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 5

Section 5.1

Probability Models of N Random Variables

Definition 5.1 Multivariate Joint CDF

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

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

Definition 5.2 *Multivariate Joint PMF*

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

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

Definition 5.3 Multivariate Joint PDF

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

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

Theorem 5.1

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

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

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

Theorem 5.2

If $X_1, \ldots, X_n$ are continuous random variables with joint PDF $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$,

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

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

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

Theorem 5.3

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

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

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

Example 5.1 Problem

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

$$S = \{s_1, s_2, \cdots, s_r\}$$

$$n_1 + n_2 + \cdots + n_r = n$$

$$P_K(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\binom{n}{k} \rightarrow \frac{n!}{k!(n-k)!}$$

$$P_{N_1,\cdots,N_r}(n_1,\cdots,n_r)$$

$$= \binom{n}{n_1, n_2, \cdots, n_r} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

$$= \binom{n}{n_1}\binom{n-n_1}{n_2} \cdots \binom{n-n_1-\cdots-n_{r-1}}{n_r} p_1^{n_1} \cdots p_r^{n_r}$$

$$= \frac{n!}{(n-n_1)!\,n_1!} \cdot \frac{(n-n_1)!}{(n-n_1-n_2)!\,n_2!} \cdots$$

Example 5.1 Solution

The solution to this problem appears in Theorem 1.19 and is repeated here:

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

$$\binom{n}{n_1,\ldots,n_r} = \frac{n!}{n_1! \cdots n_r!}$$

$$n_1 + \cdots + n_r = n$$

Example 5.2 Problem

In response to requests for information, a company sends faxes that can be 1, 2, or 3 pages in length, depending on the information requested. The PMF of L, the length of one fax is

$$P_L(l) = \begin{cases} 1/3 & l = 1, \\ 1/2 & l = 2, \\ 1/6 & l = 3, \\ 0 & \text{otherwise.} \end{cases}$$

For a set of four independent information requests:

(a) What is the joint PMF of the random variables, X, Y,and Z, the number of 1-page, 2-page, and 3-page faxes, respectively?

(b) What is $P[A] = P[$total length of four faxes is 8 pages$]$?

(c) What is $P[B] = P[$at least half of the four faxes has more than 1 page$]$?

Example 5.2 Solution

Each fax sent is an independent trial with three possible outcomes: $L = 1$, $L = 2$, and $L = 3$. Hence, the number of faxes of each length out of four faxes is described by the multinomial PMF of Example 5.1:

$$P_{X,Y,Z}(x, y, z) = \binom{4}{x, y, z} \left(\frac{1}{3}\right)^x \left(\frac{1}{2}\right)^y \left(\frac{1}{6}\right)^z .$$

The PMF is displayed numerically in Table 5.1. The final column of the table indicates that there are three outcomes in event A and 12 outcomes in event B. Adding the probabilities in the two events, we have $P[A] = 107/432$ and $P[B] = 8/9$.

Table 5.1

x (1 page)	y (2 pages)	z (3 pages)	$P_{X,Y,Z}(x, y, z)$	total pages	events
0	0	4	1/1296	12	B
0	1	3	1/108	11	B
0	2	2	1/24	10	B
0	3	1	1/12	9	B
0	4	0	1/16	8	AB
1	0	3	1/162	10	B
1	1	2	1/18	9	B
1	2	1	1/6	8	AB
1	3	0	1/6	7	B
2	0	2	1/54	8	AB
2	1	1	1/9	7	B
2	2	0	1/6	6	B
3	0	1	2/81	6	
3	1	0	2/27	5	
4	0	0	1/81	4	

The PMF $P_{X,Y,Z}(x, y, z)$ and the events A and B for Example 5.2.

Example 5.3 Problem

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

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

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

Example 5.3 Solution

$$P[A] = P\left[\max_i X_i \leq 1/2\right] = P[X_1 \leq 1/2, \ldots, X_n \leq 1/2]$$
$$= \int_0^{1/2} \cdots \int_0^{1/2} 1\, dx_1 \cdots dx_n = \frac{1}{2^n}.$$

Here we have n independent uniform $(0, 1)$ random variables. As n grows, the probability that the maximum is less than $1/2$ rapidly goes to 0.

Section 5.2

Vector Notation

Definition 5.4 Random Vector

A random vector *is a column vector* $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}'$. *Each* X_i *is a random variable.*

Definition 5.5 Vector Sample Value

A sample value of a random vector *is a column vector* $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}'$*. The* i*th component,* x_i*, of the vector* $\mathbf{x}$ *is a sample value of a random variable,* X_i*.*

Random Vector Probability

Definition 5.6 Functions

(a) *The* CDF of a random vector $\mathbf{X}$ *is* $F_{\mathbf{X}}(\mathbf{x}) = F_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$.

(b) *The* PMF of a discrete random vector $\mathbf{X}$ *is* $P_{\mathbf{X}}(\mathbf{x}) = P_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$.

(c) *The* PDF of a continuous random vector $\mathbf{X}$ *is* $f_{\mathbf{X}}(\mathbf{x}) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$.

Probability Functions of a Pair of

Definition 5.7 Random Vectors

For random vectors $\mathbf{X}$ *with* n *components and* $\mathbf{Y}$ *with* m *components:*

(a) *The* joint CDF *of* $\mathbf{X}$ *and* $\mathbf{Y}$ *is*

$$F_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) = F_{X_1,\ldots,X_n,Y_1,\ldots,Y_m}(x_1,\ldots,x_n,y_1,\ldots,y_m);$$

(b) *The* joint PMF *of discrete random vectors* $\mathbf{X}$ *and* $\mathbf{Y}$ *is*

$$P_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) = P_{X_1,\ldots,X_n,Y_1,\ldots,Y_m}(x_1,\ldots,x_n,y_1,\ldots,y_m);$$

(c) *The* joint PDF *of continuous random vectors* $\mathbf{X}$ *and* $\mathbf{Y}$ *is*

$$f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) = f_{X_1,\ldots,X_n,Y_1,\ldots,Y_m}(x_1,\ldots,x_n,y_1,\ldots,y_m).$$

Example 5.4 Problem

Random vector $\mathbf{X}$ has PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 6e^{-\mathbf{a}'\mathbf{x}} & \mathbf{x} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{a} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}'$. What is the CDF of $\mathbf{X}$?

Example 5.4 Solution

Because $\mathbf{a}$ has three components, we infer that $\mathbf{X}$ is a 3-dimensional random vector. Expanding $\mathbf{a}'\mathbf{x}$, we write the PDF as a function of the vector components,

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 6e^{-x_1-2x_2-3x_3} & x_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Applying Definition 5.7, we integrate the PDF with respect to the three variables to obtain

$$F_{\mathbf{X}}(\mathbf{x}) = \begin{cases} (1-e^{-x_1})(1-e^{-2x_2})(1-e^{-3x_3}) & x_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Quiz 5.2

Discrete random vectors $\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}'$ and $\mathbf{Y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}'$ are related by $\mathbf{Y} = \mathbf{A}\mathbf{X}$. Find the joint PMF $P_{\mathbf{Y}}(\mathbf{y})$ if $\mathbf{X}$ has joint PMF

$$P_{\mathbf{X}}(\mathbf{x}) = \begin{cases} (1-p)p^{x_3} & x_1 < x_2 < x_3; \\ & x_1, x_2, x_3 \in \{1, 2, \ldots\}, \\ 0 & \text{otherwise,} \end{cases} \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

Section 5.3

Marginal Probability Functions

Theorem 5.4

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

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

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

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

Theorem 5.5

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

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

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

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

Example 5.5 Problem

As in Quiz 5.1, the random variables $Y_1, \ldots, Y_4$ have the joint PDF

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

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

Example 5.5 Solution

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

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

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

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

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

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

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

[Continued]

Example 5.5 Solution (continued)

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

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

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

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

The complete expression is

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

Section 5.4

Independence of Random Variables and Random Vectors

Definition 5.8 N Independent Random Variables

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

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

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

Example 5.6 Problem

As in Example 5.5, random variables $Y_1, \ldots, Y_4$ have the joint PDF

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

Are $Y_1, \ldots, Y_4$ independent random variables?

Example 5.6 Solution

In Equation (5.15) of Example 5.5, we found the marginal PDF $f_{Y_1,Y_4}(y_1, y_4)$. We can use this result to show that

$$f_{Y_1}(y_1) = \int_0^1 f_{Y_1,Y_4}(y_1, y_4)\, dy_4 = 2(1 - y_1), \qquad 0 \le y_1 \le 1,$$

$$f_{Y_4}(y_4) = \int_0^1 f_{Y_1,Y_4}(y_1, y_4)\, dy_1 = 2y_4, \qquad 0 \le y_4 \le 1.$$

The full expressions for the marginal PDFs are

$$f_{Y_1}(y_1) = \begin{cases} 2(1 - y_1) & 0 \le y_1 \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$f_{Y_4}(y_4) = \begin{cases} 2y_4 & 0 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the marginal PDF $f_{Y_2,Y_3}(y_2, y_3)$ found in Equation (5.17) of Example 5.5 implies that for $0 \le y_2 \le 1$,

$$f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f_{Y_2,Y_3}(y_2, y_3)\, dy_3 = \int_0^1 4y_2(1 - y_3)\, dy_3 = 2y_2$$

[Continued]

Example 5.6 Solution (continued)

It follows that the marginal PDF of Y_2 is

$$f_{Y_2}(y_2) = \begin{cases} 2y_2 & 0 \leq y_2 \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

From Equation (5.19) for the PDF $f_{Y_3}(y_3)$ derived in Example 5.5, we have

$$\begin{aligned} f_{Y_1}(y_1) f_{Y_2}(y_2) f_{Y_3}(y_3) f_{Y_4}(y_4) & \\ &= \begin{cases} 16(1-y_1)y_2(1-y_3)y_4 & 0 \leq y_1, y_2, y_3, y_4 \leq 1, \\ 0 & \text{otherwise,} \end{cases} \qquad (1) \\ &\neq f_{Y_1,\ldots,Y_4}(y_1,\ldots,y_4). \end{aligned}$$

Therefore $Y_1, \ldots, Y_4$ are not independent random variables.

Definition 5.9 Independent and Identically Distributed (iid)

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

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

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

Definition 5.10 Independent Random Vectors

Random vectors $\mathbf{X}$ *and* $\mathbf{Y}$ *are independent if*

Discrete: $P_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) = P_{\mathbf{X}}(\mathbf{x})\,P_{\mathbf{Y}}(\mathbf{y})$

Continuous: $f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) = f_{\mathbf{X}}(\mathbf{x})\,f_{\mathbf{Y}}(\mathbf{y})$.

Example 5.7 Problem

As in Example 5.5, random variables $Y_1, \ldots, Y_4$ have the joint PDF

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

Let $\mathbf{V} = \begin{bmatrix} Y_1 & Y_4 \end{bmatrix}'$ and $\mathbf{W} = \begin{bmatrix} Y_2 & Y_3 \end{bmatrix}'$. Are $\mathbf{V}$ and $\mathbf{W}$ independent random vectors?

Example 5.7 Solution

We first note that the components of $\mathbf{V}$ are $V_1 = Y_1$, and $V_2 = Y_4$. Also, $W_1 = Y_2$, and $W_2 = Y_3$. Therefore,

$$f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w}) = f_{Y_1,\ldots,Y_4}(v_1, w_1, w_2, v_2) = \begin{cases} 4 & 0 \le v_1 \le w_1 \le 1; \\ & 0 \le w_2 \le v_2 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathbf{V} = \begin{bmatrix} Y_1 & Y_4 \end{bmatrix}'$ and $\mathbf{W} = \begin{bmatrix} Y_2 & Y_3 \end{bmatrix}'$,

$$f_{\mathbf{V}}(\mathbf{v}) = f_{Y_1,Y_4}(v_1, v_2) \qquad f_{\mathbf{W}}(\mathbf{w}) = f_{Y_2,Y_3}(w_1, w_2)$$

In Example 5.5. we found $f_{Y_1,Y_4}(y_1, y_4)$ and $f_{Y_2,Y_3}(y_2, y_3)$ in Equations (5.15) and (5.17). From these marginal PDFs, we have

$$f_{\mathbf{V}}(\mathbf{v}) = \begin{cases} 4(1-v_1)v_2 & 0 \le v_1, v_2 \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} 4w_1(1-w_2) & 0 \le w_1, w_2 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$f_{\mathbf{V}}(\mathbf{v}) f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} 16(1-v_1)v_2 w_1(1-w_2) & 0 \le v_1, v_2, w_1, w_2 \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

which is not equal to $f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w})$. Therefore $\mathbf{V}$ and $\mathbf{W}$ are not independent.

Quiz 5.4

Use the components of $\mathbf{Y} = [Y_1, \ldots, Y_4]'$ in Example 5.7 to construct two independent random vectors $\mathbf{V}$ and $\mathbf{W}$. Prove that $\mathbf{V}$ and $\mathbf{W}$ are independent.

Section 5.5

Functions of Random Vectors

Theorem 5.6

For random variable $W = g(\mathbf{X})$,

Discrete: $$P_W(w) = P[W = w] = \sum_{\mathbf{x}:g(\mathbf{x})=w} P_{\mathbf{X}}(\mathbf{x})$$

Continuous: $$F_W(w) = P[W \leq w] = \int \cdots \int_{g(\mathbf{x})\leq w} f_{\mathbf{X}}(\mathbf{x})\, dx_1 \cdots dx_n.$$

Example 5.8 Problem

Consider an experiment that consists of spinning the pointer on the wheel of circumference 1 meter in Example 3.1 n times and observing Y_n meters, the maximum position of the pointer in the n spins. Find the CDF and PDF of Y_n.

Example 5.8 Solution

If X_i is the position of the pointer on the ith spin, then $Y_n = \max\{X_1, X_2, \ldots, X_n\}$. As a result, $Y_n \le y$ if and only if each $X_i \le y$. This implies

$$P[Y_n \le y] = P[X_1 \le y, X_2 \le y, \ldots X_n \le y].$$

If we assume the spins to be independent, the events $\{X_1 \le y\}$, $\{X_2 \le y\}$, $\ldots$, $\{X_n \le y\}$ are independent events. Thus

$$P[Y_n \le y] = P[X_1 \le y] \cdots P[X_n \le y] = (P[X \le y])^n = (F_X(y))^n.$$

Example 3.2 derives that $F_X(x) = x$ for $0 \le x < 1$. Furthermore, $F_X(x) = 0$ for $x < 0$ and $F_X(x) = 1$ for $x \ge 1$ since $0 \le X \le 1$. Therefore, since the CDF of Y_n is $F_{Y_n}(y) = (F_X(y))^n$, we can write the CDF and corresponding PDF as

$$F_{Y_n}(y) = \begin{cases} 0 & y < 0, \\ y^n & 0 \le y \le 1, \\ 1 & y > 1, \end{cases} \qquad f_{Y_n}(y) = \begin{cases} ny^{n-1} & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 5.7

Let $\mathbf{X}$ be a vector of n iid random variables each with CDF $F_X(x)$ and PDF $f_X(x)$.

(a) The CDF and the PDF of $Y = \max\{X_1, \ldots, X_n\}$ are

$$F_Y(y) = (F_X(y))^n, \qquad f_Y(y) = n(F_X(y))^{n-1} f_X(y).$$

(b) The CDF and the PDF of $W = \min\{X_1, \ldots, X_n\}$ are

$$F_W(w) = 1 - (1 - F_X(w))^n, \qquad f_W(w) = n(1 - F_X(w))^{n-1} f_X(w).$$

Proof: Theorem 5.7

By definition, $f_Y(y) = P[Y \leq y]$. Because Y is the maximum value of $\{X_1, \ldots, X_n\}$, the event $\{Y \leq y\} = \{X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y\}$. Because all the random variables X_i are iid, $\{Y \leq y\}$ is the intersection of n independent events. Each of the events $\{X_i \leq y\}$ has probability $F_X(y)$. The probability of the intersection is the product of the individual probabilities, which implies the first part of the theorem: $F_Y(y) = (F_X(y))^n$. The second part is the result of differentiating $F_Y(y)$ with respect to y. The derivations of $F_W(w)$ and $f_W(w)$ are similar. They begin with the observations that $F_W(w) = 1 - P[W > w]$ and that the event $\{W > w\} = \{X_1 > w, X_2 > w, \ldots X_n > w\}$, which is the intersection of n independent events, each with probability $1 - F_X(w)$.

Theorem 5.8

For a random vector $\mathbf{X}$, the random variable $g(\mathbf{X})$ has expected value

Discrete: $$E[g(\mathbf{X})] = \sum_{x_1 \in S_{X_1}} \cdots \sum_{x_n \in S_{X_n}} g(\mathbf{x}) P_{\mathbf{X}}(\mathbf{x})$$

Continuous: $$E[g(\mathbf{X})] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\, dx_1 \cdots dx_n.$$

Theorem 5.9

When the components of $\mathbf{X}$ are independent random variables,

$$E\left[g_1(X_1)g_2(X_2)\cdots g_n(X_n)\right] = E\left[g_1(X_1)\right]E\left[g_2(X_2)\right]\cdots E\left[g_n(X_n)\right].$$

Proof: Theorem 5.9

When $\mathbf{X}$ is discrete, independence implies $P_{\mathbf{X}}(\mathbf{x}) = P_{X_1}(x_1) \cdots P_{X_n}(x_n)$. This implies

$$\begin{aligned} E[g_1(X_1)\cdots g_n(X_n)] &= \sum_{x_1 \in S_{X_1}} \cdots \sum_{x_n \in S_{X_n}} g_1(x_1)\cdots g_n(x_n) P_{\mathbf{X}}(\mathbf{x}) \\ &= \left(\sum_{x_1 \in S_{X_1}} g_1(x_1) P_{X_1}(x_1) \right) \cdots \left(\sum_{x_n \in S_{X_n}} g_n(x_n) P_{X_n}(x_n) \right) \\ &= E[g_1(X_1)]\, E[g_2(X_2)] \cdots E[g_n(X_n)]. \end{aligned}$$

The derivation is similar for independent continuous random variables.

Theorem 5.10

Given the continuous random vector $\mathbf{X}$, define the derived random vector $\mathbf{Y}$ such that $Y_k = aX_k + b$ for constants $a > 0$ and b. The CDF and PDF of $\mathbf{Y}$ are

$$F_{\mathbf{Y}}(\mathbf{y}) = F_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right), \quad f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{a^n} f_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right).$$

Proof: Theorem 5.10

We observe $\mathbf{Y}$ has CDF $F_{\mathbf{Y}}(\mathbf{y}) = P[aX_1 + b \le y_1, \ldots, aX_n + b \le y_n]$. Since $a > 0$,

$$F_{\mathbf{Y}}(\mathbf{y}) = P\left[X_1 \le \frac{y_1 - b}{a}, \ldots, X_n \le \frac{y_n - b}{a}\right] = F_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right).$$

From Theorem 5.2(b), the joint PDF of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\partial^n F_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)}{\partial y_1 \cdots \partial y_n} = \frac{1}{a^n} f_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right).$$

Theorem 5.11

If $\mathbf{X}$ is a continuous random vector and $\mathbf{A}$ is an invertible matrix, then $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ has PDF

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|\text{det}(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})\right)$$

Proof: Theorem 5.11

Let $B = \{\mathbf{y}|\mathbf{y} \leq \tilde{\mathbf{y}}\}$ so that $F_{\mathbf{Y}}(\tilde{\mathbf{y}}) = \int_B f_{\mathbf{Y}}(\mathbf{y})\, d\mathbf{y}$. Define the vector transformation $\mathbf{x} = T(\mathbf{y}) = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})$. It follows that $\mathbf{Y} \in B$ if and only if $\mathbf{X} \in T(B)$, where $T(B) = \{\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{b} \leq \tilde{\mathbf{y}}\}$ is the image of B under transformation T. This implies

$$F_{\mathbf{Y}}(\tilde{\mathbf{y}}) = P[\mathbf{X} \in T(B)] = \int_{T(B)} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

By the change-of-variable theorem (Math Fact B.13),

$$F_{\mathbf{Y}}(\tilde{\mathbf{y}}) = \int_B f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})\right) \left|\det\left(\mathbf{A}^{-1}\right)\right| \, d\mathbf{y}$$

where $|\det(\mathbf{A}^{-1})|$ is the absolute value of the determinant of $\mathbf{A}^{-1}$. Definition 5.6 for the CDF and PDF of a random vector combined with Theorem 5.2(b) imply that $f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}))|\det(\mathbf{A}^{-1})|$. The theorem follows since $|\det(\mathbf{A}^{-1})| = 1/|\det(\mathbf{A})|$.

Quiz 5.5(A)

A test of light bulbs produced by a machine has three possible outcomes: L, long life; A, average life; and R, reject. The results of different tests are independent. All tests have the following probability model: $P[L] = 0.3$, $P[A] = 0.6$, and $P[R] = 0.1$. Let X_1, X_2, and X_3 be the number of light bulbs that are L, A, and R respectively in five tests. Find the PMF $P_{\mathbf{X}}(\mathbf{x})$; the marginal PMFs $P_{X_1}(x_1)$, $P_{X_2}(x_2)$, and $P_{X_3}(x_3)$; and the PMF of $W = \max(X_1, X_2, X_3)$.

Quiz 5.5(B)

The random vector $\mathbf{X}$ has PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} e^{-x_3} & 0 \le x_1 \le x_2 \le x_3, \\ 0 & \text{otherwise.} \end{cases}$$

Find the PDF of $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$. where $\mathbf{A} = \text{diag}[2, 2, 2]$ and $\mathbf{b} = \begin{bmatrix} 4 & 4 & 4 \end{bmatrix}'$.

Section 5.6

Expected Value Vector and Correlation Matrix

Definition 5.11 Expected Value Vector

The expected value of a random vector $\mathbf{X}$ *is a column vector*

$$E[\mathbf{X}] = \mu_{\mathbf{X}} = \begin{bmatrix} E[X_1] & E[X_2] & \cdots & E[X_n] \end{bmatrix}'.$$

Definition 5.12 Expected Value of a Random Matrix

For a random matrix $\mathbf{A}$ *with the random variable* A_{ij} *as its* i, j*th element,* $E[\mathbf{A}]$ *is a matrix with* i, j*th element* $E[A_{ij}]$.

Definition 5.13 Vector Correlation

The correlation of a random vector $\mathbf{X}$ *is an* $n \times n$ *matrix* $\mathbf{R_X}$ *with* i, j*th element* $R_X(i, j) = E[X_i X_j]$*. In vector noation,*

$$\mathbf{R_X} = E\left[\mathbf{X}\mathbf{X}'\right].$$

Example 5.10

If $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$, the correlation matrix of $\mathbf{X}$ is

$$\mathbf{R_X} = \begin{bmatrix} E\left[X_1^2\right] & E\left[X_1X_2\right] & E\left[X_1X_3\right] \\ E\left[X_2X_1\right] & E\left[X_2^2\right] & E\left[X_2X_3\right] \\ E\left[X_3X_1\right] & E\left[X_3X_2\right] & E\left[X_3^2\right] \end{bmatrix} = \begin{bmatrix} E\left[X_1^2\right] & r_{X_1,X_2} & r_{X_1,X_3} \\ r_{X_2,X_1} & E\left[X_2^2\right] & r_{X_2,X_3} \\ r_{X_3,X_1} & r_{X_3,X_2} & E\left[X_3^2\right] \end{bmatrix}.$$

Definition 5.14 Vector Covariance

The covariance of a random vector $\mathbf{X}$ *is an* $n \times n$ *matrix* $\mathbf{C_X}$ *with components* $C_X(i, j) = \mathrm{Cov}[X_i, X_j]$. *In vector notation,*

$$\mathbf{C_X} = E\left[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})'\right]$$

Example 5.11

If $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$, the covariance matrix of $\mathbf{X}$ is

$$\mathbf{C_X} = \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \mathrm{Cov}\left[X_1, X_3\right] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Var}[X_2] & \mathrm{Cov}\left[X_2, X_3\right] \\ \mathrm{Cov}\left[X_3, X_1\right] & \mathrm{Cov}\left[X_3, X_2\right] & \mathrm{Var}[X_3] \end{bmatrix}$$

Theorem 5.12

For a random vector $\mathbf{X}$ with correlation matrix $\mathbf{R}_\mathbf{X}$, covariance matrix $\mathbf{C}_\mathbf{X}$, and vector expected value $\boldsymbol{\mu}_\mathbf{X}$,

$$\mathbf{C}_\mathbf{X} = \mathbf{R}_\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}'_\mathbf{X}.$$

Proof: Theorem 5.12

The proof is essentially the same as the proof of Theorem 4.16(a) with vectors replacing scalars. Cross multiplying inside the expectation of Definition 5.14 yields

$$\begin{aligned}\mathbf{C_X} &= E\left[\mathbf{X}\mathbf{X}' - \mathbf{X}\boldsymbol{\mu}'_{\mathbf{X}} - \boldsymbol{\mu}_{\mathbf{X}}\mathbf{X}' + \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}\right] \\ &= E\left[\mathbf{X}\mathbf{X}'\right] - E\left[\mathbf{X}\boldsymbol{\mu}'_{\mathbf{X}}\right] - E\left[\boldsymbol{\mu}_{\mathbf{X}}\mathbf{X}'\right] + E\left[\boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}\right].\end{aligned}$$

Since $E[\mathbf{X}] = \boldsymbol{\mu}_{\mathbf{X}}$ is a constant vector,

$$\mathbf{C_X} = \mathbf{R_X} - E\left[\mathbf{X}\right]\boldsymbol{\mu}'_{\mathbf{X}} - \boldsymbol{\mu}_{\mathbf{X}} E\left[\mathbf{X}'\right] + \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}} = \mathbf{R_X} - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}.$$

Example 5.12 Problem

Find the expected value $E[\mathbf{X}]$, the correlation matrix $\mathbf{R}_{\mathbf{X}}$, and the covariance matrix $\mathbf{C}_{\mathbf{X}}$ of the 2-dimensional random vector $\mathbf{X}$ with PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 2 & 0 \leq x_1 \leq x_2 \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example 5.12 Solution

The elements of the expected value vector are

$$E[X_i] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_i\, dx_1 dx_2, \quad i = 1, 2.$$

The integrals are $E[X_1] = 1/3$ and $E[X_2] = 2/3$, so that $\boldsymbol{\mu}_{\mathbf{X}} = E[\mathbf{X}] = \begin{bmatrix} 1/3 & 2/3 \end{bmatrix}'$. The elements of the correlation matrix are

$$E\left[X_1^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^2 f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_1^2\, dx_1 dx_2,$$

$$E\left[X_2^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2^2 f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_2^2\, dx_1 dx_2,$$

$$E[X_1 X_2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_1 x_2\, dx_1 dx_2.$$

[Continued]

Example 5.12 Solution (continued)

These integrals are $E[X_1{}^2] = 1/6$, $E[X_2{}^2] = 1/2$, and $E[X_1 X_2] = 1/4$.

Therefore,

$$\mathbf{R_X} = \begin{bmatrix} 1/6 & 1/4 \\ 1/4 & 1/2 \end{bmatrix}.$$

We use Theorem 5.12 to find the elements of the covariance matrix.

$$\mathbf{C_X} = \mathbf{R_X} - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}'_\mathbf{X} = \begin{bmatrix} 1/6 & 1/4 \\ 1/4 & 1/2 \end{bmatrix} - \begin{bmatrix} 1/9 & 2/9 \\ 2/9 & 4/9 \end{bmatrix} = \begin{bmatrix} 1/18 & 1/36 \\ 1/36 & 1/18 \end{bmatrix}.$$

Definition 5.15 Vector Cross-Correlation

The cross-correlation of random vectors*, $\mathbf{X}$ with n components and $\mathbf{Y}$ with m components, is an $n \times m$ matrix $\mathbf{R}_{\mathbf{XY}}$ with i, jth element $R_{XY}(i, j) = E[X_i Y_j]$, or, in vector notation,*

$$\mathbf{R}_{\mathbf{XY}} = E\left[\mathbf{X}\mathbf{Y}'\right].$$

Definition 5.16 Vector Cross-Covariance

The cross-covariance of a pair of random vectors $\mathbf{X}$ *with* n *components and* $\mathbf{Y}$ *with* m *components is an* $n \times m$ *matrix* $\mathbf{C}_{\mathbf{XY}}$ *with* i, j *th element* $C_{\mathbf{XY}}(i, j) = \mathrm{Cov}[X_i, Y_j]$, *or, in vector notation,*

$$\mathbf{C}_{\mathbf{XY}} = E\left[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})'\right].$$

Theorem 5.13

$\mathbf{X}$ is an n-dimensional random vector with expected value $\boldsymbol{\mu}_{\mathbf{X}}$, correlation $\mathbf{R}_{\mathbf{X}}$, and covariance $\mathbf{C}_{\mathbf{X}}$. The m-dimensional random vector $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$, where $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{b}$ is an m-dimensional vector, has expected value $\boldsymbol{\mu}_{\mathbf{Y}}$, correlation matrix $\mathbf{R}_{\mathbf{Y}}$, and covariance matrix $\mathbf{C}_{\mathbf{Y}}$ given by

$$\begin{aligned}
\boldsymbol{\mu}_{\mathbf{Y}} &= \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}, \\
\mathbf{R}_{\mathbf{Y}} &= \mathbf{A}\mathbf{R}_{\mathbf{X}}\mathbf{A}' + (\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}})\mathbf{b}' + \mathbf{b}(\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}})' + \mathbf{b}\mathbf{b}', \\
\mathbf{C}_{\mathbf{Y}} &= \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}'.
\end{aligned}$$

Proof: Theorem 5.13

We derive the formulas for the expected value and covariance of $\mathbf{Y}$. The derivation for the correlation is similar. First, the expected value of $\mathbf{Y}$ is

$$\boldsymbol{\mu}_{\mathbf{Y}} = E\left[\mathbf{A}\mathbf{X} + \mathbf{b}\right] = \mathbf{A}E\left[\mathbf{X}\right] + E\left[\mathbf{b}\right] = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}.$$

It follows that $\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}} = \mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})$. This implies

$$\begin{aligned}
\mathbf{C}_{\mathbf{Y}} &= E\left[(\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}))(\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}))'\right] \\
&= E\left[\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}))(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'\mathbf{A}'\right] = \mathbf{A}E\left[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'\right]\mathbf{A}' = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}'.
\end{aligned}$$

Example 5.13 Problem

Given random vector $\mathbf{X}$ defined in Example 5.12, let $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 6 & 3 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 0 \\ -2 \\ -2 \end{bmatrix}.$$

Find the expected value $\boldsymbol{\mu}_{\mathbf{Y}}$, the correlation $\mathbf{R}_{\mathbf{Y}}$, and the covariance $\mathbf{C}_{\mathbf{Y}}$.

Example 5.13 Solution

From the matrix operations of Theorem 5.13, we obtain $\boldsymbol{\mu}_{\mathbf{Y}} = \begin{bmatrix} 1/3 & 2 & 3 \end{bmatrix}'$ and

$$\mathbf{R}_{\mathbf{Y}} = \begin{bmatrix} 1/6 & 13/12 & 4/3 \\ 13/12 & 7.5 & 9.25 \\ 4/3 & 9.25 & 12.5 \end{bmatrix}; \qquad \mathbf{C}_{\mathbf{Y}} = \begin{bmatrix} 1/18 & 5/12 & 1/3 \\ 5/12 & 3.5 & 3.25 \\ 1/3 & 3.25 & 3.5 \end{bmatrix}.$$

Theorem 5.14

The vectors $\mathbf{X}$ and $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ have cross-correlation $\mathbf{R_{XY}}$ and cross-covariance $\mathbf{C_{XY}}$ given by

$$\mathbf{R_{XY}} = \mathbf{R_X}\mathbf{A}' + \boldsymbol{\mu}_{\mathbf{X}}\mathbf{b}', \qquad \mathbf{C_{XY}} = \mathbf{C_X}\mathbf{A}'.$$

Example 5.14 Problem

Continuing Example 5.13 for random vectors $\mathbf{X}$ and $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$, calculate

(a) The cross-correlation matrix $\mathbf{R}_{\mathbf{XY}}$ and the cross-covariance matrix $\mathbf{C}_{\mathbf{XY}}$.

(b) The correlation coefficients ρ_{Y_1,Y_3} and ρ_{X_2,Y_1}.

Example 5.14 Solution

(a) Direct matrix calculation using Theorem 5.14 yields

$$\mathbf{R_{XY}} = \begin{bmatrix} 1/6 & 13/12 & 4/3 \\ 1/4 & 5/3 & 29/12 \end{bmatrix}; \qquad \mathbf{C_{XY}} = \begin{bmatrix} 1/18 & 5/12 & 1/3 \\ 1/36 & 1/3 & 5/12 \end{bmatrix}.$$

(b) Referring to Definition 4.8 and recognizing that $\text{Var}[Y_i] = C_{\mathbf{Y}}(i,i)$, we have

$$\rho_{Y_1,Y_3} = \frac{\text{Cov}\left[Y_1, Y_3\right]}{\sqrt{\text{Var}[Y_1]\,\text{Var}[Y_3]}} = \frac{C_Y(1,3)}{\sqrt{C_Y(1,1)C_Y(3,3)}} = 0.756$$

Similarly,

$$\rho_{X_2,Y_1} = \frac{\text{Cov}\left[X_2, Y_1\right]}{\sqrt{\text{Var}[X_2]\,\text{Var}[Y_1]}} = \frac{C_{\mathbf{XY}}(2,1)}{\sqrt{C_{\mathbf{X}}(2,2)C_{\mathbf{Y}}(1,1)}} = 1/2.$$

Section 5.7

Gaussian Random Vectors

Definition 5.17 Gaussian Random Vector

$\mathbf{X}$ *is the Gaussian* $(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ *random vector with expected value* $\boldsymbol{\mu}_{\mathbf{X}}$ *and covariance* $\mathbf{C}_{\mathbf{X}}$ *if and only if*

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}[det(\mathbf{C}_{\mathbf{X}})]^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{X}})'\mathbf{C}_{\mathbf{X}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{X}})\right)$$

where det$(\mathbf{C}_{\mathbf{X}})$*, the determinant of* $\mathbf{C}_{\mathbf{X}}$*, satisfies det*$(\mathbf{C}_{\mathbf{X}}) > 0$.

Theorem 5.15

A Gaussian random vector $\mathbf{X}$ has independent components if and only if $\mathbf{C_X}$ is a diagonal matrix.

Proof: Theorem 5.15

First, if the components of $\mathbf{X}$ are independent, then for $i \neq j$, X_i and X_j are independent. By Theorem 4.27(c), $\text{Cov}[X_i, X_j] = 0$. Hence the off-diagonal terms of $\mathbf{C_X}$ are all zero. If $\mathbf{C_X}$ is diagonal, then

$$\mathbf{C_X} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix} \quad \text{and} \quad \mathbf{C_X}^{-1} = \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_n^2 \end{bmatrix}.$$

It follows that $\mathbf{C_X}$ has determinant $\det(\mathbf{C_X}) = \prod_{i=1}^n \sigma_i^2$ and that

$$(\mathbf{x} - \boldsymbol{\mu}_\mathbf{X})'\mathbf{C_X}^{-1}(\mathbf{x} - \boldsymbol{\mu}_\mathbf{X}) = \sum_{i=1}^n \frac{(X_i - \mu_i)^2}{\sigma_i^2}.$$

From Definition 5.17, we see that

$$\begin{aligned} f_\mathbf{X}(\mathbf{x}) &= \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i^2} \exp\left(- \sum_{i=1}^n (x_i - \mu_i)/2\sigma_i^2 \right) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-(x_i - \mu_i)^2/2\sigma_i^2\right). \end{aligned}$$

Thus $f_\mathbf{X}(\mathbf{x}) = \prod_{i=1}^n f_{X_i}(x_i)$, implying $X_1, \ldots, X_n$ are independent.

Example 5.15 Problem

Consider the outdoor temperature at a certain weather station. On May 5, the temperature measurements in units of degrees Fahrenheit taken at 6 AM, 12 noon, and 6 PM are all Gaussian random variables, X_1, X_2, X_3 with variance 16 degrees^2. The expected values are 50 degrees, 62 degrees, and 58 degrees respectively. The covariance matrix of the three measurements is

$$\mathbf{C_X} = \begin{bmatrix} 16.0 & 12.8 & 11.2 \\ 12.8 & 16.0 & 12.8 \\ 11.2 & 12.8 & 16.0 \end{bmatrix}.$$

(a) Write the joint PDF of X_1, X_2 using the algebraic notation of Definition 4.17.

(b) Write the joint PDF of X_1, X_2 using vector notation.

(c) Write the joint PDF of $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$ using vector notation.

Example 5.15 Solution

(a) First we note that X_1 and X_2 have expected values $\mu_1 = 50$ and $\mu_2 = 62$, variances $\sigma_1^2 = \sigma_2^2 = 16$, and covariance $\text{Cov}[X_1, X_2] = 12.8$. It follows from Definition 4.8 that the correlation coefficient is

$$\rho_{X_1,X_2} = \frac{\text{Cov}\,[X_1, X_2]}{\sigma_1\sigma_2} = \frac{12.8}{16} = 0.8.$$

From Definition 4.17, the joint PDF is

$$f_{X_1,X_2}(x_1, x_2) = \frac{\exp\left(-\frac{(x_1-50)^2-1.6(x_1-50)(x_2-62)+(x_2-62)^2}{19.2}\right)}{60.3}.$$

(b) Let $\mathbf{W} = \begin{bmatrix} X_1 & X_2 \end{bmatrix}'$ denote a vector representation for random variables X_1 and X_2. From the covariance matrix $\mathbf{C_X}$, we observe that the 2×2 submatrix in the upper left corner is the covariance matrix of the random vector $\mathbf{W}$. Thus [Continued]

Example 5.15 Solution (continued)

$$\boldsymbol{\mu}_{\mathbf{W}} = \begin{bmatrix} 50 \\ 62 \end{bmatrix}, \qquad \mathbf{C}_{\mathbf{W}} = \begin{bmatrix} 16.0 & 12.8 \\ 12.8 & 16.0 \end{bmatrix}.$$

We observe that $\det(\mathbf{C}_{\mathbf{W}}) = 92.16$ and $\det(\mathbf{C}_{\mathbf{W}})^{1/2} = 9.6$. From Definition 5.17, the joint PDF of $\mathbf{W}$ is

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{1}{60.3} \exp\left(-\frac{1}{2}(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{W}})^T \mathbf{C}_{\mathbf{W}}^{-1} (\mathbf{w} - \boldsymbol{\mu}_{\mathbf{W}}) \right).$$

(c) For the joint PDF of $\mathbf{X}$, we note that $\mathbf{X}$ has expected value $\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} 50 & 62 & 58 \end{bmatrix}'$ and that $\det(\mathbf{C}_{\mathbf{X}})^{1/2} = 22.717$. Thus

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{357.8} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}}) \right).$$

Theorem 5.16

Given an n-dimensional Gaussian random vector $\mathbf{X}$ with expected value $\boldsymbol{\mu}_{\mathbf{X}}$ and covariance $\mathbf{C}_{\mathbf{X}}$, and an $m \times n$ matrix $\mathbf{A}$ with rank$(\mathbf{A}) = m$,

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$$

is an m-dimensional Gaussian random vector with expected value $\boldsymbol{\mu}_{\mathbf{Y}} = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}$ and covariance $\mathbf{C}_{\mathbf{Y}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}'$.

Proof: Theorem 5.16

The proof of Theorem 5.13 contains the derivations of $\boldsymbol{\mu}_{\mathbf{Y}}$ and $\mathbf{C}_{\mathbf{Y}}$. Our proof that $\mathbf{Y}$ has a Gaussian PDF is confined to the special case when $m = n$ and $\mathbf{A}$ is an invertible matrix. The case of $m < n$ is addressed in Problem 5.7.9. When $m = n$, we use Theorem 5.11 to write

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b})\right) \\ &= \frac{\exp\left(-\frac{1}{2}[\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b})-\boldsymbol{\mu}_{\mathbf{X}}]'\mathbf{C}_{\mathbf{X}}^{-1}[\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b})-\boldsymbol{\mu}_{\mathbf{X}}]\right)}{(2\pi)^{n/2}\,|\det(\mathbf{A})|\,|\det(\mathbf{C}_{\mathbf{X}})|^{1/2}}. \end{aligned}$$

In the exponent of $f_{\mathbf{Y}}(\mathbf{y})$, we observe that

$$\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b})-\boldsymbol{\mu}_{\mathbf{X}} = \mathbf{A}^{-1}[\mathbf{y}-(\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}}+\mathbf{b})] = \mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}}),$$

since $\boldsymbol{\mu}_{\mathbf{Y}} = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}$.

[Continued]

Proof: Theorem 5.16 (continued)

Applying (5.79) to (5.78) yields

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}[\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})]'\mathbf{C}_{\mathbf{X}}^{-1}[\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})]\right)}{(2\pi)^{n/2}\,|\det(\mathbf{A})|\,|\det(\mathbf{C}_{\mathbf{X}})|^{1/2}}.$$

Using the identities $|\det(\mathbf{A})||\det(\mathbf{C}_{\mathbf{X}})|^{1/2} = |\det(\mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}')|^{1/2}$ and $(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$, we can write

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})'(\mathbf{A}')^{-1}\mathbf{C}_{\mathbf{X}}^{-1}\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})\right)}{(2\pi)^{n/2}\left|\det\left(\mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}'\right)\right|^{1/2}}.$$

Since $(\mathbf{A}')^{-1}\mathbf{C}_{\mathbf{X}}^{-1}\mathbf{A}^{-1} = (\mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}')^{-1}$, we see from Equation (5.81) that $\mathbf{Y}$ is a Gaussian vector with expected value $\boldsymbol{\mu}_{\mathbf{Y}}$ and covariance matrix $\mathbf{C}_{\mathbf{Y}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}'$.

Example 5.16 Problem

Continuing Example 5.15, use the formula $Y_i = (5/9)(X_i - 32)$ to convert the three temperature measurements to degrees Celsius.

(a) What is $\boldsymbol{\mu}_{\mathbf{Y}}$, the expected value of random vector $\mathbf{Y}$?

(b) What is $\mathbf{C}_{\mathbf{Y}}$, the covariance of random vector $\mathbf{Y}$?

(c) Write the joint PDF of $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix}'$ using vector notation.

Example 5.16 Solution

(a) In terms of matrices, we observe that $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ where

$$\mathbf{A} = \begin{bmatrix} 5/9 & 0 & 0 \\ 0 & 5/9 & 0 \\ 0 & 0 & 5/9 \end{bmatrix}, \qquad \mathbf{b} = -\frac{160}{9}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

(b) Since $\boldsymbol{\mu}_\mathbf{X} = \begin{bmatrix} 50 & 62 & 58 \end{bmatrix}'$, from Theorem 5.16,

$$\boldsymbol{\mu}_\mathbf{Y} = \mathbf{A}\boldsymbol{\mu}_\mathbf{X} + \mathbf{b} = \begin{bmatrix} 10 \\ 50/3 \\ 130/9 \end{bmatrix}.$$

(c) The covariance of $\mathbf{Y}$ is $\mathbf{C}_\mathbf{Y} = \mathbf{A}\mathbf{C}_\mathbf{X}\mathbf{A}'$. We note that $\mathbf{A} = \mathbf{A}' = (5/9)\mathbf{I}$ where $\mathbf{I}$ is the 3×3 identity matrix. Thus $\mathbf{C}_\mathbf{Y} = (5/9)^2\mathbf{C}_\mathbf{X}$ and $\mathbf{C}_\mathbf{Y}^{-1} = (9/5)^2\mathbf{C}_\mathbf{X}^{-1}$. The PDF of $\mathbf{Y}$ is

$$f_\mathbf{Y}(\mathbf{y}) = \frac{1}{24.47}\exp\left(-\frac{81}{50}(\mathbf{y} - \boldsymbol{\mu}_\mathbf{Y})^T\mathbf{C}_\mathbf{X}^{-1}(\mathbf{y} - \boldsymbol{\mu}_\mathbf{Y})\right).$$

Definition 5.18 Standard Normal Random Vector

The n-dimensional standard normal random vector $\mathbf{Z}$ *is the n-dimensional Gaussian random vector with* $E[\mathbf{Z}] = \mathbf{0}$ *and* $\mathbf{C}_{\mathbf{Z}} = \mathbf{I}$.

Theorem 5.17

For a Gaussian $(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ random vector, let $\mathbf{A}$ be an $n \times n$ matrix with the property $\mathbf{A}\mathbf{A}' = \mathbf{C}_\mathbf{X}$. The random vector

$$\mathbf{Z} = \mathbf{A}^{-1}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})$$

is a standard normal random vector.

Proof: Theorem 5.17

Applying Theorem 5.16 with $\mathbf{A}$ replaced by $\mathbf{A}^{-1}$, and $\mathbf{b} = \mathbf{A}^{-1}\mu_{\mathbf{X}}$, we have that $\mathbf{Z}$ is a Gaussian random vector with expected value

$$E[\mathbf{Z}] = E\left[\mathbf{A}^{-1}(\mathbf{X} - \mu_{\mathbf{X}})\right] = \mathbf{A}^{-1}E\left[\mathbf{X} - \mu_{\mathbf{X}}\right] = \mathbf{0},$$

and covariance

$$\mathbf{C}_{\mathbf{Z}} = \mathbf{A}^{-1}\mathbf{C}_{\mathbf{X}}(\mathbf{A}^{-1})' = \mathbf{A}^{-1}\mathbf{A}\mathbf{A}'(\mathbf{A}')^{-1} = \mathbf{I}.$$

Theorem 5.18

Given the n-dimensional standard normal random vector $\mathbf{Z}$, an invertible $n \times n$ matrix $\mathbf{A}$, and an n-dimensional vector $\mathbf{b}$,

$$\mathbf{X} = \mathbf{A}\mathbf{Z} + \mathbf{b}$$

is an n-dimensional Gaussian random vector with expected value $\boldsymbol{\mu}_{\mathbf{X}} = \mathbf{b}$ and covariance matrix $\mathbf{C}_{\mathbf{X}} = \mathbf{A}\mathbf{A}'$.

Proof: Theorem 5.18

By Theorem 5.16, $\mathbf{X}$ is a Gaussian random vector with expected value

$$\mu_{\mathbf{X}} = E[\mathbf{X}] = E\left[\mathbf{A}\mathbf{Z} + \mu_{\mathbf{X}}\right] = \mathbf{A}E[\mathbf{Z}] + \mathbf{b} = \mathbf{b}.$$

The covariance of $\mathbf{X}$ is

$$\mathbf{C}_{\mathbf{X}} = \mathbf{A}\mathbf{C}_{\mathbf{Z}}\mathbf{A}' = \mathbf{A}\mathbf{I}\mathbf{A}' = \mathbf{A}\mathbf{A}'.$$

Theorem 5.19

For a Gaussian vector $\mathbf{X}$ with covariance $\mathbf{C_X}$, there always exists a matrix $\mathbf{A}$ such that $\mathbf{C_X} = \mathbf{A}\mathbf{A}'$.

Proof: Theorem 5.19

To verify this fact, we connect some simple facts:

- In Problem 5.6.9, we ask the reader to show that every random vector $\mathbf{X}$ has a positive semidefinite covariance matrix $\mathbf{C_X}$. By Math Fact B.17, every eigenvalue of $\mathbf{C_X}$ is nonnegative.
- The definition of the Gaussian vector PDF requires the existence of $\mathbf{C_X^{-1}}$. Hence, for a Gaussian vector $\mathbf{X}$, all eigenvalues of $\mathbf{C_X}$ are nonzero. From the previous step, we observe that all eigenvalues of $\mathbf{C_X}$ must be positive.
- Since $\mathbf{C_X}$ is a real symmetric matrix, Math Fact B.15 says it has a singular value decomposition (SVD) $\mathbf{C_X} = \mathbf{UDU}'$ where $\mathbf{D} = \text{diag}[d_1, \ldots, d_n]$ is the diagonal matrix of eigenvalues of $\mathbf{C_X}$. Since each d_i is positive, we can define $\mathbf{D}^{1/2} = \text{diag}[\sqrt{d_1}, \ldots, \sqrt{d_n}]$, and we can write

$$\mathbf{C_X} = \mathbf{UD}^{1/2}\mathbf{D}^{1/2}\mathbf{U}' = \left(\mathbf{UD}^{1/2}\right)\left(\mathbf{UD}^{1/2}\right)'.$$

We see that $\mathbf{A} = \mathbf{UD}^{1/2}$.