Probability and Stochastic Processes

A Friendly Introduction for Electrical and Computer Engineers

SECOND EDITION

Roy D. Yates   David J. Goodman

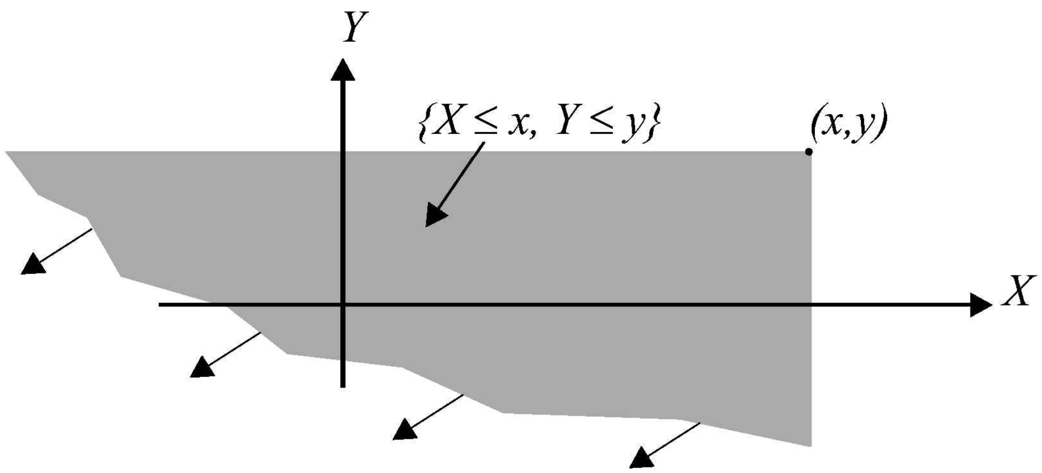
Definitions, Theorems, Proofs, Examples, Quizzes, Problems, Solutions

Chapter 4

Section 4.1

Joint Cumulative Distribution Function

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*

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

$$= P[A \cap B]$$

where $A = \{X \leq x\}$, $B = \{Y \leq y\}$

Theorem 4.1

For any pair of random variables, X, Y,

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

(b) $F_X(x) = F_{X,Y}(x, \infty)$, $= P[X \le x, Y \le \infty] = P[X \le x]$

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

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

$\hookrightarrow P[X \le -\infty, Y \le y] = P[\phi \cap B] = P[\phi] = 0$

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

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

$\hookrightarrow P[X \le \infty, Y \le \infty] = P[S \cap S] = P[S] = 1$

Section 4.2

Joint Probability Mass Function

Definition 4.2 *(PMF)*

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

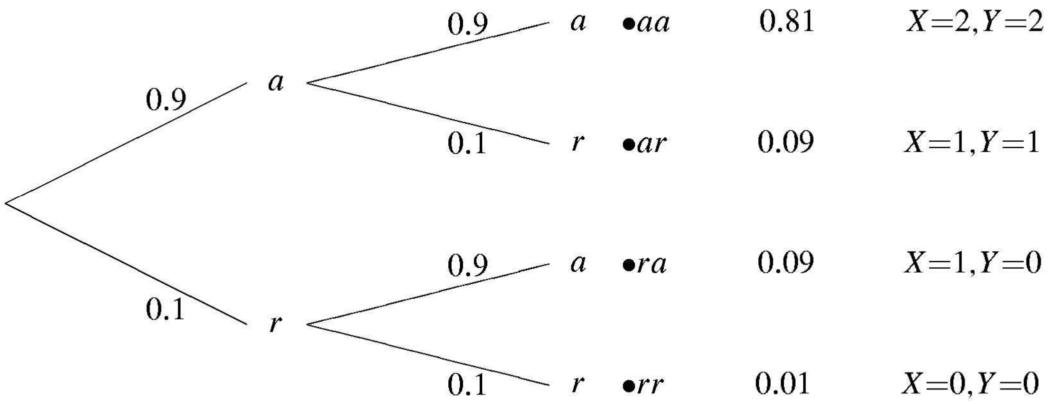
$$P_{X,Y}(x, y) = P[X = x, Y = y].$$

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.

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

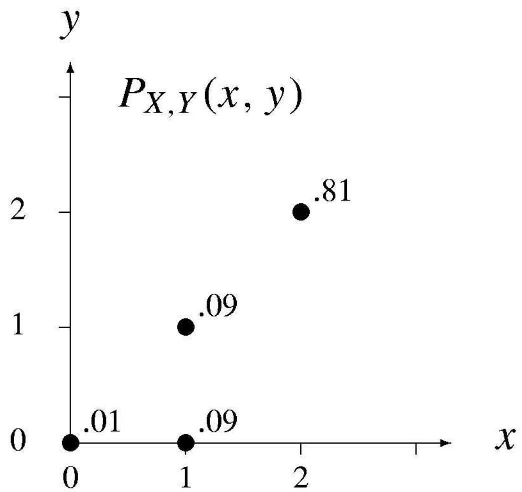
$$P[aa] = 0.81, \qquad P[ar] = P[ra] = 0.09, \qquad P[rr] = 0.01.$$

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:



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

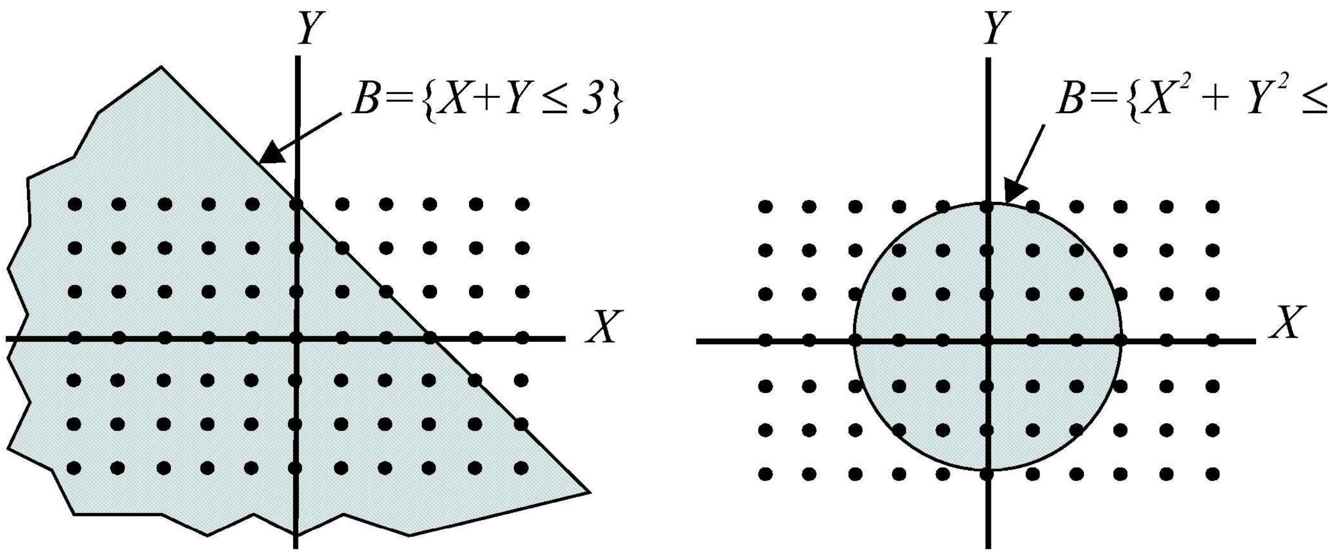
[Continued]

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

Figure 4.2



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

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,

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

Section 4.3

Marginal PMF

Theorem 4.3

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

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

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	

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.

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

Section 4.4

Joint Probability Density Function

Definition 4.3 Joint Probability Density Function (PDF)

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

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

Theorem 4.4

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

Theorem 4.5

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

$$P[x_1 < \mathrm{x} \le x_2, y_1 < \mathrm{y} \le y_2] = P[x_1 < \mathrm{x} \le x_2 \;, \mathrm{y} \le y_2] - P[x_1 < \mathrm{x} \le x_2 \;, \mathrm{y} \le y_1]$$

$$= F_{\mathrm{x,y}}(x_2, y_2) - F_{\mathrm{x,y}}(x_1, y_2) - F_{\mathrm{x,y}}(x_2, y_1) + F_{\mathrm{x,y}}(x_1, y_1)$$

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

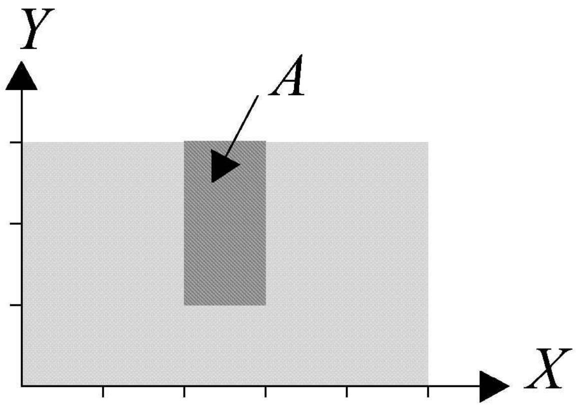
$$f_{X,Y}(x, y) = \begin{cases} c & 0 \leq x \leq 5, 0 \leq y \leq 3, \\ 0 & \text{otherwise.} \end{cases}$$

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

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:

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



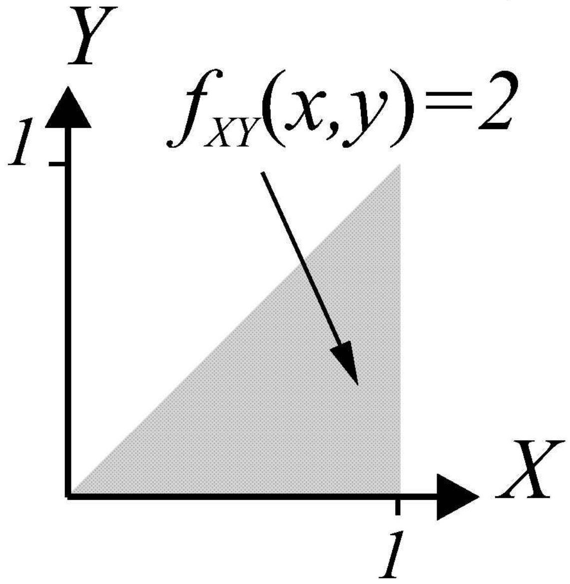
Therefore, $c = 1/15$. The small dark rectangle in the diagram is the event $A = \{2 \le X < 3, 1 \le Y < 3\}$. $P[A]$ is the integral of the PDF over this rectangle, which is

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

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



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

Example 4.5 Solution

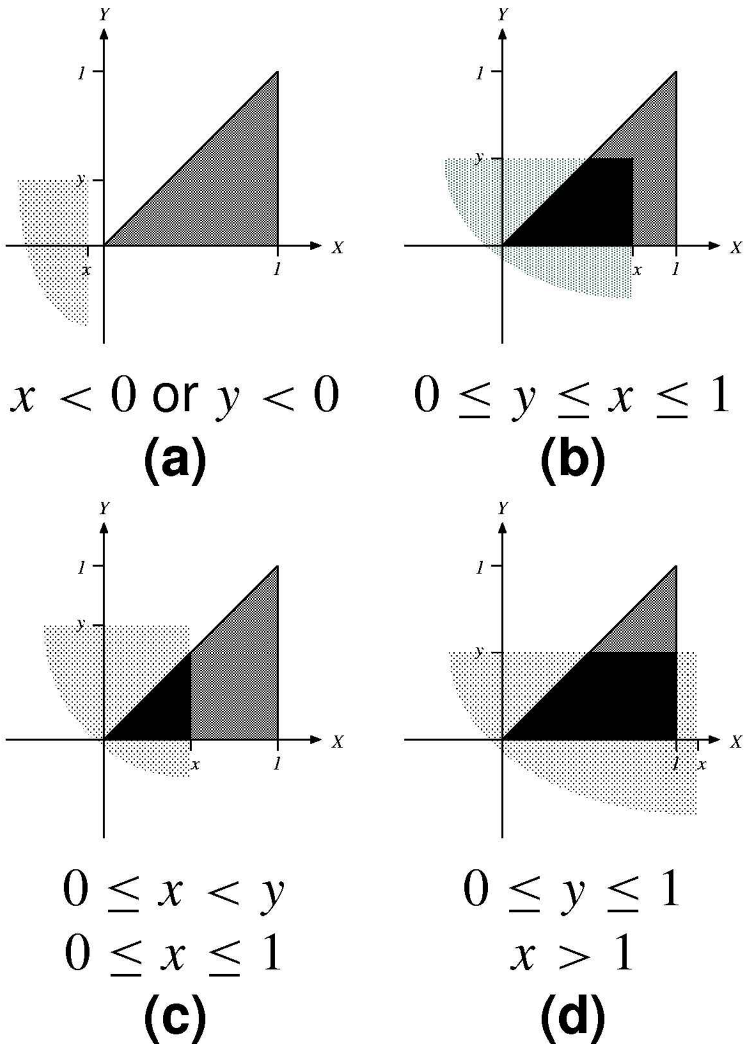
The joint CDF can be found using Definition 4.3 in which we integrate the joint PDF $f_{X,Y}(x, y)$ over the area shown in Figure 4.1. To perform the integration it is extremely useful to draw a diagram that clearly shows the area with nonzero probability, and then to use the diagram to derive the limits 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]

Figure 4.3



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

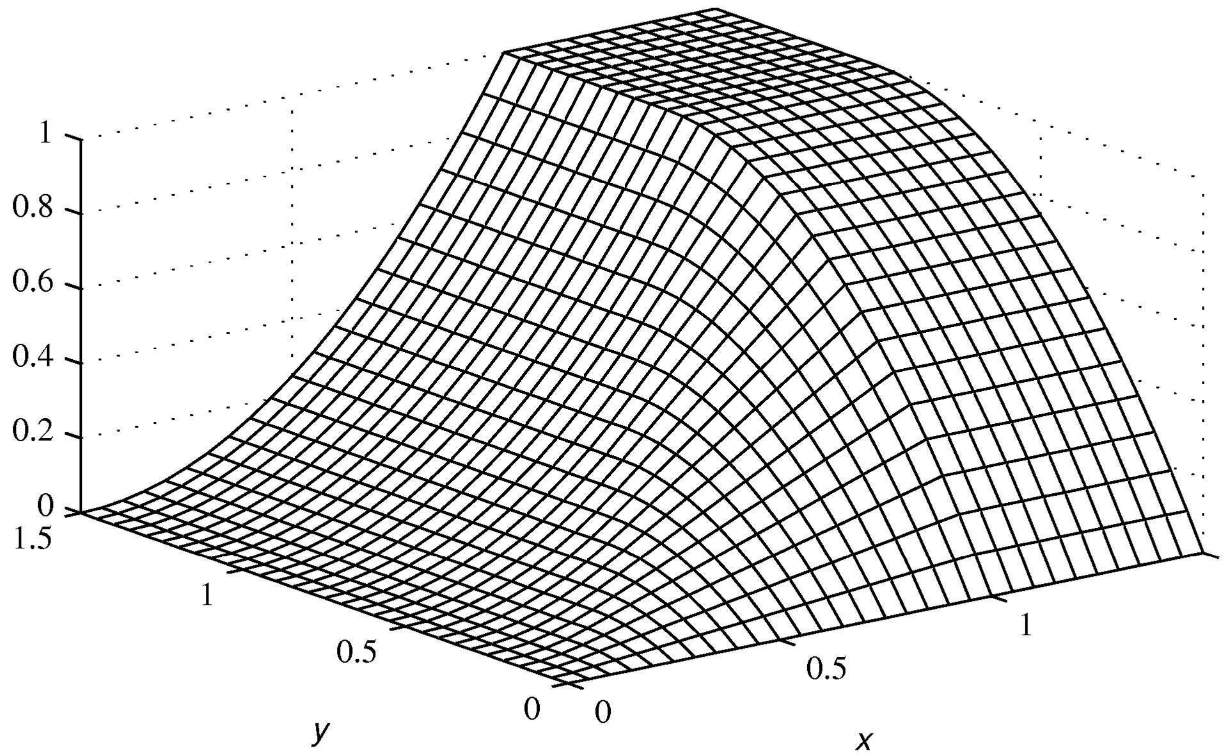
$$F_{X,Y}(x, y) = \int_0^y \int_v^1 2\,du\,dv = 2y - y^2 \qquad \text{(Figure 4.3d)}$$

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.

Example 4.6 Problem

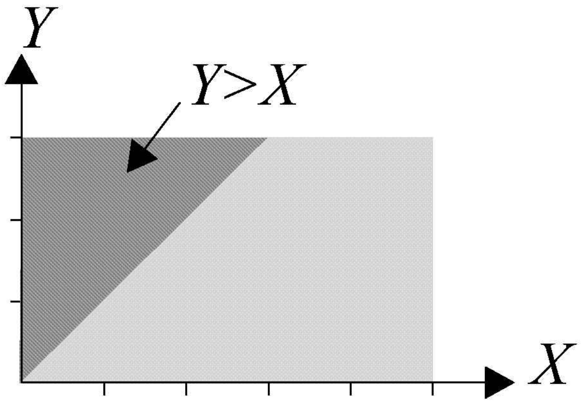
As in Example 4.4, random variables X and Y have joint PDF

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

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,



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

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

Section 4.5

Marginal PDF

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

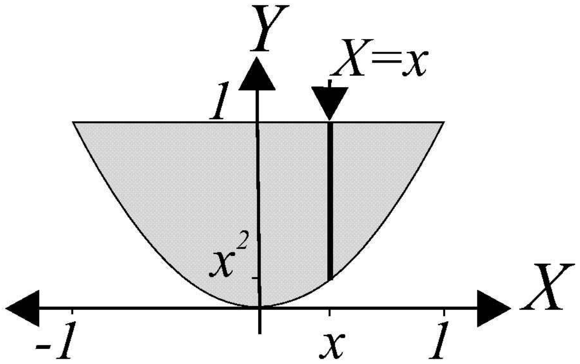
The joint PDF of X and Y is

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

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

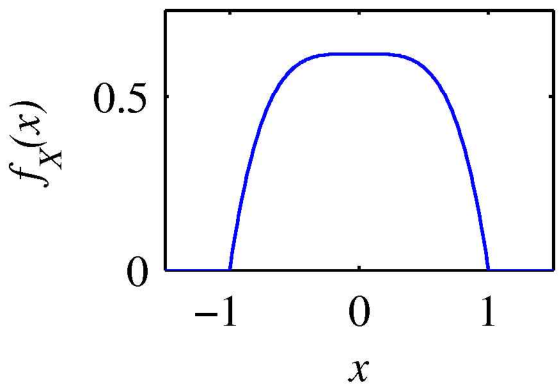
Example 4.7 Solution

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

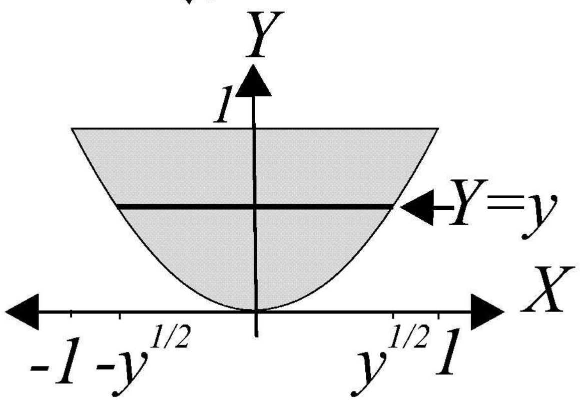


$$f_X(x) = \begin{cases} 5(1 - x^4)/8 & -1 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

[Continued]

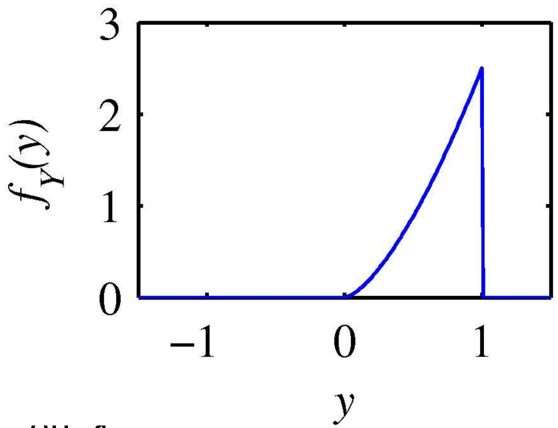
Example 4.7 Solution (continued)

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 $0 \le y \le 1$,



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

The complete marginal PDF of Y is



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

Section 4.6

Functions of Two Random Variables

Functions of Two Random Variables

Let the random variable Z be defined as a function of X and Y:

$$Z = g(X,Y).$$

The cdf of Z is found similar to the functions of single random variable. If there exists D_z such that:

$$\{Z \le z\} = \{g(X,Y) \le z\} = \{(x,y) \in D_z\},$$

then, the cdf of Z is found by,

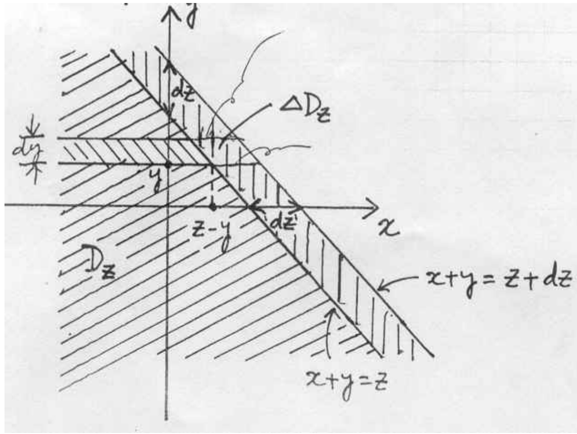
$$F_Z(z) = P\{Z \le z\} = P\{(X,Y) \in D_z\} = \iint_{D_z} f_{X,Y}(x,y)dxdy.$$

The pdf of Z is found by taking the derivative of $F_Z(z)$ with respect to z such that,

$$f_Z(z) = \frac{d}{dz} F_Z(z).$$

1. $Z = X + Y$

$$F_Z(z) = \iint_{D_z} f_{X,Y}(x,y)dxdy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{X,Y}(x,y)dxdy$$
$$f_Z(z) = \frac{\partial F_Z(z)}{\partial z} = \frac{\partial}{\partial z}\int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{X,Y}(x,y)dxdy$$
$$= \int_{-\infty}^{\infty} f_{X,Y}(z-y,y)\,dy$$



If x and y are independent, the following hold

$$f_{X,Y}(z-y,y) = f_X(z-y)f_Y(y).$$

In this case, Eq. (*) reads

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy.$$

This is a convolution integral.

Example ($Z = X + Y$)

X and Y are independent r.v. with

$$f_X(x) = \alpha e^{-\alpha x} u(x); \quad f_Y(y) = \beta e^{-\beta y} u(y).$$

Find $f_z(z)$ when $Z = X + Y$.

$$f_Z(z) = \int_0^z f_X(z-y) f_Y(y) dy$$

$$= \alpha\beta \int_0^z e^{-\alpha(z-y)} e^{-\beta y} dy$$

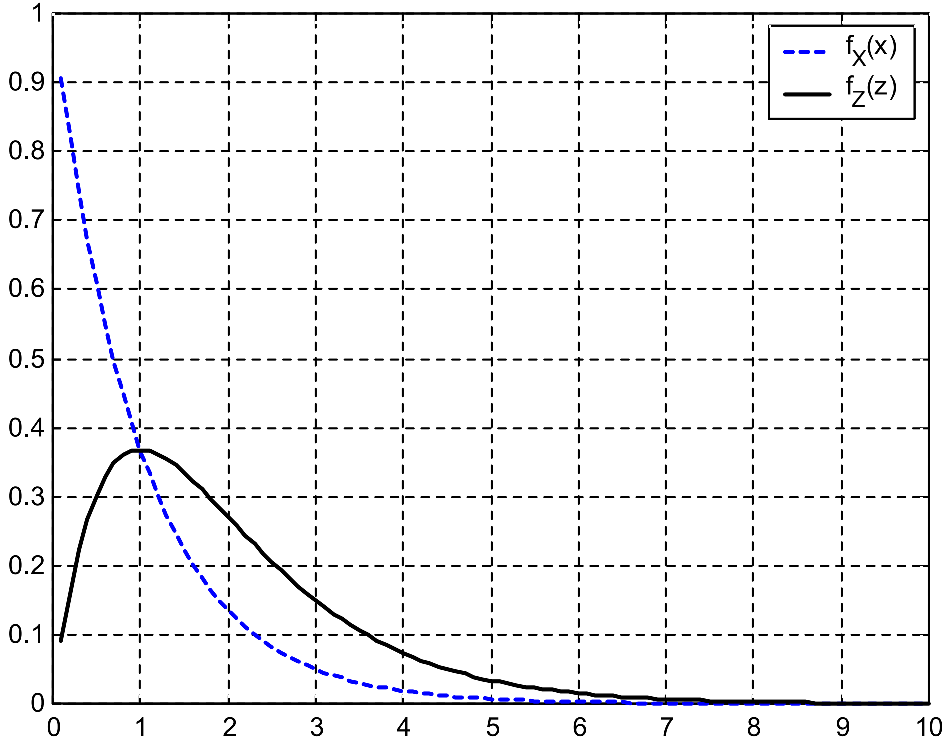
$$= \frac{\alpha\beta}{\beta - \alpha}(e^{-\alpha z} - e^{-\beta z}), \quad \beta \neq \alpha.$$

($f_X(x) = 0 \;\; for \;\; x < 0 \;\; => \;\; f_X(z-y) = 0 \;\; for \;\; y > z \;\; and \;\; f_Y(y) = 0 \;\; for \;\; y < 0$)

If $\alpha = \beta$,

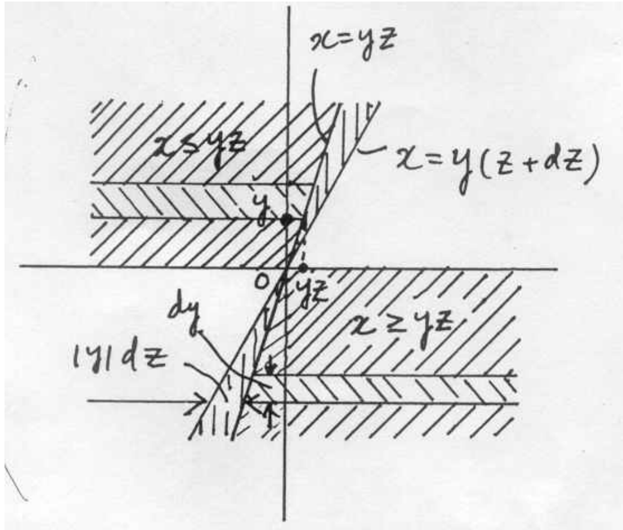
$$f_Z(z) = \alpha^2 z e^{-\alpha z}.$$

In this case, note that x and y are exponential random variables and z is an Erlang random variable with parameter m=2. Let's plot $f_X(x)$ and $f_Z(z)$ for $\alpha = 1$.



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
8
9
10
$f_X(x)$
$f_Z(z)$

2. $Z = X/Y$



$$F_Z(z) = \iint_{D_Z} f_{X,Y}(x,y)\,dxdy$$
$$= \int_0^\infty \int_{-\infty}^{yz} f_{X,Y}(x,y)\,dxdy + \int_{-\infty}^0 \int_{yz}^\infty f_{X,Y}(x,y)\,dxdy$$

$$\left(\frac{x}{y} \le z, \quad y > 0 \quad => x \le yz; \quad \frac{x}{y} \le z, \quad y < 0 \quad => \quad x \ge yz\right)$$

$$f_Z(z) = \frac{d}{dz}\int_0^\infty \int_{-\infty}^{yz} f_{X,Y}(x,y)\,dxdy + \int_{-\infty}^0 \int_{yz}^\infty f_{X,Y}(x,y)\,dxdy$$
$$= \int_0^\infty y\, f_{X,Y}(zy,y)\,dy + \int_{-\infty}^0 (-y)\, f_{X,Y}(zy,y)\,dy$$
$$= \int_{-\infty}^\infty |\,y\,|\, f_{X,Y}(zy,y)\,dy$$

Example 4.10 Problem

X and Y have the joint PDF

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

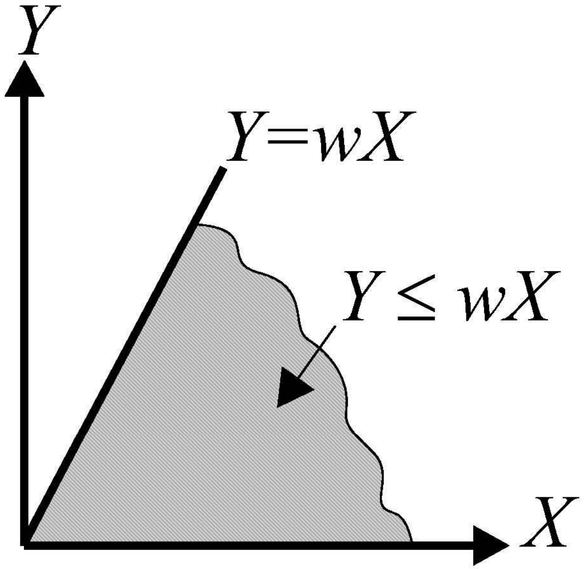
Find the PDF of $W = Y/X$.

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

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

Section 4.7

Expected Values

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 4.13

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

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

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

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

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

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,

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

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

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

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

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

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

Definition 4.4 Covariance

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

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

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

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

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

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

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,

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

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

Example 4.12 Solution

By Definition 4.5,

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

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

$$E[X] = (1)(0.18) + (2)(0.81) = 1.80,$$
$$E[Y] = (1)(0.09) + (2)(0.81) = 1.71.$$

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* $\text{Cov}[X, Y] = 0$.

Definition 4.8 Correlation Coefficient

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

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

Theorem 4.17

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

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,

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

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

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

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

Theorem 4.18

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

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

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 $\rho_{X,Y}$

Section 4.8

Conditioning by an Event

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*

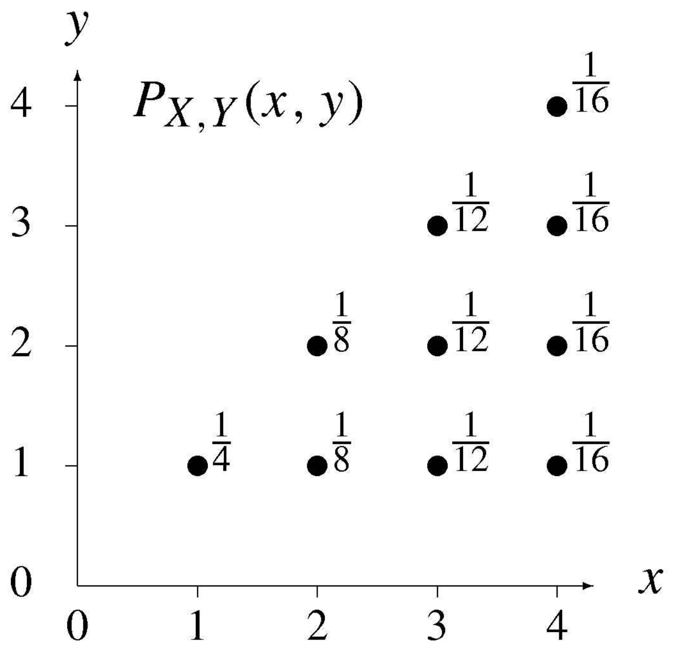
$$P_{X,Y|B}(x, y) = P[X = x, Y = y|B].$$

Theorem 4.19

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

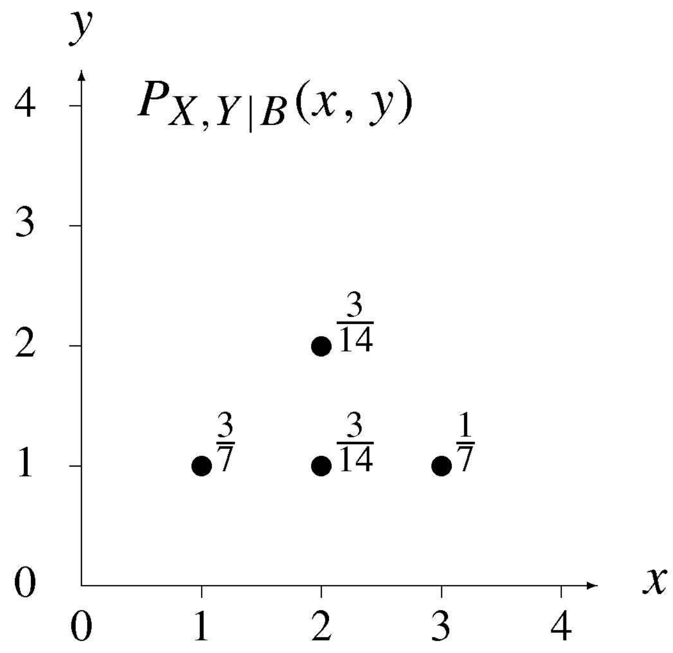
$$P_{X,Y|B}(x, y) = \begin{cases} \dfrac{P_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

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 $x + y \leq 4$. 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

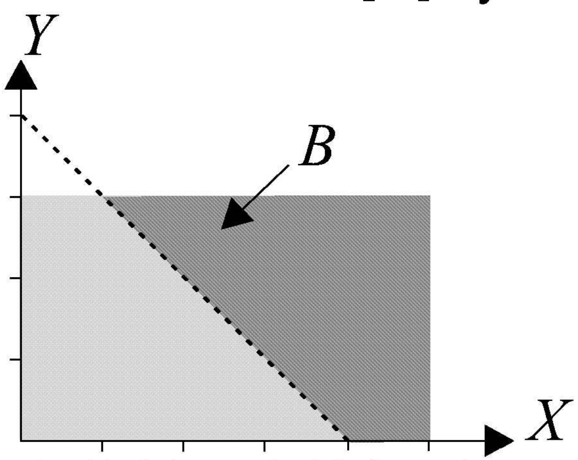
X and Y are random variables with joint PDF

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

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

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*

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

Theorem 4.21

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

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,

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

Similarly,

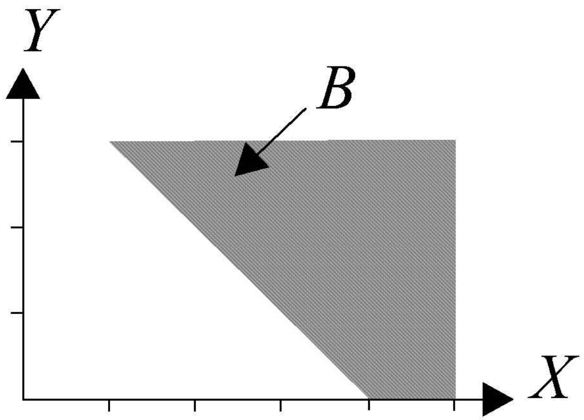
$$E\left[W^2|B\right] = \sum_{x,y} (x + y)^2 P_{X,Y|B}\left(x, y\right)$$
$$= 2^2\frac{3}{7} + 3^2\frac{3}{14} + 4^2\frac{1}{7} + 4^2\frac{3}{14} = \frac{131}{14}.$$

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

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

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

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 $\text{Var}[V|A]$?

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 $\text{Var}[V|A]$, 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

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

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 $\text{Var}[W|B]$?

Section 4.9

Conditioning by a Random Variable

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 $P_Y(y) > 0$,

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

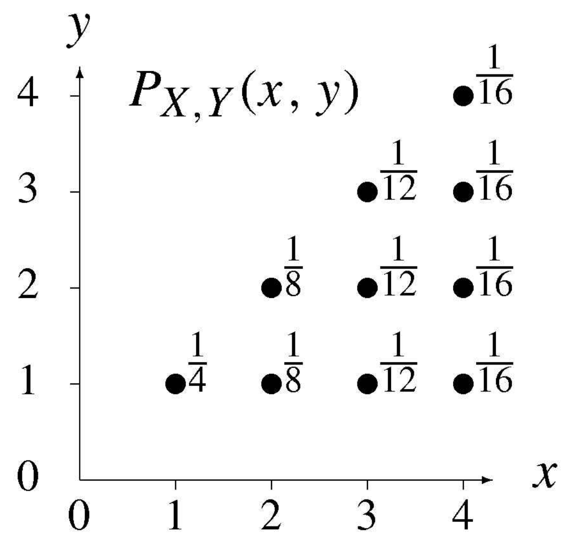
Proof: Theorem 4.22

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

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

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

$$E[Y|X = 1] = 1, \qquad E[Y|X = 2] = 1.5,$$
$$E[Y|X = 3] = 2, \qquad E[Y|X = 4] = 2.5.$$

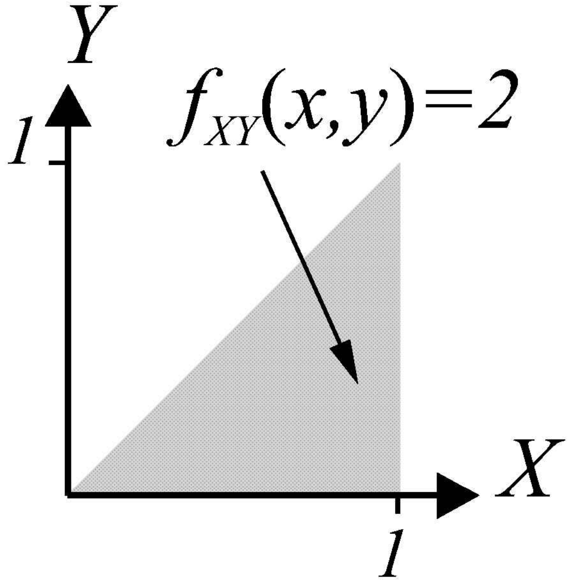
Definition 4.13 Conditional PDF

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

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

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 $0 \le x \le 1$, 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

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

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

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

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

Theorem 4.24

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

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*

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

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

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

$$= \int_{y}^{1} \frac{1}{1-y} x\, dx = \left. \frac{x^2}{2(1-y)} \right|_{x=y}^{x=1} = \frac{1+y}{2}.$$

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

Theorem 4.25 Iterated Expectation

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

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

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

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.

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(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 $\text{Var}[X|Y = 1/2]$?

Section 4.10

Independent Random Variables

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

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

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

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

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

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

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

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) $\text{Cov}[X, Y] = \rho_{X,Y} = 0$,

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

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

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

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

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

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

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

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

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

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

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.

Section 4.11

Bivariate Gaussian Random Variables

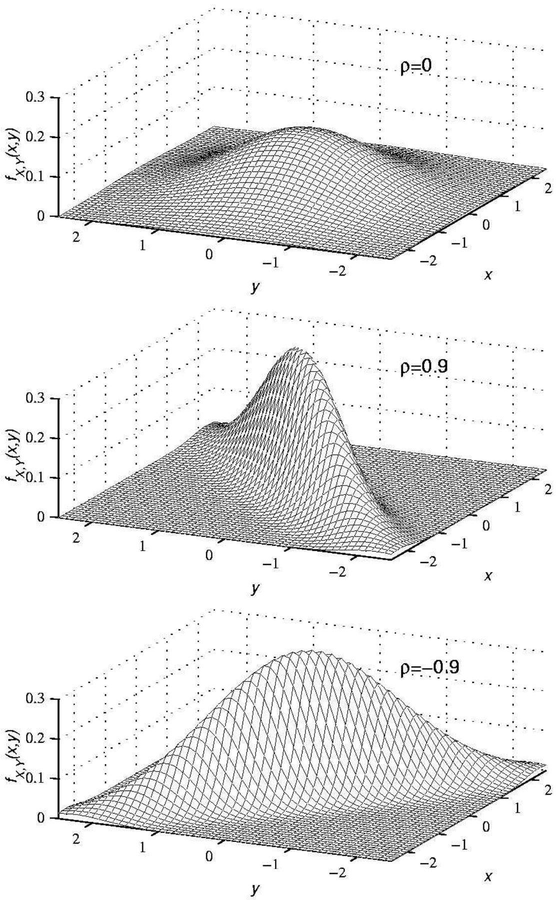
Definition 4.17 Bivariate Gaussian Random Variables

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

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

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 three values of ρ.

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

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

$$= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2}\, dy}_{1}$$

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

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

Theorem 4.30

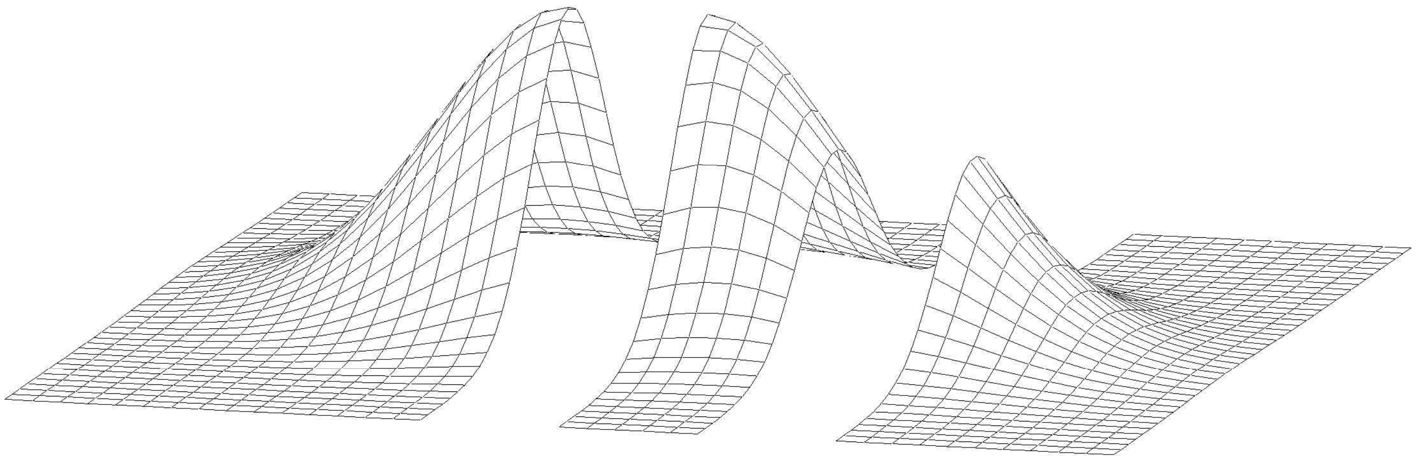
If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of X given Y is

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

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



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,

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

because the integral in the final expression is $Var\left[X\right] = \sigma_1^2$.

Theorem 4.32

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

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 $Y = 2$?