$. Theorem 4.3 gives

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

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

For the PMF of Y, we obtain

$$P_Y(0) = \sum_{x=0}^{2} P_{X,Y}(x, 0) = 0.10 \qquad P_Y(1) = \sum_{x=0}^{2} P_{X,Y}(x, 1) = 0.09$$

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

Referring to the matrix representation of $P_{X,Y}(x, y)$ in Example 4.1, we observe that each value of $P_X(x)$ is the result of adding all the entries in one row of the matrix. Each value of $P_Y(y)$ is a column sum.

[Continued]

Example 4.3 Solution (continued)

We display $P_X(x)$ and $P_Y(y)$ by rewriting the matrix in Example 4.1 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	

margins

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 2.1(b), which states that the PMF of any random variable must sum to 1. The complete marginal PMF, $P_Y(y)$, appears in the bottom row of the table, and $P_X(x)$ appears in the last column of the table.

$$P_X(x) = \begin{cases} 0.01 & x=0, \\ 0.18 & x=1, \\ 0.81 & x=2, \\ 0 & \text{otherwise.} \end{cases} \qquad P_Y(y) = \begin{cases} 0.1 & y=0, \\ 0.09 & y=1, \\ 0.81 & y=2, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 4.3

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

Quiz 4.3 Solution

By Theorem 4.3, the marginal PMF of H is

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

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

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

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	

4.4 Joint Probability Density Function

Joint Probability Density

Definition 4.3 Function (PDF)

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

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

Notice the substitutions of variables (u, v)

- For two random variables X and Y, the joint PDF $f_{X,Y}(x,y)$ measures probability per unit area. In particular, from the definition of the PDF,

$$P[x < X \leq x + dx, y < Y \leq y + dy] = f_{X,Y}(x, y)dxdy$$

- $f_{X,Y}(x,y)$ is the derivative of the CDF $F_{X,Y}(x,y)$.

Theorem 4.4

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

- **Definition 4.3** and **Theorem 4.4** demonstrate that the joint CDF $F_{X,Y}(x,y)$ and the joint PDF $f_{X,Y}(x,y)$ are equivalent probability models for random variables X and Y.
- PDF is more useful for problem solving. Typically, it is very difficult to use $F_{X,Y}(x,y)$ to calculate the probabilities of events. (reference Theorem 4.5)

Theorem 4.5

$$P[x_1 < X \le x_2, y_1 < Y \le y_2] = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)$$

- The steps need to prove the theorem are outlined in **Problem 4.1.5.** The theorem says that to find the probability that an outcome is in a rectangle, it is necessary to evaluate the joint CDF at all four corners. When the probability of interest corresponds to a nonrectangular area, the joint CDF is much harder to use.
- Of course, not every function $f_{X,Y}(x,y)$ is a valid joint PDF (Example 4.5). Properties (e) and (f) of **Theorem 4.1** for the CDF $F_{X,Y}(x,y)$ imply corresponding properties for the PDF.

Problem 4.1.5

In this problem, we prove Theorem 4.5.

(a) Sketch the following events on the X, Y plane:

$$
\begin{aligned}
A &= \{X \le x_1, y_1 < Y \le y_2\}, \\
B &= \{x_1 < X \le x_2, Y \le y_1\}, \\
C &= \{x_1 < X \le x_2, y_1 < Y \le y_2\}.
\end{aligned}
$$

(b) Express the probability of the events A, B, and $A \cup B \cup C$ in terms of the joint CDF $F_{X,Y}(x, y)$.

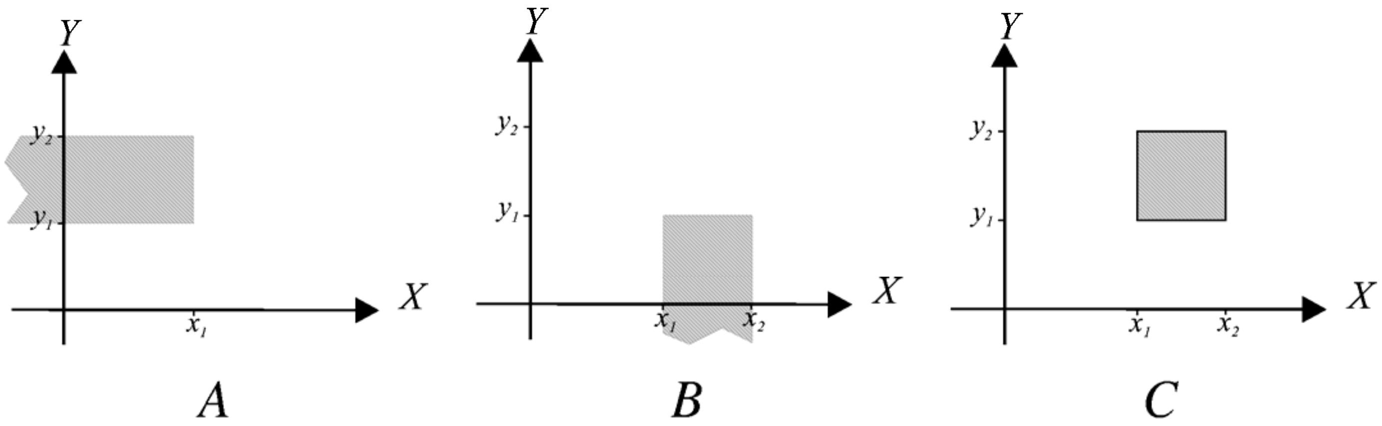
(c) Use the observation that events A, B, and C are mutually exclusive to prove Theorem 4.5.

Problem 4.1.5 Solution

In this problem, we prove Theorem 4.5 which states

$$P[x_1 < X \le x_2, y_1 < Y \le y_2] = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)$$

(a) The events A, B, and C are



(b) In terms of the joint CDF $F_{X,Y}(x, y)$, we can write

$$P[A] = F_{X,Y}(x_1, y_2) - F_{X,Y}(x_1, y_1)$$
$$P[B] = F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_1)$$
$$P[A \cup B \cup C] = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_1)$$

Problem 4.1.5 Solution (continued)

(c) Since A, B, and C are mutually exclusive,

$$P[A \cup B \cup C] = P[A] + P[B] + P[C]$$

However, since we want to express

$$P[C] = P[x_1 < X \leq x_2, y_1 < Y \leq y_2]$$

in terms of the joint CDF $F_{X,Y}(x, y)$, we write

$$\begin{aligned} P[C] &= P[A \cup B \cup C] - P[A] - P[B] \\ &= F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1) \end{aligned}$$

which completes the proof of the theorem.

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

(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 4.7

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

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

Example 4.4 Problem

Random variables X and Y have joint PDF

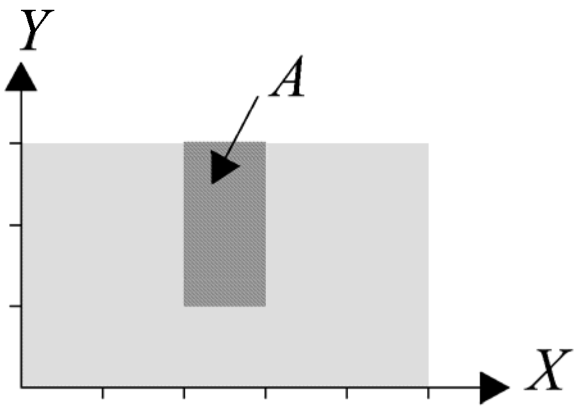
$$f_{X,Y}(x,y) = \begin{cases} c & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases}$$

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

Example 4.4 Solution

The large rectangle in the diagram is the area of nonzero probability. Theorem 4.6 states that the integral of the joint PDF over this rectangle is 1:

$$1 = \int_0^5 \int_0^3 c\, dy\, dx = 15c.$$



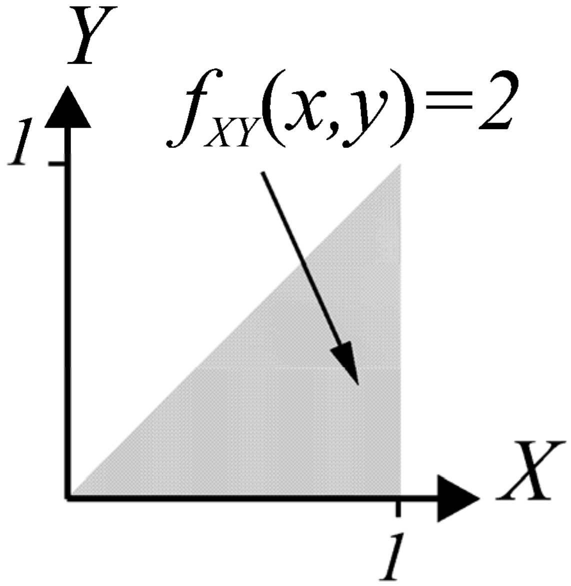
Therefore, $c = 1/15$. The small dark rectangle in the diagram is the event $A = \{2 \le X < 3, 1 \le Y < 3\}$. $P[A]$ is the integral of the PDF over this rectangle, which is

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

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

Example 4.5 Problem

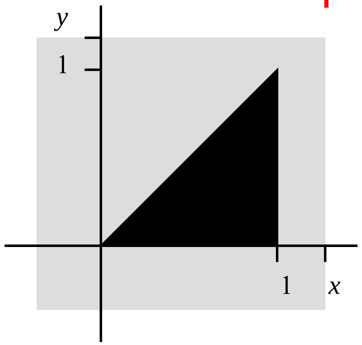
Find the joint CDF $F_{X,Y}(x, y)$ when X and Y have joint PDF



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

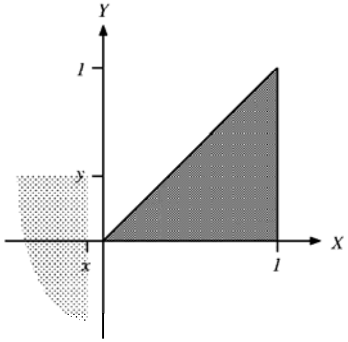
Figure 4.3

Simple cases



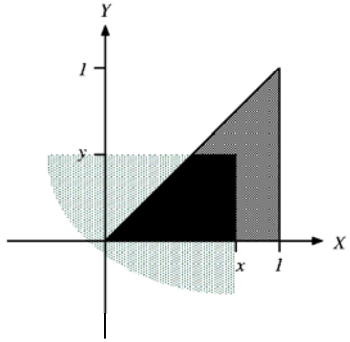
$x > 1$ and $y > 1$

(e)



$x < 0$ or $y < 0$

(a)

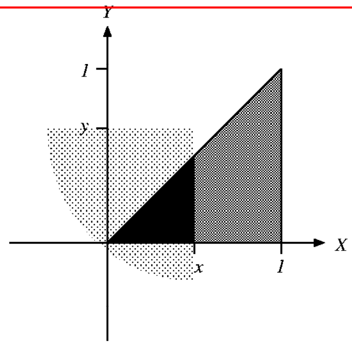


$0 \leq y \leq x \leq 1$

(b)

$$\int_0^y \int_v^x 2dudv$$

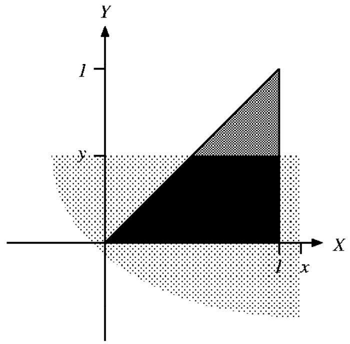
$$\int_0^x \int_v^x 2dudv$$



$0 \leq x < y$

$0 \leq x \leq 1$

(c)



$0 \leq y \leq 1$

$x > 1$

(d)

$$\int_0^y \int_v^1 2dudv$$

Example 4.5 Solution

The joint CDF can be found using Definition 4.3 in which we integrate the joint PDF $f_{X,Y}(x, y)$ over the area shown in Figure 4.1. 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 4.3.

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 4.3. First, we note that with $x < 0$ or $y < 0$, the triangle is completely outside the region of integration as shown in Figure 4.3a. 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 4.3e that the triangle is completely inside the region of integration and we infer from Theorem 4.6 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 4.3b), the integral is

$$F_{X,Y}(x, y) = \int_0^y \int_v^x 2\,du\,dv = 2xy - y^2 \qquad \text{(Figure 4.3b)}.$$

[Continued]

Example 4.5 Solution (continued)

In Figure 4.3c, (x, y) is above the triangle, and the integral is

$$F_{X,Y}(x, y) = \int_0^x \int_v^x 2\,du\,dv = x^2 \qquad \text{(Figure 4.3c).}$$

The remaining situation to consider is shown in Figure 4.3d when (x, y) is to the right of the triangle of nonzero probability, in which case the integral is

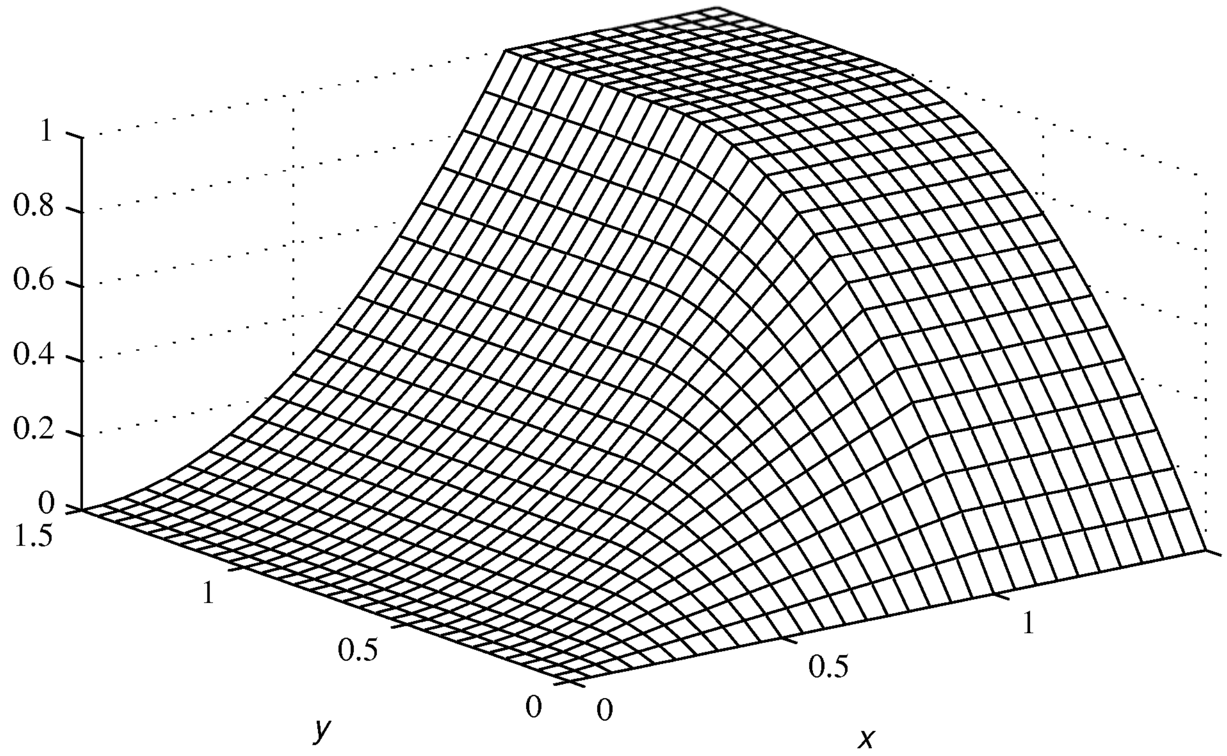
$$F_{X,Y}(x, y) = \int_0^y \int_v^1 2\,du\,dv = 2y - y^2 \qquad \text{(Figure 4.3d)}$$

The resulting CDF, corresponding to the five cases of Figure 4.3, is

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

In Figure 4.4, 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 $f_{X,Y}(x, y)$.

Figure 4.4



A graph of the joint CDF $F_{X,Y}(x, y)$ of Example 4.5.

Comments on Example 4.5

- It takes careful study to verify that $F_{X,Y}(x,y)$ is a valid CDF that satisfies the properties of Theorem 4.1, or even that it is defined for all values x and y.
- Comparing the joint PDF with the joint CDF, we see that
  - The PDF indicates clearly that X, Y occurs with equal probability in all areas of the same size in the triangular region $0 \le y \le x \le 1$.
  - The joint CDF completely hides this simple, important property of the probability model.

Example 4.6 Problem

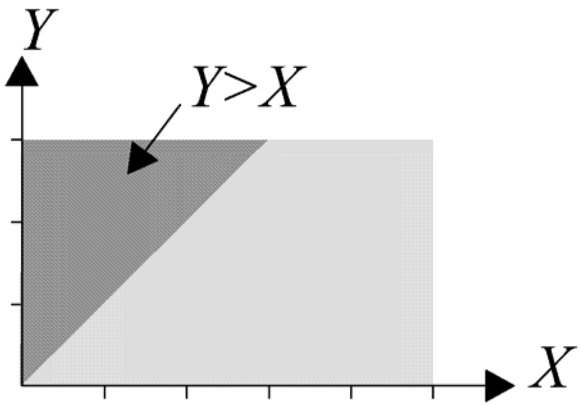
As in Example 4.4, random variables X and Y have joint PDF

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

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

Example 4.6 Solution

Applying Theorem 4.7, we integrate the density $f_{X,Y}(x, y)$ over the part of the X, Y plane satisfying $Y > X$. In this case,



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

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

Quiz 4.4

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

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

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

Quiz 4.4 Solution

To find the constant c, we apply $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$. Specifically,

$$
\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy &= \int_0^2 \int_0^1 cxy\, dx\, dy \\
&= c\int_0^2 y \left(x^2/2 \Big|_0^1 \right) dy \\
&= (c/2)\int_0^2 y\, dy = (c/4) y^2 \Big|_0^2 = c
\end{aligned}
$$

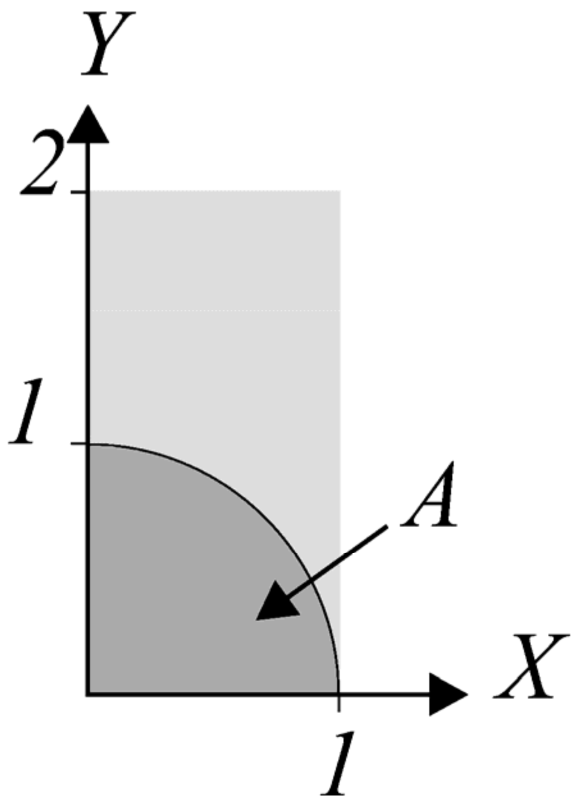
Thus $c = 1$. To calculate $P[A]$, we write

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

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

[Continued]

Quiz 4.4 Solution (continued)



$$P[A] = \int_0^{\pi/2} \int_0^1 r^2 \sin\theta \cos\theta \, r \, dr \, d\theta$$

$$= \left(\int_0^1 r^3 \, dr \right) \left(\int_0^{\pi/2} \sin\theta \cos\theta \, d\theta \right)$$

$$= \left(r^4/4 \Big|_0^1 \right) \left(\frac{\sin^2\theta}{2} \Big|_0^{\pi/2} \right) = 1/8$$

$$\int \sin\theta \cos\theta d\theta = \int \sin\theta d \sin\theta = \frac{\sin^2\theta}{2} + C$$

$$\int \sin\theta \cos\theta d\theta = \frac{1}{2} \int \sin 2\theta d\theta = -\frac{\cos 2\theta}{2} + C$$

4.5 Marginal PDF

For discrete random variables

- PMF $P_{X,Y}(x,y)$ ➔ marginal PMF: $P_X(x)$, $P_Y(y)$

For continuous random variables

- PDF $f_{X,Y}(x,y)$ ➔ marginal PDF: $f_X(x)$, $f_Y(y)$

Theorem 4.8

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 4.8

From the definition of the joint PDF, we can write

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

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 4.7 Problem

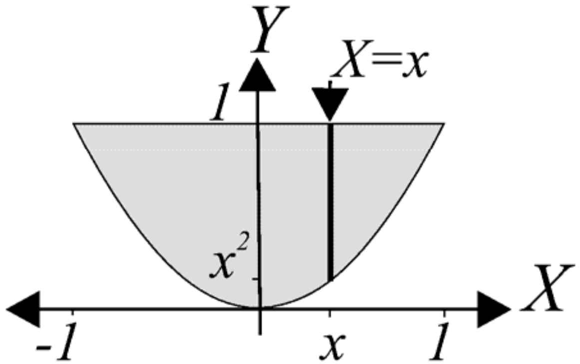
The joint PDF of X and Y is

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

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

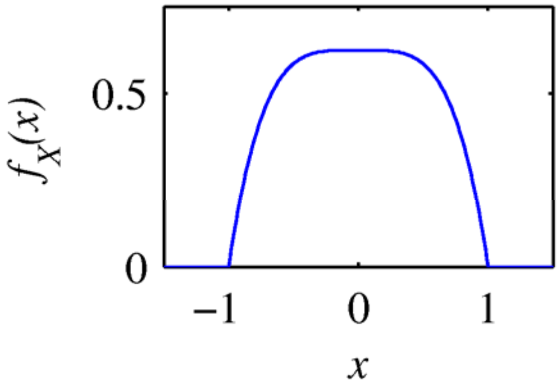
Example 4.7 Solution

We use Theorem 4.8 to find the marginal PDF $f_X(x)$. When $x < -1$ or when $x > 1$, $f_{X,Y}(x, y) = 0$, and therefore $f_X(x) = 0$. For $-1 \le x \le 1$,



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

The complete expression for the marginal PDF of X is

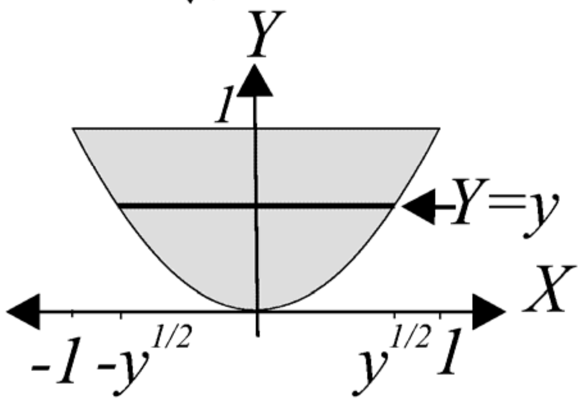


$$f_X(x) = \begin{cases} 5(1 - x^4)/8 & -1 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

[Continued]

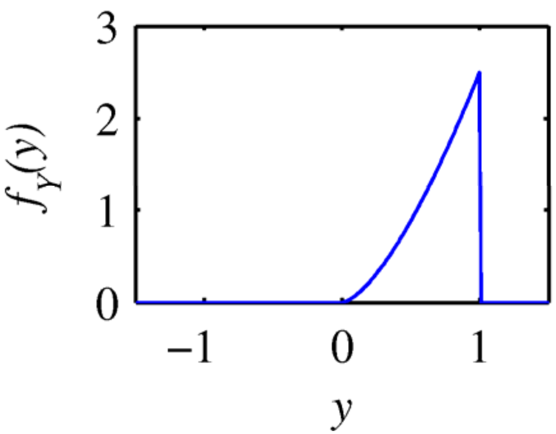
Example 4.7 Solution (continued)

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 $0 \le y \le 1$,



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

The complete marginal PDF of Y is



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

Quiz 4.5

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

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

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

Quiz 4.5 Solution

By Theorem 4.8, the marginal PDF of X is

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

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

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

The complete expression for the PDF of X is

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

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

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

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

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

4.6 Functions of Two Random Variables

Examples

- Example 1:
  - X: the amplitude of the signal transmitted by a radio station as a random variable.
  - Y: the attenuation of the signal as it travels to the antenna of a moving car as another random variable.
  - $W = X / Y$: the amplitude of the signal at the radio receiver in the car.
- Example 2:
  - X, Y: the amplitudes of the signals arriving at the two antennas. The radio receiver connected to the two antennas.
  - Selection diversity combining: $W = X$ if $|X| > |Y|$ and $W = Y$, otherwise.
  - Equal gain combining: $W = X + Y$.
  - Maximal radio combining: $W = aX + bY$.

- Probability model for X and Y
  - PMF: $P_{X,Y}(x,y)$ or PDF: $f_{X,Y}(x,y)$
- $W = g(X, Y)$, how to derive a probability model for W?
- X, Y : discrete random variables (Theorem 4.9)
  - S_W: the range of W, is a countable set corresponding to all possible values of $g(X, Y)$
  - $P_W(w) = P[W = w]$: adding the values of $P_{X,Y}(x,y)$ corresponding to the x, y pairs for which $g(X, Y) = w$.
- X, Y : continuous random variables (Theorem 4.10)
  - $W = g(X, Y)$ is a continuous random variable.
  - $f_W(w)$: find the CDF $F_W(w)$ first and then calculate the derivative.

Theorem 4.9

For discrete random variables X and Y, the derived random variable $W = g(X, Y)$ has PMF

$$P_W(w) = \sum_{(x,y):g(x,y)=w} P_{X,Y}(x, y).$$

Example 4.8 Problem Discrete Random Variables

A firm sends out two kinds of promotional facsimiles. One kind contains only text and requires 40 seconds to transmit each page. The other kind contains grayscale pictures that take 60 seconds per page. Faxes can be 1, 2, or 3 pages long. Let the random variable L represent the length of a fax in pages. $S_L = \{1, 2, 3\}$. Let the random variable T represent the time to send each page. $S_T = \{40, 60\}$. After observing many fax transmissions, the firm derives the following probability model:

$P_{L,T}(l,t)$	$t = 40$ sec	$t = 60$ sec
$l = 1$ page	0.15	0.1
$l = 2$ pages	0.3	0.2
$l = 3$ pages	0.15	0.1

Let $D = g(L, T) = LT$ be the total duration in seconds of a fax transmission. Find the range S_D, the PMF $P_D(d)$, and the expected value $E[D]$.

Example 4.8 Solution

By examining the six possible combinations of L and T we find that the possible values of D are $S_D = \{40, 60, 80, 120, 180\}$. For the five elements of S_D, we find the following probabilities:

$$
\begin{aligned}
&P_D(40) = P_{L,T}(1, 40) = 0.15, && P_D(120) = P_{L,T}(3, 40) + P_{L,T}(2, 60) = 0.35, \\
&P_D(60) = P_{L,T}(1, 60) = 0.1, && P_D(180) = P_{L,T}(3, 60) = 0.1, \\
&P_D(80) = P_{L,T}(2, 40) = 0.3, && P_D(d) = 0; \quad d \neq 40, 60, 80, 120, 180.
\end{aligned}
$$

The expected duration of a fax transmission is

$$
\begin{aligned}
E[D] &= \sum_{d \in S_D} d P_D(d) \\
&= (40)(0.15) + 60(0.1) + 80(0.3) + 120(0.35) + 180(0.1) = 96 \text{ sec.}
\end{aligned}
$$

- When X and Y are continuous random variables and $g(x, y)$ is a continuous function, $W = g(X, Y)$ is a continuous random variable.
- To find the PDF, $f_W(w)$, it is usually helpful to first find the CDF $F_W(w)$ and then calculate the derivative. Viewing $\{W \leq w\}$ as an event *A*, we can apply **Theorem 4.7**.

Theorem 4.10

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

$$F_W(w) = P[W \leq w] = \iint_{g(x,y) \leq w} f_{X,Y}(x, y)\, dx\, dy.$$

- Once we obtain the CDF $F_W(w)$, it is generally straightforward to calculate the derivative $f_W(w) = dF_W(w) / dw$.
- However, for most functions $g(x, y)$, performing the integration to find $F_W(w)$ can be a tedious process. Fortunately, there are convenient techniques for finding $f_W(w)$ for certain functions that arise in many applications.
- The most important function $g(X, Y) = X + Y$, is the subject of Chapter 6.
- Another interesting function is the maximum of two random variables. The following theorem follows from the observation that $\{W \leq w\} = \{X \leq w\} \cap \{Y \leq w\}$

Theorem 4.11

For continuous random variables X and Y, the CDF of $W = \max(X, Y)$ is

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

Example 4.9 Problem

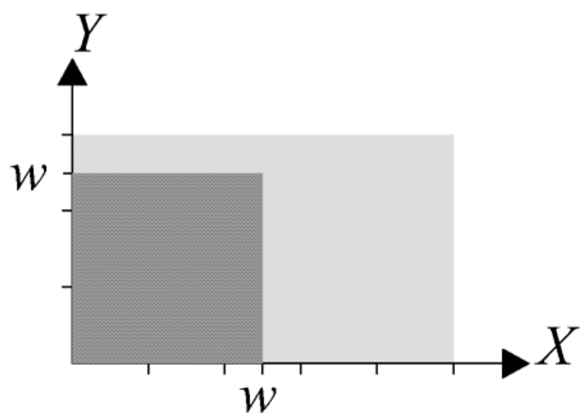
In Examples 4.4 and 4.6, X and Y have joint PDF

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

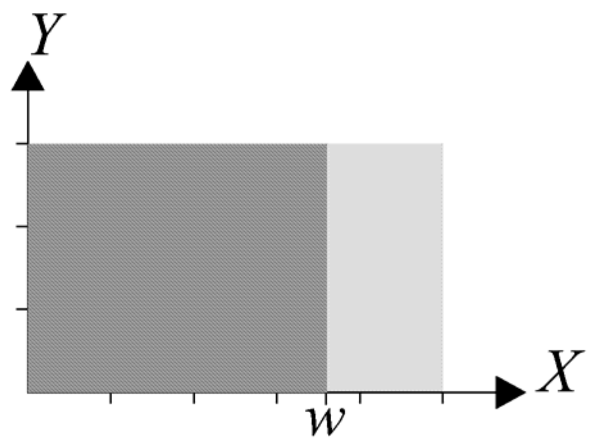
Find the PDF of $W = \max(X, Y)$.

Example 4.9 Solution

Because $X \geq 0$ and $Y \geq 0$, $W \geq 0$. Therefore, $F_W(w) = 0$ for $w < 0$. Because $X \leq 5$ and $Y \leq 3$, $W \leq 5$. Thus $F_W(w) = 1$ for $w \geq 5$. For $0 \leq w \leq 5$, diagrams provide a guide to calculating $F_W(w)$. Two cases, $0 \leq w \leq 3$ and $3 \leq w \leq 5$, are shown here:



$0 \leq w \leq 3$



$3 \leq w \leq 5$

[Continued]

Example 4.9 Solution (continued)

When $0 \leq w \leq 3$, Theorem 4.11 yields

$$F_W(w) = \int_0^w \int_0^w \frac{1}{15}\,dx\,dy = w^2/15.$$

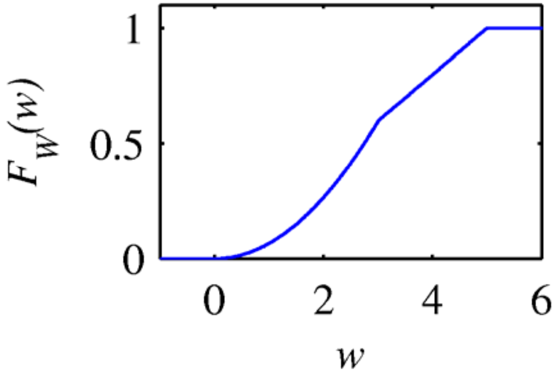
Because the joint PDF is uniform, we see this probability is just the area w^2 times the value of the joint PDF over that area. When $3 \leq w \leq 5$, the integral over the region $\{X \leq w, Y \leq w\}$ becomes

$$F_W(w) = \int_0^w \left(\int_0^3 \frac{1}{15}\,dy \right) dx = \int_0^w \frac{1}{5}\,dx = w/5,$$

which is the area $3w$ times the value of the joint PDF over that area. [Continued]

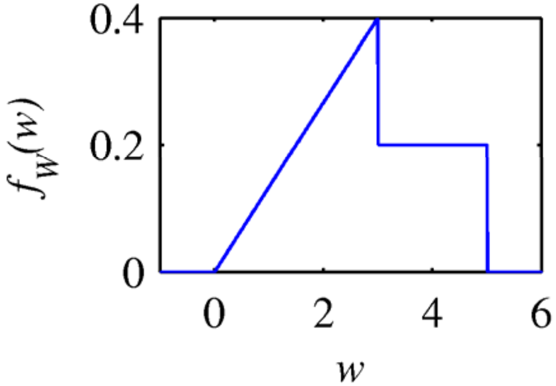
Example 4.9 Solution (continued)

Combining the parts, we can write the joint CDF:



$$F_W(w) = \begin{cases} 0 & w < 0, \\ w^2/15 & 0 \le w \le 3, \\ w/5 & 3 < w \le 5, \\ 1 & w > 5. \end{cases}$$

By taking the derivative, we find the corresponding joint PDF:



$$f_W(w) = \begin{cases} 2w/15 & 0 \le w \le 3, \\ 1/5 & 3 < w \le 5, \\ 0 & \text{otherwise.} \end{cases}$$

Example 4.10 Problem

X and Y have the joint PDF

$$f_{X,Y}(x, y) = \begin{cases} \lambda\mu e^{-(\lambda x+\mu y)} & x \geq 0, y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

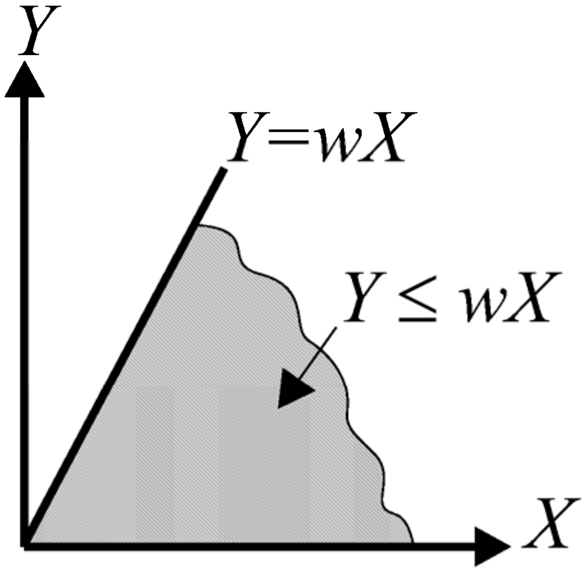
Find the PDF of $W = Y/X$.

Example 4.10 Solution

First we find the CDF:

$$F_W(w) = P[Y/X \le w] = P[Y \le wX].$$

For $w < 0$, $F_W(w) = 0$. For $w \ge 0$, we integrate the joint PDF $f_{X,Y}(x,y)$ over the region of the X, Y plane for which $Y \le wX$, $X \ge 0$, and $Y \ge 0$ as shown:



$$\begin{aligned} P[Y \le wX] &= \int_0^\infty \left(\int_0^{wx} f_{X,Y}(x,y)\, dy \right) dx \\ &= \int_0^\infty \lambda e^{-\lambda x} \left(\int_0^{wx} \mu e^{-\mu y}\, dy \right) dx \\ &= \int_0^\infty \lambda e^{-\lambda x} \left(1 - e^{-\mu w x}\right) dx \\ &= 1 - \frac{\lambda}{\lambda + \mu w} \end{aligned}$$

[Continued]

Example 4.10 Solution (continued)

Therefore,

$$F_W(w) = \begin{cases} 0 & w < 0, \\ 1 - \frac{\lambda}{\lambda + \mu w} & \omega \geq 0. \end{cases}$$

Differentiating with respect to w, we obtain

$$f_W(w) = \begin{cases} \lambda\mu/(\lambda + \mu w)^2 & w \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 4.6(A)

Two computers use modems and a telephone line to transfer e-mail and Internet news every hour. At the start of a data call, the modems at each end of the line negotiate a speed that depends on the line quality. When the negotiated speed is low, the computers reduce the amount of news that they transfer. The number of bits transmitted L and the speed B in bits per second have the joint PMF

$P_{L,B}(l, b)$	$b = 14,400$	$b = 21,600$	$b = 28,800$
$l = 518,400$	0.2	0.1	0.05
$l = 2,592,000$	0.05	0.1	0.2
$l = 7,776,000$	0	0.1	0.2

Let T denote the number of seconds needed for the transfer. Express T as a function of L and B. What is the PMF of T?

Quiz 4.6(A) Solution

The time required for the transfer is $T = L/B$. For each pair of values of L and B, we can calculate the time T needed for the transfer. We can write these down on the table for the joint PMF of L and B as follows:

$P_{L,B}(l,b)$	$b = 14,400$	$b = 21,600$	$b = 28,800$
$l = 518,400$	0.20 $(T{=}36)$	0.10 $(T{=}24)$	0.05 $(T{=}18)$
$l = 2,592,000$	0.05 $(T{=}180)$	0.10 $(T{=}120)$	0.20 $(T{=}90)$
$l = 7,776,000$	0.00 $(T{=}540)$	0.10 $(T{=}360)$	0.20 $(T{=}270)$

From the table, writing down the PMF of T is straightforward.

$$P_T(t) = \begin{cases} 0.05 & t = 18 \\ 0.1 & t = 24 \\ 0.2 & t = 36, 90 \\ 0.1 & t = 120 \\ 0.05 & t = 180 \\ 0.2 & t = 270 \\ 0.1 & t = 360 \\ 0 & \text{otherwise} \end{cases}$$

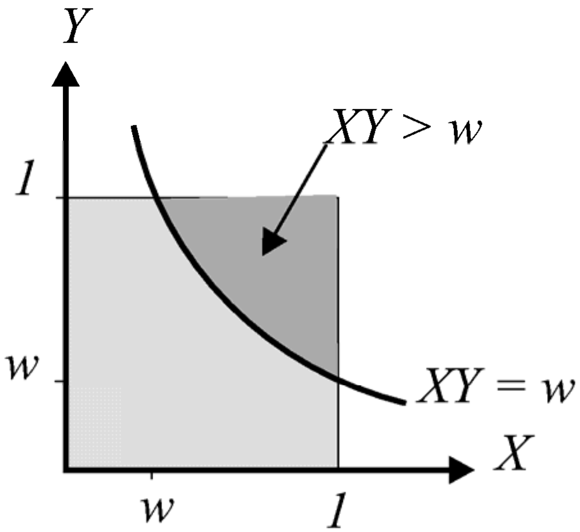
Quiz 4.6(B)

Find the CDF and the PDF of $W = XY$ when random variables X and Y have joint PDF

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

Quiz 4.6(B) Solution

First, we observe that since $0 \leq X \leq 1$ and $0 \leq Y \leq 1$, $W = XY$ satisfies $0 \leq W \leq 1$. Thus $f_W(0) = 0$ and $f_W(1) = 1$. For $0 < w < 1$, we calculate the CDF $F_W(w) = P[W \leq w]$. As shown below, integrating over the region $W \leq w$ is fairly complex. The calculus is simpler if we integrate over the region $XY > w$. Specifically,



$$
\begin{aligned}
F_W(w) &= 1 - P[XY > w] \\
&= 1 - \int_w^1 \int_{w/x}^1 dy\, dx \\
&= 1 - \int_w^1 (1 - w/x)\, dx \\
&= 1 - \left(x - w \ln x \Big|_{x=w}^{x=1} \right) \\
&= 1 - (1 - w + w \ln w) = w - w \ln w
\end{aligned}
$$

[Continued]

Quiz 4.6(B) Solution (continued)

The complete expression for the CDF is

$$F_W(w) = \begin{cases} 0 & w < 0 \\ w - w \ln w & 0 \le w \le 1 \\ 1 & w > 1 \end{cases}$$

By taking the derivative of the CDF, we find the PDF is

$$f_W(w) = \frac{dF_W(w)}{dw} = \begin{cases} 0 & w < 0 \\ -\ln w & 0 \le w \le 1 \\ 0 & w > 1 \end{cases}$$

4.7 Expected Values

- There are many situations in which we are interested only in the expected value of a derived random variable $W = g(X, Y)$, not the entire probability model.
- In these situations, we can obtain the expected value directly from $P_{X,Y}(x, y)$ or $f_{X,Y}(x, y)$ without taking the trouble to compute $P_W(w)$ or $f_W(w)$.

Theorem 4.12

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

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

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

- Theorem 2.10 $\quad E[Y] = \mu_Y = \sum_{x \in S_X} g(x) P_X(x)$
- Theorem 3.4 $\quad E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

Example 4.11 Problem

In Example 4.8, compute $E[D]$ directly from $P_{L,T}(l,t)$.

Example 4.11 Solution

Applying Theorem 4.12 to the discrete random variable D, we obtain

$$
\begin{aligned}
E[D] &= \sum_{l=1}^{3} \sum_{t=40,60} lt P_{L,T}(l,t) \\
&= (1)(40)(0.15) + (1)60(0.1) + (2)(40)(0.3) + (2)(60)(0.2) \\
&\quad + (3)(40)(0.15) + (3)(60)(0.1) = 96 \text{ sec},
\end{aligned}
$$

which is the same result obtained in Example 4.8 after calculating $P_D(d)$.

Theorem 4.13

$$E[g_1(X,Y)+\cdots+g_n(X,Y)] = E[g_1(X,Y)]+\cdots+E[g_n(X,Y)].$$

Proof: Theorem 4.13

Let $g(X,Y) = g_1(X,Y)+\cdots+g_n(X,Y)$. For discrete random variables X, Y, Theorem 4.12 states

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

We can break the double summation into n double summations:

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

By Theorem 4.12, the ith double summation on the right side is $E[g_i(X,Y)]$, thus

$$E[g(X,Y)] = E[g_1(X,Y)]+\cdots+E[g_n(X,Y)].$$

For continuous random variables, Theorem 4.12 says

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

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

Theorem 4.14

For any two random variables X and Y,

$$E[X + Y] = E[X] + E[Y].$$

- An important consequence of this theorem is that we can find the expected sum of two random variables from the separate models: $P_X(x)$ and $P_Y(y)$ or $f_X(x)$ and $f_Y(y)$. We do not need a complete probability model embodied in $P_{X,Y}(x, y)$ or $f_{X,Y}(x, y)$.
- By contrast, the variance of $X + Y$ depends on the entire joint PMF or joint CDF.

Theorem 4.15

The variance of the sum of two random variables is

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2E[(X-\mu_X)(Y-\mu_Y)].$$

Proof: Theorem 4.15

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

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

We observe that each of the three terms in the preceding expected values is a function of X and Y. Therefore, Theorem 4.13 implies

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

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

Definition 4.4 Covariance

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

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

$$\sigma_{XY} = \mathrm{Cov}\left[X, Y\right] = E\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right]$$

Definition 4.5 Correlation

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

Theorem 4.16

(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 $r_{X,Y} = E[X^2] = E[Y^2]$.

Proof: Theorem 4.16

Cross-multiplying inside the expected value of Definition 4.4 yields

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

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

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

Note that in the expression $E[\mu_Y X]$, μ_Y is a constant. Referring to Theorem 2.12, we set $a = \mu_Y$ and $b = 0$ to obtain $E[\mu_Y X] = \mu_Y E[X] = \mu_Y \mu_X$. The same reasoning demonstrates that $E[\mu_X Y] = \mu_X E[Y] = \mu_X \mu_Y$. Therefore,

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

The other relationships follow directly from the definitions and Theorem 4.15.

Example 4.12 Problem

For the integrated circuits tests in Example 4.1, we found in Example 4.3 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 $\mathrm{Cov}[X, Y]$.

Example 4.12 Solution

By Definition 4.5,

$$
\begin{aligned}
r_{X,Y} = E[XY] &= \sum_{x=0}^{2}\sum_{y=0}^{2} xy P_{X,Y}(x, y) \\
&= (1)(1)0.09 + (2)(2)0.81 = 3.33.
\end{aligned}
$$

To use Theorem 4.16(a) to find the covariance, we find

$$
\begin{aligned}
E[X] &= (1)(0.18) + (2)(0.81) = 1.80, \\
E[Y] &= (1)(0.09) + (2)(0.81) = 1.71.
\end{aligned}
$$

Therefore, by Theorem 4.16(a), $\text{Cov}[X, Y] = 3.33 - (1.80)(1.71) = 0.252$.

Definition 4.6 Orthogonal Random Variables

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

Definition 4.7 Uncorrelated Random Variables

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

- **Orthogonal** means zero correlation, while **uncorrelated** means zero covariance

Definition 4.8 Correlation Coefficient

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

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

- Note that the units of the covariance and the correlation are the product of the units of X and Y. Thus if X has units of kilograms and Y has units of seconds, then $\mathrm{Cov}[X, Y]$ and $r_{X,Y}$ have units of kilograms-seconds.
- By contrast, $\rho_{X,Y}$ is a **dimensionless** quantity.

Theorem 4.17

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

Proof: Theorem 4.17

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,

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

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

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

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

Comments on Correlation Coefficient

- $\rho_{X,Y}$ describes the information we gain about Y by observing X.
- $\rho_{X,Y} > 0$, suggests that when X is high relative to its expected value, Y also tends to be high, and when X is low, Y is likely to be low.
- $\rho_{X,Y} < 0$, suggests that a high value of X is likely to be accompanied by a low value of Y and that a low value of X is likely to be accompanied by a high value of Y.
- A linear relationship between X and Y produces the extreme values, $\rho_{X,Y} = \pm 1$.

Theorem 4.18

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

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

- The proof is in Problem 4.7.7.

Examples of Correlation Coefficient

- X is the height of a student. Y is the weight of the same student. $0 < \rho_{X,Y} < 1$.
- 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 resister measured in degrees Kelvin. $\rho_{X,Y} = 1$.
- 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. $\rho_{X,Y} = 0$.

Problem 4.7.7

For a random variable X, let $Y = aX + b$. Show that if $a > 0$ then $\rho_{X,Y} = 1$. Also show that if $a < 0$, then $\rho_{X,Y} = -1$.

Problem 4.7.7 Solution

First, we observe that Y has mean $\mu_Y = a\mu_X + b$ and variance $\text{Var}[Y] = a^2 \text{Var}[X]$. The covariance of X and Y is

$$\text{Cov}[X, Y] = E[(X - \mu_X)(aX + b - a\mu_X - b)] = aE\left[(X - \mu_X)^2\right] = a \text{Var}[X]$$

The correlation coefficient is

$$\rho_{X,Y} = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}} = \frac{a \text{Var}[X]}{\sqrt{\text{Var}[X]}\sqrt{a^2 \text{Var}[X]}} = \frac{a}{|a|}$$

When $a > 0$, $\rho_{X,Y} = 1$. When $a < 0$, $\rho_{X,Y} = 1$.

Quiz 4.7(A)

Random variables L and T given in Example 4.8 have joint PMF

$P_{L,T}(l,t)$	$t = 40$ sec	$t = 60$ 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.

(1) $E[L]$ and $\mathrm{Var}[L]$

(2) $E[T]$ and $\mathrm{Var}[T]$

(3) The correlation $r_{L,T} = E[LT]$

(4) The covariance $\mathrm{Cov}[L,T]$

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

Quiz 4.7(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	

(1) The expected value of L is

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

Since the second moment of L is

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

the variance of L is

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

(2) The expected value of T is

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

The second moment of T is

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

[Continued]

Quiz 4.7(A) Solution (continued)

Thus

$$\text{Var}[T] = E\left[T^2\right] - (E[T])^2 = 2400 - 48^2 = 96.$$

(3) The correlation is

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

(4) From Theorem 4.16(a), the covariance of L and T is

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

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

Quiz 4.7(B)

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

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

Find the following quantities.

(1) $E[X]$ and $\text{Var}[X]$

(2) $E[Y]$ and $\text{Var}[Y]$

(3) The correlation $r_{X,Y} = E[XY]$

(4) The covariance $\text{Cov}[X, Y]$

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

Quiz 4.7(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 $0 \le x \le 1$,

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

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

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

The complete expressions for the marginal PDFs are

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} y/2 & 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$$

From the marginal PDFs, it is straightforward to calculate the various expectations.

(1) The first and second moments of X are

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

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

[Continued]

Quiz 4.7(B) Solution (continued)

The variance of X is $\text{Var}[X] = E[X^2] - (E[X])^2 = 1/18$.

(2) The first and second moments of Y are

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

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

The variance of Y is $\text{Var}[Y] = E[Y^2] - (E[Y])^2 = 2 - 16/9 = 2/9$.

(3) The correlation of X and Y is

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

(4) The covariance of X and Y is

$$\text{Cov}[X, Y] = E[XY] - E[X]\, E[Y] = \frac{8}{9} - \left(\frac{2}{3}\right)\left(\frac{4}{3}\right) = 0.$$

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

4.8 Conditioning by an Event

- An experiment produces two random variables, X and Y. Outcome (x, y) is an element of an event, B. We use the information $(x, y) \in B$ to construct a new probability model.
- If X and Y are discrete, the new model is a conditional joint PMF, the ratio of the joint PMF to $P[B]$.
- If X and Y are continuous, the new model is a conditional joint PDF, defined as the ratio of the join PDF to $P[B]$.
- Same intuition as **Definition 1.6**.
- Section 4.9 considers the special case of an event that corresponds to an observation of one of the two random variables: either $B = \{X = x\}$, or $B = \{Y = y\}$.

Definition 4.9 *Conditional Joint PMF*

For discrete random variables X *and* Y *and an event,* B *with* $P[B] > 0$*, the* conditional joint PMF *of* X *and* Y *given* B *is*

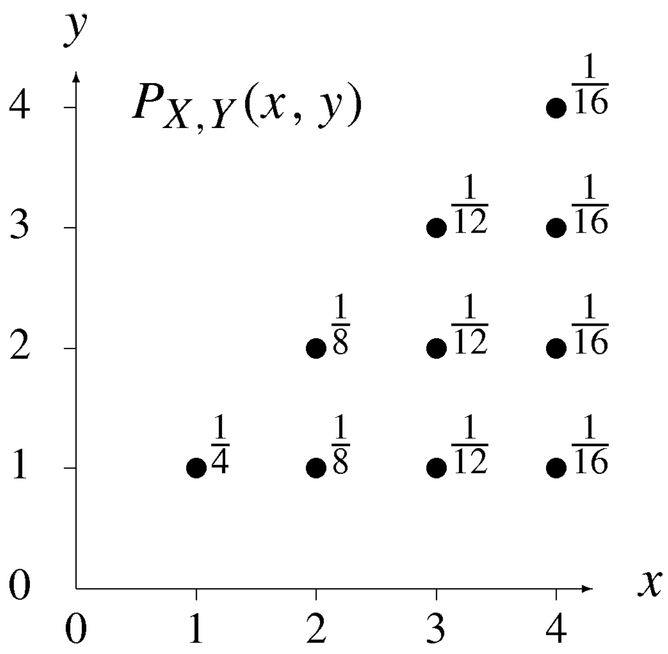
$$P_{X,Y|B}(x, y) = P[X = x, Y = y|B].$$

Theorem 4.19

For any event B, a region of the X, Y plane with $P[B] > 0$,

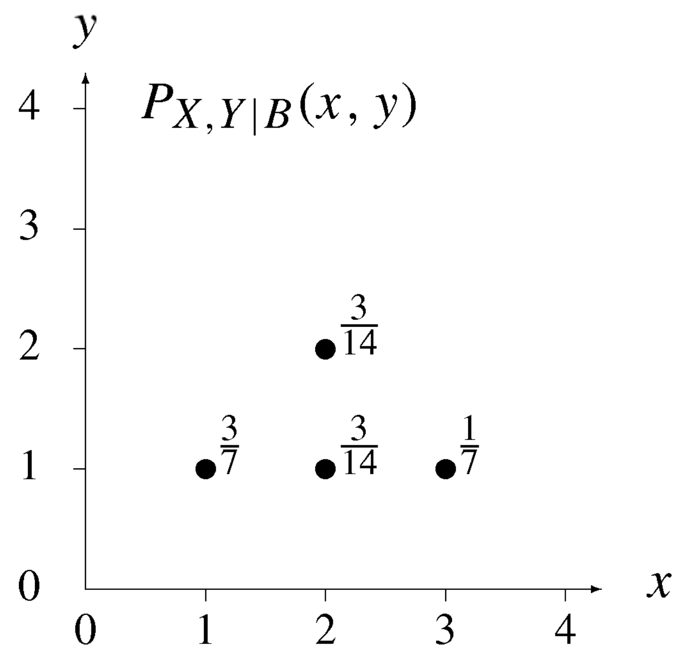
$$P_{X,Y|B}(x, y) = \begin{cases} \dfrac{P_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Example 4.13 Problem



Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$ as shown. Let B denote the event $X + Y \leq 4$. Find the conditional PMF of X and Y given B.

Example 4.13 Solution



Event $B = \{(1, 1), (2, 1), (2, 2), (3, 1)\}$ consists of all points (x, y) such that $x + y \le 4$. By adding up the probabilities of all outcomes in B, we find

$$P[B] = P_{X,Y}(1, 1) + P_{X,Y}(2, 1) + P_{X,Y}(2, 2) + P_{X,Y}(3, 1) = \frac{7}{12}.$$

The conditional PMF $P_{X,Y|B}(x, y)$ is shown on the left.

Definition 4.10 Conditional Joint PDF

Given an event B with $P[B] > 0$, the conditional joint probability density function of X and Y is

$$f_{X,Y|B}(x, y) = \begin{cases} \dfrac{f_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \textit{otherwise.} \end{cases}$$

Example 4.14 Problem

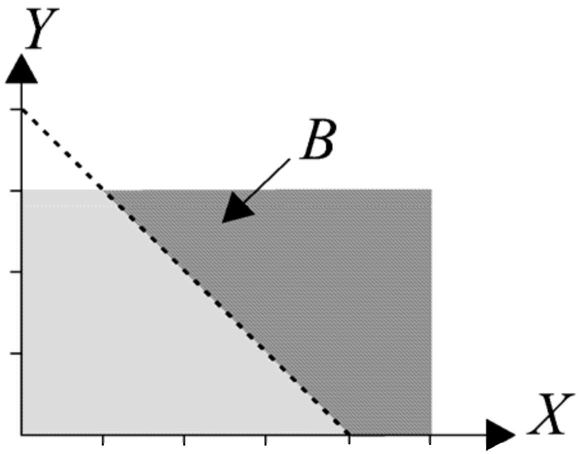
X and Y are random variables with joint PDF

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

Find the conditional PDF of X and Y given the event $B = \{X + Y \ge 4\}$.

Example 4.14 Solution

We calculate $P[B]$ by integrating $f_{X,Y}(x, y)$ over the region B.



$$P[B] = \int_0^3 \int_{4-y}^5 \frac{1}{15}\,dx\,dy = \frac{1}{15}\int_0^3 (1+y)\,dy = 1/2.$$

Definition 4.10 leads to the conditional joint PDF

$$f_{X,Y|B}(x, y) = \begin{cases} 2/15 & 0 \le x \le 5, 0 \le y \le 3, x + y \ge 4, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 4.20 Conditional Expected Value

For random variables X and Y and an event B of nonzero probability, the conditional expected value of $W = g(X, Y)$ given B is

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

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

Definition 4.11 Conditional variance

The conditional variance *of the random variable* $W = g(X, Y)$ *is*

$$\operatorname{Var}[W|B] = E\left[\left(W - \mu_{W|B}\right)^2 |B\right].$$

$$\sigma_{W|B}^2 = \operatorname{Var}[W \mid B] = E\left[\left(W - \mu_{W|B}\right)^2 \middle| B\right]$$

Theorem 4.21

$$\operatorname{Var}[W|B] = E\left[W^2|B\right] - (\mu_{W|B})^2.$$

Example 4.15 Problem

Continuing Example 4.13, find the conditional expected value and the conditional variance of $W = X + Y$ given the event $B = \{X + Y \leq 4\}$.

Example 4.15 Solution

We recall from Example 4.13 that $P_{X,Y|B}(x, y)$ has four points with nonzero probability: $(1, 1)$, $(1, 2)$, $(1, 3)$, and $(2, 2)$. Their probabilities are $3/7$, $3/14$, $1/7$, and $3/14$, respectively. Therefore,

$$E[W|B] = \sum_{x,y} (x + y) P_{X,Y|B}(x, y)$$
$$= 2\frac{3}{7} + 3\frac{3}{14} + 4\frac{1}{7} + 4\frac{3}{14} = \frac{41}{14}.$$

Similarly,

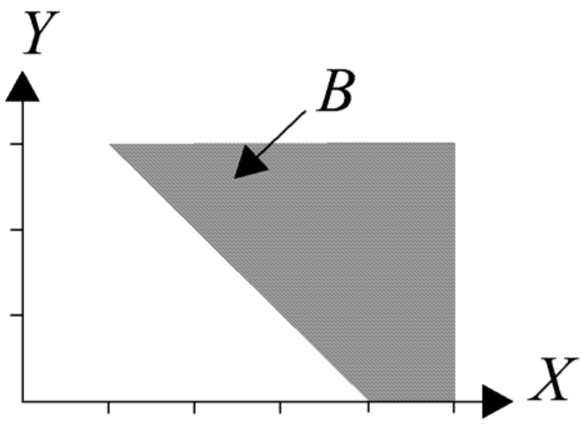
$$E\left[W^2|B\right] = \sum_{x,y} (x + y)^2 P_{X,Y|B}(x, y)$$
$$= 2^2\frac{3}{7} + 3^2\frac{3}{14} + 4^2\frac{1}{7} + 4^2\frac{3}{14} = \frac{131}{14}.$$

The conditional variance is $\text{Var}[W|B] = E[W^2|B] - (E[W|B])^2 = (131/14) - (41/14)^2 = 153/196$.

Example 4.16 Problem

Continuing Example 4.14, find the conditional expected value of $W = XY$ given the event $B = \{X + Y \geq 4\}$.

Example 4.16 Solution



For the event B shown in the adjacent graph, Example 4.14 showed that the conditional PDF of X, Y given B is

$$f_{X,Y|B}(x, y) = \begin{cases} 2/15 & 0 \le x \le 5, 0 \le y \le 3, (x, y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

From Theorem 4.20,

$$\begin{aligned} E[XY|B] &= \int_0^3 \int_{4-y}^5 \frac{2}{15} xy \, dx \, dy \\ &= \frac{1}{15} \int_0^3 \left(x^2 \Big|_{4-y}^5 \right) y \, dy \\ &= \frac{1}{15} \int_0^3 \left(9y + 8y^2 - y^3 \right) dy = \frac{123}{20}. \end{aligned}$$

Quiz 4.8(A)

From Example 4.8, random variables L and T have joint PMF

$P_{L,T}(l,t)$	$t = 40$ sec	$t = 60$ sec
$l = 1$ page	0.15	0.1
$l = 2$ pages	0.3	0.2
$l = 3$ pages	0.15	0.1

For random variable $V = LT$, we define the event $A = \{V > 80\}$. Find the conditional PMF $P_{L,T|A}(l,t)$ of L and T given A. What are $E[V|A]$ and $\mathrm{Var}[V|A]$?

Quiz 4.8(A) Solution

Since the event $V > 80$ occurs only for the pairs $(L, T) = (2, 60)$, $(L, T) = (3, 40)$ and $(L, T) = (3, 60)$,

$$P[A] = P[V > 80] = P_{L,T}(2, 60) + P_{L,T}(3, 40) + P_{L,T}(3, 60) = 0.45$$

By Definition 4.9,

$$P_{L,T|A}(l, t) = \begin{cases} \frac{P_{L,T}(l,t)}{P[A]} & lt > 80 \\ 0 & \text{otherwise} \end{cases}$$

We can represent this conditional PMF in the following table:

$P_{L,T\|A}(l, t)$	$t = 40$	$t = 60$
$l = 1$	0	0
$l = 2$	0	4/9
$l = 3$	1/3	2/9

[Continued]

Quiz 4.8(A) Solution (continued)

The conditional expectation of V can be found from the conditional PMF.

$$E[V|A] = \sum_l \sum_t lt\, P_{L,T|A}(l,t)$$

$$= (2 \cdot 60)\frac{4}{9} + (3 \cdot 40)\frac{1}{3} + (3 \cdot 60)\frac{2}{9} = 133\,\frac{1}{3}$$

For the conditional variance $\mathrm{Var}[V|A]$, we first find the conditional second moment

$$E\left[V^2|A\right] = \sum_l \sum_t (lt)^2 P_{L,T|A}(l,t)$$

$$= (2 \cdot 60)^2\frac{4}{9} + (3 \cdot 40)^2\frac{1}{3} + (3 \cdot 60)^2\frac{2}{9} = 18,400$$

It follows that

$$\mathrm{Var}[V|A] = E\left[V^2|A\right] - (E[V|A])^2 = 622\,\frac{2}{9}$$

Quiz 4.8(B)

Random variables X and Y have the joint PDF

$$f_{X,Y}(x, y) = \begin{cases} xy/4000 & 1 \leq x \leq 3, 40 \leq y \leq 60, \\ 0 & \text{otherwise.} \end{cases}$$

For random variable $W = XY$, we define the event $B = \{W > 80\}$. Find the conditional joint PDF $f_{X,Y|B}(l, t)$ of X and Y given B. What are $E[W|B]$ and $\text{Var}[W|B]$?

Quiz 4.8(B) Solution

For continuous random variables X and Y, we first calculate the probability of the conditioning event.

$$P[B] = \iint_B f_{X,Y}(x, y)\, dx\, dy = \int_{40}^{60} \int_{80/y}^{3} \frac{xy}{4000}\, dx\, dy$$

A little calculus yields

$$P[B] = \int_{40}^{60} \frac{y}{4000} \left(\left. \frac{x^2}{2} \right|_{80/y}^{3} \right) dy = \int_{40}^{60} \frac{y}{4000} \left(\frac{9}{2} - \frac{3200}{y^2} \right) dy = \frac{9}{8} - \frac{4}{5} \ln \frac{3}{2}$$

In fact, $P[B] \approx 0.801$. The conditional PDF of X and Y is

$$\begin{aligned} f_{X,Y|B}(x, y) &= \begin{cases} f_{X,Y}(x, y)/P[B] & (x, y) \in B \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} Kxy & 40 \le y \le 60, 80/y \le x \le 3 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

where $K = (4000P[B])^{-1}$. The conditional expectation of W given event B is [Continued]

Quiz 4.8(B) Solution (continued)

$$E\left[W|B\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y|B}\left(x, y\right)\, dx\, dy = \int_{40}^{60}\int_{80/y}^{3} Kx^2y^2\, dx\, dy.$$

These next steps are just calculus:

$$E\left[W|B\right] = (K/3)\int_{40}^{60} y^2x^3\Big|_{x=80/y}^{x=3}\, dy = (K/3)\int_{40}^{60}\left(27y^2 - 80^3/y\right)\, dy$$

$$= (K/3)\left(9y^3 - 80^3\ln y\right)\Big|_{40}^{60} \approx 120.78$$

The conditional second moment of K given B is

$$E\left[W^2|B\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (xy)^2 f_{X,Y|B}\left(x, y\right)\, dx\, dy = \int_{40}^{60}\int_{80/y}^{3} Kx^3y^3\, dx\, dy$$

With a final bit of calculus,

$$E\left[W^2|B\right] = (K/4)\int_{40}^{60} y^3x^4\Big|_{x=80/y}^{x=3}\, dy = (K/4)\int_{40}^{60}\left(81y^3 - 80^4/y\right)\, dy$$

$$= (K/4)\left((81/4)y^4 - 80^4\ln y\right)\Big|_{40}^{60} \approx 16{,}116.10$$

It follows that $\mathrm{Var}[W|B] = E[W^2|B] - (E[W|B])^2 \approx 1528.30$.

4.9 Conditioning by a Random Variable

- Discuss the special case in which the partial knowledge consists of the value of one of the random variables: either $B = \{X = x\}$ or $B = \{Y = y\}$

- Learning $\{Y = y\}$ changes our knowledge of random variables X, Y. We now have complete knowledge of Y and modified knowledge of X.

- The new model is either a conditional PMF of X given Y or a conditional PDF of X given Y.
  - X, Y: discrete. $P_{X,Y|B}(x, y)$ and $E[g(X, Y)|B]$
  - X, Y: continuous. We can not apply section 4.8 directly because $P[B] = P[Y = y] = 0$. We define a conditional PDF as **the ratio of the joint PDF to the marginal PDF**.

Definition 4.12 Conditional PMF

For any event $Y = y$ *such that* $P_Y(y) > 0$*, the* conditional PMF *of* X *given* $Y = y$ *is*

$$P_{X|Y}(x|y) = P[X = x|Y = y].$$

Theorem 4.22

For random variables X and Y with joint PMF $P_{X,Y}(x, y)$, and x and y such that $P_X(x) > 0$ and $P_Y(y) > 0$,

$$P_{X,Y}(x, y) = P_{X|Y}(x|y)\, P_Y(y) = P_{Y|X}(y|x)\, P_X(x) .$$

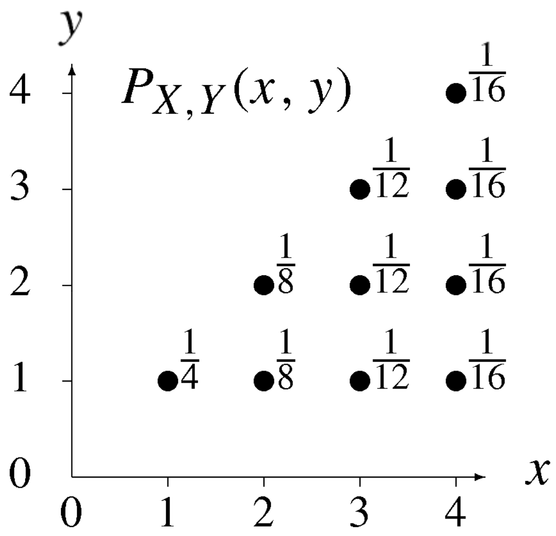
Proof: Theorem 4.22

Referring to Definition 4.12, Definition 1.6, and Theorem 4.3, we observe that

$$P_{X|Y}(x|y) = P[X = x|Y = y] = \frac{P[X = x, Y = y]}{P[Y = y]} = \frac{P_{X,Y}(x, y)}{P_Y(y)}.$$

Hence, $P_{X,Y}(x, y) = P_{X|Y}(x|y) P_Y(y)$. The proof of the second part is the same with X and Y reversed.

Example 4.17 Problem



Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$, as given in Example 4.13 and repeated in the accompanying graph. Find the conditional PMF of Y given $X = x$ for each $x \in S_X$.

Example 4.17 Solution

To apply Theorem 4.22, we first find the marginal PMF $P_X(x)$. By Theorem 4.3, $P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y)$. For a given $X = x$, we sum the nonzero probablities along the vertical line $X = x$. That is,

$$P_X(x) = \begin{cases} 1/4 & x = 1, \\ 1/8 + 1/8 & x = 2, \\ 1/12 + 1/12 + 1/12 & x = 3, \\ 1/16 + 1/16 + 1/16 + 1/16 & x = 4, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1/4 & x = 1, \\ 1/4 & x = 2, \\ 1/4 & x = 3, \\ 1/4 & x = 4, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 4.22 implies that for $x \in \{1, 2, 3, 4\}$,

$$P_{Y|X}(y|x) = \frac{P_{X,Y}(x, y)}{P_X(x)} = 4P_{X,Y}(x, y).$$

[Continued]

Example 4.17 Solution (continued)

For each $x \in \{1, 2, 3, 4\}$, $P_{Y|X}(y|x)$ is a different PMF.

$$P_{Y|X}(y|1) = \begin{cases} 1 & y = 1, \\ 0 & \text{otherwise.} \end{cases} \qquad P_{Y|X}(y|2) = \begin{cases} 1/2 & y \in \{1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

$$P_{Y|X}(y|3) = \begin{cases} 1/3 & y \in \{1, 2, 3\}, \\ 0 & \text{otherwise.} \end{cases} \qquad P_{Y|X}(y|4) = \begin{cases} 1/4 & y \in \{1, 2, 3, 4\}, \\ 0 & \text{otherwise.} \end{cases}$$

Given $X = x$, the conditional PMF of Y is the discrete uniform $(1, x)$ random variable.

Theorem 4.23 Conditional Expected Value of a Function

X and Y are discrete random variables. For any $y \in S_Y$, the conditional expected value of $g(X, Y)$ given $Y = y$ is

$$E[g(X, Y)|Y = y] = \sum_{x \in S_X} g(x, y) P_{X|Y}(x|y).$$

- The conditional expected value of X given $Y = y$ is a special case of Theorem 4.23:

$$E[X \mid Y = y] = \sum_{x \in S_X} x P_{X|Y}(x \mid y)$$

Example 4.18 Problem

In Example 4.17, we derived the following conditional PMFs: $P_{Y|X}(y|1)$, $P_{Y|X}(y|2)$, $P_{Y|X}(y|3)$, and $P_{Y|X}(y|4)$. Find $E[Y|X = x]$ for $x = 1, 2, 3, 4$.

Example 4.18 Solution

Applying Theorem 4.23 with $g(x, y) = x$, we calculate

y

$$E[Y|X = 1] = 1, \qquad E[Y|X = 2] = 1.5,$$

$$E[Y|X = 3] = 2, \qquad E[Y|X = 4] = 2.5.$$

Definition 4.13 Conditional PDF

For y such that $f_Y(y) > 0$, the conditional PDF of X given $\{Y = y\}$ is

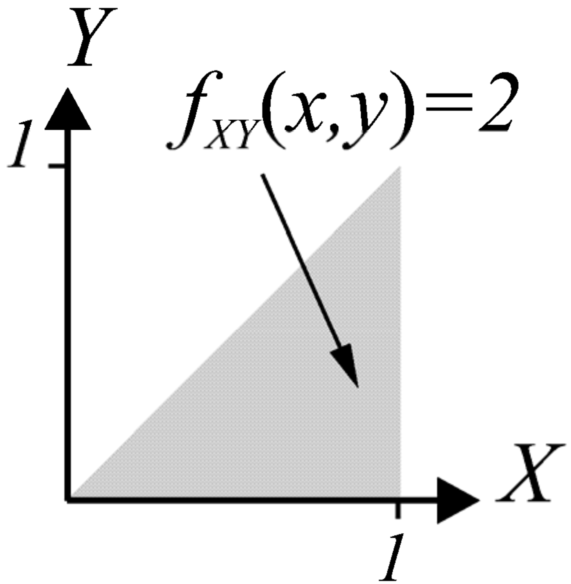
$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

- Definition 4.13 implies

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

Example 4.19 Problem

Returning to Example 4.5, random variables X and Y have joint PDF



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

For $0 \le x \le 1$, find the conditional PDF $f_{Y|X}(y|x)$. For $0 \le y \le 1$, find the conditional PDF $f_{X|Y}(x|y)$.

Example 4.19 Solution

For $0 \leq x \leq 1$, Theorem 4.8 implies

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy = \int_0^x 2\, dy = 2x.$$

The conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \begin{cases} 1/x & 0 \leq y \leq x, \\ 0 & \text{otherwise.} \end{cases}$$

Given $X = x$, we see that Y is the uniform $(0, x)$ random variable. For $0 \leq y \leq 1$, Theorem 4.8 implies

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx = \int_y^1 2\, dx = 2(1 - y).$$

Furthermore, Equation (4.102) implies

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \begin{cases} 1/(1 - y) & y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Conditioned on $Y = y$, we see that X is the uniform $(y, 1)$ random variable.

Theorem 4.24

$$f_{X,Y}(x, y) = f_{Y|X}(y|x)\, f_X(x) = f_{X|Y}(x|y)\, f_Y(y)\,.$$

- For each y with $f_Y(y) > 0$, the conditional PDF $f_{X|Y}(x|y)$ gives us a new probability model of X. We can use this model in any way that we use $f_X(x)$, the model we have in the absence of knowledge of Y.
- Most important, we can find expected values with respect to $f_{X|Y}(x|y)$ just as we do in **Chapter 3** with respect to $f_X(x)$.

Definition 4.14 Conditional Expected Value of a Function

For continuous random variables X and Y and any y such that $f_Y(y) > 0$, the conditional expected value *of $g(X, Y)$ given $Y = y$ is*

$$E[g(X,Y)|Y=y] = \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)\, dx.$$

- The conditional expected value of X given $Y=y$ is a special case of **Definition 4.14**.

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

Comments

- $E[X]$: a number derived from the probability model of X
- Case 1: $E[X|B]$: the same, a **number**.
- Case 2: $E[X|Y=y]$
  - Conditional expected value is a different number for each possible observation $y \in S_Y$. Therefore, $E[X|Y=y]$, is a deterministic function of the observation y.
  - When we perform an experiment and observe $Y=y$, $E[X|Y=y]$ is a function of the random variable Y.
  - We use the notation $E[X|Y]$ to denote this function of the random variable Y.
  - $E[X|Y]$ is a **random variable!**

Definition 4.15 Conditional Expected Value

The conditional expected value $E[X|Y]$ *is a function of random variable* Y *such that if* $Y = y$ *then* $E[X|Y] = E[X|Y = y]$.

Example 4.20 Problem

For random variables X and Y in Example 4.5, we found in Example 4.19 that the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(1-y) & y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find the conditional expected values $E[X|Y=y]$ and $E[X|Y]$.

Example 4.20 Solution

Given the conditional PDF $f_{X|Y}(x|y)$, we perform the integration

$$\begin{aligned} E[X|Y=y] &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx \\ &= \int_y^1 \frac{1}{1-y} x\, dx = \left. \frac{x^2}{2(1-y)} \right|_{x=y}^{x=1} = \frac{1+y}{2}. \end{aligned}$$

Since $E[X|Y=y] = (1+y)/2$, $E[X|Y] = (1+Y)/2$.

Theorem 4.25 Iterated Expectation

$$E\left[E\left[X|Y\right]\right] = E\left[X\right].$$

- An interesting property of the random variable $E[X|Y]$ is its expected value $E[E[X|Y]]$. We find $E[E[X|Y]]$ in two steps: first we calculate $g(y) = E[X|Y = y]$ and then we apply Theorem 3.4 to evaluate $E[g(y)]$. This two-step process is known as **iterated expectation**.

Proof: Theorem 4.25

We consider continuous random variables X and Y and apply Theorem 3.4:

$$E\left[E\left[X|Y\right]\right] = \int_{-\infty}^{\infty} E\left[X|Y=y\right] f_Y\left(y\right)\,dy.$$

To obtain this formula from Theorem 3.4, we have used $E[X|Y=y]$ in place of $g(x)$ and $f_Y(y)$ in place of $f_X(x)$. Next, we substitute the right side of Equation (4.108) for $E[X|Y=y]$:

$$E\left[E\left[X|Y\right]\right] = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x f_{X|Y}\left(x|y\right)\,dx\right) f_Y\left(y\right)\,dy.$$

Rearranging terms in the double integral and reversing the order of integration, we obtain:

$$E\left[E\left[X|Y\right]\right] = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X|Y}\left(x|y\right) f_Y\left(y\right)\,dy\,dx.$$

Next, we apply Theorem 4.24 and Theorem 4.8 to infer that the inner integral is simply $f_X(x)$. Therefore,

$$E\left[E\left[X|Y\right]\right] = \int_{-\infty}^{\infty} x f_X\left(x\right)\,dx.$$

The proof is complete because the right side of this formula is the definition of $E[X]$. A similar derivation (using sums instead of integrals) proves the theorem for discrete random variables.

Theorem 4.26

$$E\left[E\left[g(X)|Y\right]\right] = E\left[g(X)\right].$$

- Theorem 4.26 decomposes the calculation of $E[g(X)]$ into two steps: the calculation of $E[g(X)|Y=y]$, followed by he averaging of $E[g(X)|Y=y]$ over the distribution of Y. This is another example of iterated expectation.
- In Section 4.11, we will see that the iterated expectation can both facilitate understanding as well as simplify calculations.

Example 4.21 Problem

At noon on a weekday, we begin recording new call attempts at a telephone switch. Let X denote the arrival time of the first call, as measured by the number of seconds after noon. Let Y denote the arrival time of the second call. In the most common model used in the telephone industry, X and Y are continuous random variables with joint PDF

$$f_{X,Y}(x, y) = \begin{cases} \lambda^2 e^{-\lambda y} & 0 \le x < y, \\ 0 & \text{otherwise.} \end{cases}$$

where $\lambda > 0$ calls/second is the average arrival rate of telephone calls. Find the marginal PDFs $f_X(x)$ and $f_Y(y)$ and the conditional PDFs $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$.

Example 4.21 Solution

For $x < 0$, $f_X(x) = 0$. For $x \geq 0$, Theorem 4.8 gives $f_X(x)$:

$$f_X(x) = \int_x^\infty \lambda^2 e^{-\lambda y}\, dy = \lambda e^{-\lambda x}.$$

Referring to Appendix A.2, we see that X is an exponential random variable with expected value $1/\lambda$. Given $X = x$, the conditional PDF of Y is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \begin{cases} \lambda e^{-\lambda(y-x)} & y > x, \\ 0 & \text{otherwise.} \end{cases}$$

To interpret this result, let $U = Y - X$ denote the interarrival time, the time between the arrival of the first and second calls. Problem 4.10.15 asks the reader to show that given $X = x$, U has the same PDF as X. That is, U is an exponential (λ) random variable.

[Continued]

Example 4.21 Solution (continued)

Now we can find the marginal PDF of Y. For $y < 0$, $f_Y(y) = 0$. Theorem 4.8 implies

$$f_Y(y) = \begin{cases} \int_0^y \lambda^2 e^{-\lambda y}\, dx = \lambda^2 y e^{-\lambda y} & y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Y is the Erlang $(2, \lambda)$ random variable (Appendix A). Given $Y = y$, the conditional PDF of X is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \begin{cases} 1/y & 0 \leq x < y, \\ 0 & \text{otherwise.} \end{cases}$$

Under the condition that the second call arrives at time y, the time of arrival of the first call is the uniform $(0, y)$ random variable.

- In Example 4.21, Joint PDF → two conditional PDFs.
- Often in practice situation, condition PDF + marginal PDF → joint PDF + the other conditional PDF.

Example 4.22 Problem

Let R be the uniform $(0, 1)$ random variable. Given $R = r$, X is the uniform $(0, r)$ random variable. Find the conditional PDF of R given X.

Example 4.22 Solution

The problem definition states that

$$f_R(r) = \begin{cases} 1 & 0 \le r < 1, \\ 0 & \text{otherwise,} \end{cases} \qquad f_{X|R}(x|r) = \begin{cases} 1/r & 0 \le x < r < 1, \\ 0 & \text{otherwise.} \end{cases}$$

It follows from Theorem 4.24 that the joint PDF of R and X is

$$f_{R,X}(r,x) = f_{X|R}(x|r)\, f_R(r) = \begin{cases} 1/r & 0 \le x < r < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Now we can find the marginal PDF of X from Theorem 4.8. For $0 < x < 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{R,X}(r,x)\, dr = \int_x^1 \frac{dr}{r} = -\ln x.$$

By the definition of the conditional PDF,

$$f_{R|X}(r|x) = \frac{f_{R,X}(r,x)}{f_X(x)} = \begin{cases} \frac{1}{-r \ln x} & x \le r \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 4.9(A)

The probability model for random variable A is

$$P_A(a) = \begin{cases} 0.4 & a = 0, \\ 0.6 & a = 2, \\ 0 & \text{otherwise.} \end{cases}$$

The conditional probability model for random variable B given A is

$$P_{B|A}(b|0) = \begin{cases} 0.8 & b = 0, \\ 0.2 & b = 1, \\ 0 & \text{otherwise,} \end{cases} \qquad P_{B|A}(b|2) = \begin{cases} 0.5 & b = 0, \\ 0.5 & b = 1, \\ 0 & \text{otherwise.} \end{cases}$$

(1) What is the probability model for A and B? Write the joint PMF $P_{A,B}(a,b)$ as a table.

(2) If $A = 2$, what is the conditional expected value $E[B|A = 2]$?

(3) If $B = 0$, what is the conditional PMF $P_{A|B}(a|0)$?

(4) If $B = 0$, what is the conditional variance $\text{Var}[A|B = 0]$ of A?

Quiz 4.9(A) Solution

(1) The joint PMF of A and B can be found from the marginal and conditional PMFs via $P_{A,B}(a,b) = P_{B|A}(b|a)P_A(a)$. Incorporating the information from the given conditional PMFs can be confusing, however. Consequently, we can note that A has range $S_A = \{0,2\}$ and B has range $S_B = \{0,1\}$. A table of the joint PMF will include all four possible combinations of A and B. The general form of the table is

$P_{A,B}(a,b)$	$b=0$	$b=1$
$a=0$	$P_{B\|A}(0\|0)P_A(0)$	$P_{B\|A}(1\|0)P_A(0)$
$a=2$	$P_{B\|A}(0\|2)P_A(2)$	$P_{B\|A}(1\|2)P_A(2)$

Substituting values from $P_{B|A}(b|a)$ and $P_A(a)$, we have

$P_{A,B}(a,b)$	$b=0$	$b=1$
$a=0$	$(0.8)(0.4)$	$(0.2)(0.4)$
$a=2$	$(0.5)(0.6)$	$(0.5)(0.6)$

or

$P_{A,B}(a,b)$	$b=0$	$b=1$
$a=0$	0.32	0.08
$a=2$	0.3	0.3

(2) Given the conditional PMF $P_{B|A}(b|2)$, it is easy to calculate the conditional expectation

$$E[B|A=2] = \sum_{b=0}^{1} b P_{B|A}(b|2) = (0)(0.5) + (1)(0.5) = 0.5$$

[Continued]

Quiz 4.9(A) Solution (continued)

(3) From the joint PMF $P_{A,B}(a,b)$, we can calculate the the conditional PMF

$$P_{A|B}(a|0) = \frac{P_{A,B}(a,0)}{P_B(0)} = \begin{cases} 0.32/0.62 & a=0 \\ 0.3/0.62 & a=2 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 16/31 & a=0 \\ 15/31 & a=2 \\ 0 & \text{otherwise} \end{cases}$$

(4) We can calculate the conditional variance $\mathrm{Var}[A|B=0]$ using the conditional PMF $P_{A|B}(a|0)$. First we calculate the conditional expected value

$$E[A|B=0] = \sum_a a P_{A|B}(a|0) = 0(16/31) + 2(15/31) = 30/31$$

The conditional second moment is

$$E\left[A^2|B=0\right] = \sum_a a^2 P_{A|B}(a|0) = 0^2(16/31) + 2^2(15/31) = 60/31$$

The conditional variance is then

$$\mathrm{Var}[A|B=0] = E\left[A^2|B=0\right] - (E[A|B=0])^2 = \frac{960}{961}$$

Quiz 4.9(B)

The PDF of random variable X and theconditional PDF of random variable Y given X are

$$f_X(x) = \begin{cases} 3x^2 & 0 \le x \le 1, \\ 0 & \text{otherwise}, \end{cases} \qquad f_{Y|X}(y|x) = \begin{cases} 2y/x^2 & 0 \le y \le x, 0 < x \le 1, \\ 0 & \text{otherwise}. \end{cases}$$

(1) What is the probability model for X and Y? Find $f_{X,Y}(x, y)$.

(2) If $X = 1/2$, find the conditional PDF $f_{Y|X}(y|1/2)$.

(3) If $Y = 1/2$, what is the conditional PDF $f_{X|Y}(x|1/2)$?

(4) If $Y = 1/2$, what is the conditional variance $\mathrm{Var}[X|Y = 1/2]$?

Quiz 4.9(B) Solution

(1) The joint PDF of X and Y is

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)\, f_X(x) = \begin{cases} 6y & 0 \le y \le x, 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(2) From the given conditional PDF $f_{Y|X}(y|x)$,

$$f_{Y|X}(y|1/2) = \begin{cases} 8y & 0 \le y \le 1/2 \\ 0 & \text{otherwise} \end{cases}$$

(3) The conditional PDF of Y given $X = 1/2$ is $f_{X|Y}(x|1/2) = f_{X,Y}(x, 1/2)/f_Y(1/2)$. To find $f_Y(1/2)$, we integrate the joint PDF.

$$f_Y(1/2) = \int_{-\infty}^{\infty} f_{X,1/2}()\, dx = \int_{1/2}^{1} 6(1/2)\, dx = 3/2$$

Thus, for $1/2 \le x \le 1$,

$$f_{X|Y}(x|1/2) = \frac{f_{X,Y}(x, 1/2)}{f_Y(1/2)} = \frac{6(1/2)}{3/2} = 2$$

(4) From the pervious part, we see that given $Y = 1/2$, the conditional PDF of X is uniform $(1/2, 1)$. Thus, by the definition of the uniform (a, b) PDF,

$$\mathrm{Var}[X|Y = 1/2] = \frac{(1 - 1/2)^2}{12} = \frac{1}{48}$$

4.10 Independent Random Variables

- Definition 1.7 stated that events A and B are independent if and only if the probability of the intersection is the product of individual probabilities, $P[AB] = P[A]P[B]$.
- Applying this idea of independence to random variables, we say that X and Y are independent random variables if and only if the event $\{X = x\}$ and $\{Y = y\}$ are independent for all $x \in S_X$, and all $y \in S_y$.

Definition 4.16 Independent Random Variables

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

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

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

- Theorem 4.22 implies that if X and Y are independent discrete random variables, then

$$P_{X|Y}(x|y) = P_X(x),\ P_{Y|X}(y|x) = P_Y(y)$$

- Theorem 4.24 implies that if X and Y are independent continuous random variables, then

$$f_{X|Y}(x|y) = f_X(x),\ f_{Y|X}(y|x) = f_Y(y)$$

Example 4.23 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 4.23 Solution

The marginal PDFs of X and Y are

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

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 4.24 Problem

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

Are U and V independent?

Example 4.24 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

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

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

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

Comments

- The region of nonzero probability plays a crucial role in determining whether random variables are independent.
- To infer that X and Y are independent, it is necessary to verify the functional equalities in **Definition 4.16** for all $x \in S_X$ and $y \in S_Y$.
- The interpretation of independent random variables is a generalization of the interpretation of independent events.
- When X and Y are independent random variables,
  - The conditional PMF/PDF of X given $Y = y$ is the same for all $y \in S_Y$. And vice versa.
  - The conditional PMF/PDF is identical to corresponding marginal PMF/PDF.
  - Observing $Y = y$ des not alter out probability model for X. $Y = y$ provides no information about X. And vice versa.

Theorem 4.27

For independent random variables X and Y,

(a) $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$,

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

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

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

(e) $E[X|Y = y] = E[X]$ for all $y \in S_Y$,

(f) $E[Y|X = x] = E[Y]$ for all $x \in S_X$.

(c) states that independent random variables are uncorrelated. $\mathrm{Cov}[X, Y] = 0$ is a necessary property for independence, it is not sufficient.

Proof: Theorem 4.27

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 $P_{X,Y}(x, y) = P_X(x)P_Y(y)$,

$$E[g(X)h(Y)] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x)h(y)P_X(x)\,P_Y(y)$$

$$= \left(\sum_{x \in S_X} g(x)P_X(x)\right)\left(\sum_{y \in S_Y} h(y)P_Y(y)\right) = E[g(X)]\,E[h(Y)].$$

If $g(X) = X$, and $h(Y) = Y$, this equation implies $r_{X,Y} = E[XY] = E[X]E[Y]$. This equation and Theorem 4.16(a) imply $\text{Cov}[X, Y] = 0$. As a result, Theorem 4.16(b) implies $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$. Furthermore, $\rho_{X,Y} = \text{Cov}[X, Y]/(\sigma_X\sigma_Y) = 0$.

Since $P_{X|Y}(x|y) = P_X(x)$,

$$E[X|Y = y] = \sum_{x \in S_X} xP_{X|Y}(x|y) = \sum_{x \in S_X} xP_X(x) = E[X].$$

Since $P_{Y|X}(y|x) = P_Y(y)$,

$$E[Y|X = x] = \sum_{y \in S_Y} yP_{Y|X}(y|x) = \sum_{y \in S_Y} yP_Y(y) = E[Y].$$

Example 4.25 Problem

Random variables X and Y have a joint PMF given by the following matrix

$P_{X,Y}(x,y)$	$y=-1$	$y=0$	$y=1$
$x=-1$	0	0.25	0
$x=1$	0.25	0.25	0.25

Are X and Y independent? Are X and Y uncorrelated?

Example 4.25 Solution

For the marginal PMFs, we have $P_X(-1) = 0.25$ and $P_Y(-1) = 0.25$. Thus

$$P_X(-1)\, P_Y(-1) = 0.0625 \neq P_{X,Y}(-1, -1) = 0,$$

and we conclude that X and Y are not independent.

To find $\text{Cov}[X, Y]$, we calculate

$$E[X] = 0.5, \qquad E[Y] = 0, \qquad E[XY] = 0.$$

Therefore, Theorem 4.16(a) implies

$$\text{Cov}[X, Y] = E[XY] - E[X]\, E[Y] = \rho_{X,Y} = 0,$$

and by definition X and Y are uncorrelated.

Quiz 4.10(A)

Random variables X and Y in Example 4.1 and random variables Q and G in Quiz 4.2 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

(1) Are X and Y independent?

(2) Are Q and G independent?

Quiz 4.10(A) Solution

(1) For random variables X and Y from Example 4.1, we observe that $P_Y(1) = 0.09$ and $P_X(0) = 0.01$. However,

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

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 $P_Y(y) = 0$.

(2) For random variables Q and G from Quiz 4.2, 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 4.2.

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

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 4.10(B)

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

$$f_X(x) = \begin{cases} 1 - x/2 & 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

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

(2) Find the CDF of $Z = \max(X_1, X_2)$.

Quiz 4.10(B) Solution

(1) Since X_1 and X_2 are independent,

$$
\begin{aligned}
f_{X_1,X_2}(x_1,x_2) &= f_{X_1}(x_1)\, f_{X_2}(x_2) \\
&= \begin{cases} (1 - x_1/2)(1 - x_2/2) & 0 \le x_1 \le 2, 0 \le x_2 \le 2 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

(2) Let $F_X(x)$ denote the CDF of both X_1 and X_2. The CDF of $Z = \max(X_1, X_2)$ is found by observing that $Z \le z$ iff $X_1 \le z$ and $X_2 \le z$. That is,

$$
\begin{aligned}
P[Z \le z] &= P[X_1 \le z, X_2 \le z] \\
&= P[X_1 \le z]\, P[X_2 \le z] = [F_X(z)]^2
\end{aligned}
$$

[Continued]

Quiz 4.10(B) Solution (continued)

To complete the problem, we need to find the CDF of each X_i. From the PDF $f_X(x)$, the CDF is

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy = \begin{cases} 0 & x < 0 \\ x - x^2/4 & 0 \le x \le 2 \\ 1 & x > 2 \end{cases}$$

Thus for $0 \le z \le 2$,

$$F_Z(z) = (z - z^2/4)^2$$

The complete expression for the CDF of Z is

$$F_Z(z) = \begin{cases} 0 & z < 0 \\ (z - z^2/4)^2 & 0 \le z \le 2 \\ 1 & z > 1 \end{cases}$$

4.11 Bivariate Gaussian Random Variables

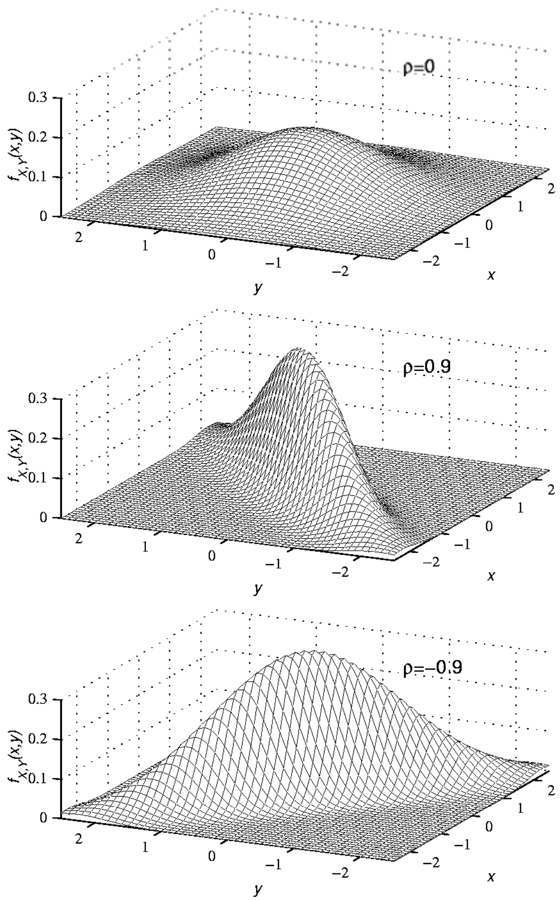
Definition 4.17 Bivariate Gaussian Random Variables

Random variables X *and* Y *have a* bivariate Gaussian PDF *with parameters* μ_1, σ_1, μ_2, σ_2, *and* ρ *if*

$$f_{X,Y}(x, y) = \frac{\exp\left[-\dfrac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

where μ_1 *and* μ_2 *can be any real numbers,* $\sigma_1 > 0$, $\sigma_2 > 0$, *and* $-1 < \rho < 1$.

Figure 4.5



The Joint Gaussian PDF $f_{X,Y}(x, y)$ for $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and

- Figure 4.5 illustrates the bivariate Gaussian PDF for $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and three values of ρ.
- $\rho = 0$: the joint PDF has the circular symmetry of a sombrero.
- $\rho = 0.9$: the joint PDF forms a ridge over the line $x = y$, and
- $\rho = -0.9$: the joint PDF forms a ridge over the line $x = -y$.
- The ridge becomes increasingly steep as $\rho \to \pm 1$.

- To examine mathematically the properties of the bivariate Gaussian PDF, we define

$$\tilde{\mu}_2(x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \qquad \tilde{\sigma}_2 = \sigma_2\sqrt{1-\rho^2}$$

  and manipulate the formula in Definition 4.17 to obtain the following expression for the joint Gaussian PDF:

$$f_{X,Y}(x,y) = \frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{-(x-\mu_1)^2}{2\sigma_1^2}}\frac{1}{\tilde{\sigma}_2\sqrt{2\pi}}e^{-\frac{-(y-\tilde{\mu}_2)^2}{2\tilde{\sigma}_2^2}} \qquad (4.146)$$

- Equation (4.146) expresses $f_{X,Y}(x, y)$ as the product of two Gaussian PDFs. This formula plays a key role in the proof of the following theorem. (reference theorem 4.29, 4.30)

Theorem 4.28

If X and Y are the bivariate Gaussian random variables in Definition 4.17, X is the Gaussian (μ_1, σ_1) random variable and Y is the Gaussian (μ_2, σ_2) random variable:

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} \qquad f_Y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(y-\mu_2)^2/2\sigma_2^2}.$$

Proof: Theorem 4.28

Integrating $f_{X,Y}(x, y)$ in Equation (4.146) over all y, we have

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy \\
&= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2}\, dy}_{1}
\end{aligned}
$$

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_1, σ_1) random variable. The same reasoning with the roles of X and Y reversed leads to the formula for $f_Y(y)$.

Theorem 4.29

If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2},$$

where, given $X = x$, the conditional expected value and variance of Y are

$$\tilde{\mu}_2(x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \qquad \tilde{\sigma}_2^2 = \sigma_2^2(1 - \rho^2).$$

- Given the marginal PDFs of X and Y, we use Definition 4.13 to find the conditional PDFs.

- Theorem 4.29 is the result of dividing $f_{X,Y}(x, y)$ in Equation (4.146) by $f_X(x)$ to obtain $f_{Y|X}(y|x)$.
- The cross section of Figure 4.6 illustrate the conditional PDF. The figure is a graph of $f_{X,Y}(x, y) = f_{Y|X}(y|x)\, f_X(x)$. Since X is a constant on each cross section, the cross section is scaled picture of $f_{Y|X}(y|x)$. As Theorem 4.29 indicates, the cross section has the Gaussian bell shape.
- Corresponding to Theorem 4.29, this conditional PDF of X given Y is also Gaussian. This conditional PDF is found by dividing $f_{X,Y}(x, y)$ by $f_Y(y)$ to obtain $f_{X|Y}(x|y)$.

Theorem 4.30

If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of X given Y is

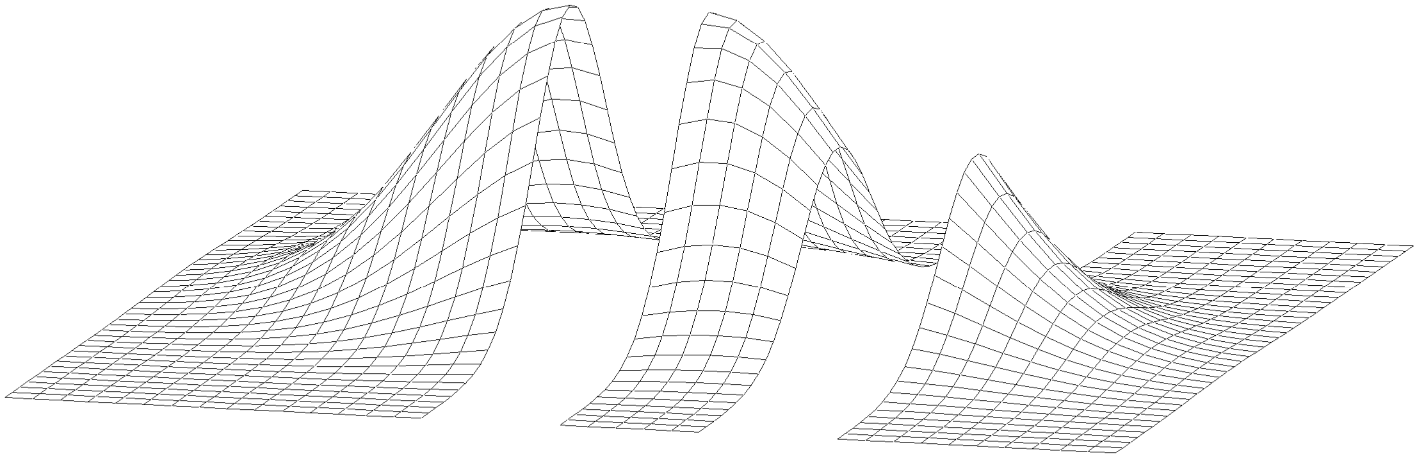
$$f_{X|Y}(x|y) = \frac{1}{\tilde{\sigma}_1\sqrt{2\pi}} e^{-(x-\tilde{\mu}_1(y))^2/2\tilde{\sigma}_1^2},$$

where, given $Y = y$, the conditional expected value and variance of X are

$$\tilde{\mu}_1(y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2) \qquad \tilde{\sigma}_1^2 = \sigma_1^2(1 - \rho^2).$$

Figure 4.6

$$f_{X,Y}(x, y) = f_{Y|X}(y \mid x) f_X(x)$$



Cross-sectional view of the joint Gaussian PDF with $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and $\rho = 0.9$. Theorem 4.29 confirms that the bell shape of the cross section occurs because the conditional PDF $f_{Y|X}(y|x)$ is Gaussian.

Theorem 4.31

Bivariate Gaussian random variables X and Y in Definition 4.17 have correlation coefficient

$$\rho_{X,Y} = \rho.$$

Proof: Theorem 4.31

Substituting μ_1, σ_1, μ_2, and σ_2 for μ_X, σ_X, μ_Y, and σ_Y in Definition 4.4 and Definition 4.8, we have

$$\rho_{X,Y} = \frac{E\left[(X-\mu_1)(Y-\mu_2)\right]}{\sigma_1\sigma_2}.$$

To evaluate this expected value, we use the substitution

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) f_X(x)$$

in the double integral of Theorem 4.12. The result can be expressed as

$$\begin{aligned}\rho_{X,Y} &= \frac{1}{\sigma_1\sigma_2}\int_{-\infty}^{\infty}(x-\mu_1)\left(\int_{-\infty}^{\infty}(y-\mu_2)\, f_{Y|X}(y|x)\, dy\right) f_X(x)\, dx \\ &= \frac{1}{\sigma_1\sigma_2}\int_{-\infty}^{\infty}(x-\mu_1)\, E\left[Y-\mu_2|X=x\right] f_X(x)\, dx\end{aligned}$$

Because $E[Y|X=x] = \tilde{\mu}_2(x)$ in Theorem 4.29, it follows that

$$E\left[Y-\mu_2|X=x\right] = \tilde{\mu}_2(x) - \mu_2 = \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$

Therefore,

$$\rho_{X,Y} = \frac{\rho}{\sigma_1^2}\int_{-\infty}^{\infty}(x-\mu_1)^2\, f_X(x)\, dx = \rho,$$

- From Theorem 4.31, we observe that if X and Y are uncorrelated, then $\rho = 0$ and,
- From Theorem 4.29 and 4.30, $f_{Y|X}(y|x) = f_Y(y)$ and $f_{X|Y}(x|y) = f_X(x)$.
- Thus we have the following theorem.

Theorem 4.32

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

– From theorem 4.31, we observe that if X and Y are uncorrelated, then ρ = 0 and, from Theorem 4.29 and 4.30, $f_{Y|X}(y|x) = f_Y(y)$ and $f_{X|Y}(x|y) = f_X(x)$.
– Theorem 4.31: $\rho_{X,Y} = \rho$.
– Theorem 4.17: $|\rho_{X,Y}| = |\rho| < 1$.
– The inequality for conditional variance:

$$\mathrm{Var}[Y \mid X = x] = \sigma_2^2\left(1-\rho^2\right) \le \sigma_2^2$$
$$\mathrm{Var}[X \mid Y = y] = \sigma_1^2\left(1-\rho^2\right) \le \sigma_1^2$$

– For $\rho \neq 0$, learning the value of one of the random variables leads to a model of the other random variable with reduced variance. This suggests that learning the value of Y reduces our uncertainty regarding X.

Quiz 4.11

Let X and Y be jointly Gaussian $(0, 1)$ random variables with correlation coefficient $1/2$.

(1) What is the joint PDF of X and Y?

(2) What is the conditional PDF of X given $Y = 2$?

Quiz 4.11 Solution

This problem just requires identifying the various terms in Definition 4.17 and Theorem 4.29. Specifically, from the problem statement, we know that $\rho = 1/2$,

$$\mu_1 = \mu_X = 0, \qquad \mu_2 = \mu_Y = 0,$$

and that

$$\sigma_1 = \sigma_X = 1, \qquad \sigma_2 = \sigma_Y = 1.$$

(1) Applying these facts to Definition 4.17, we have

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

(2) By Theorem 4.30, the conditional expected value and standard deviation of X given $Y = y$ are

$$E[X|Y = y] = y/2 \qquad \tilde{\sigma}_X = \sigma_1^2(1 - \rho^2) = \sqrt{3/4}.$$

When $Y = y = 2$, we see that $E[X|Y = 2] = 1$ and $\mathrm{Var}[X|Y = 2] = 3/4$. The conditional PDF of X given $Y = 2$ is simply the Gaussian PDF

$$f_{X|Y}(x|2) = \frac{1}{\sqrt{3\pi/2}} e^{-2(x-1)^2/3}.$$