# Section 8.1

## Vector Notation

# **Definition 8.1** 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 8.2 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 Vectors: Notation

- Following our convention for random variables, the uppercase $\mathbf{X}$ is the random vector and the lowercase $\mathbf{x}$ is a sample value of $\mathbf{X}$.

- However, we also use boldface capitals such as $\mathbf{A}$ and $\mathbf{B}$ to denote matrices with components that are not random variables.

- It will be clear from the context whether $\mathbf{A}$ is a matrix of numbers, a matrix of random variables, or a random vector.

# Definition 8.3 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*

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1,\ldots,X_n}(x_1,\ldots,x_n).$$

## Definition 8.4 Probability Functions of a Pair of 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 8.1 Problem

Random vector $\mathbf{X}$ has PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 6e^{-\mathbf{a}'\mathbf{x}} & \mathbf{x} \geq 0, \\ 0 & \text{otherwise}, \end{cases} \tag{8.1}$$

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

# Example 8.1 Solution

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

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 6e^{-x_1-2x_2-3x_3} & x_i \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{8.2}$$

Applying Definition 8.4, we integrate the PDF with respect to the three variables to obtain

$$F_{\mathbf{X}}(\mathbf{x}) = \begin{cases} (1-e^{-x_1})(1-e^{-2x_2})(1-e^{-3x_3}) & x_i \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{8.3}$$

# Quiz 8.1

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

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

# Quiz 8.1 Solution

By definition of $\mathbf{A}$, $Y_1 = X_1$, $Y_2 = X_2 - X_1$ and $Y_3 = X_3 - X_2$. Since $0 < X_1 < X_2 < X_3$, each $Y_i$ must be a strictly positive integer. Thus, for $y_1, y_2, y_3 \in \{1, 2, \ldots\}$,

$$\begin{aligned}
P_{\mathbf{Y}}(\mathbf{y}) &= \mathrm{P}\left[Y_1 = y_1, Y_2 = y_2, Y_3 = y_3\right] \\
&= \mathrm{P}\begin{bmatrix} X_1 = & y_1, \\ X_2 - X_1 = & y_2, \\ X_3 - X_2 = & y_3 \end{bmatrix} \\
&= \mathrm{P}\begin{bmatrix} X_1 = & y_1, \\ X_2 = & y_2 + y_1, \\ X_3 = & y_3 + y_2 + y_1 \end{bmatrix} \\
&= P_{\mathbf{X}}(y_1, y_2 + y_1, y_3 + y_2 + y_1) \\
&= (1-p)^3 p^{y_1 + y_2 + y_3}. \qquad (1)
\end{aligned}$$

With $\mathbf{a} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}'$ and $q = 1 - p$, the joint PMF of $\mathbf{Y}$ is

$$P_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} qp^{\mathbf{a}'\mathbf{y}} & y_1, y_2, y_3 \in \{1, 2, \ldots\}, \\ 0 & \text{otherwise.} \end{cases}$$

# Section 8.2

## Independent Random Variables and Random Vectors

# Definition 8.5 Independent Random Vectors

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

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

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

# Example 8.2 Problem

As in Example 5.22, random variables $Y_1, \ldots, Y_4$ have the joint PDF

$$f_{Y_1,\ldots,Y_4}(y_1,\ldots,y_4) = \begin{cases} 4 & 0 \le y_1 \le y_2 \le 1, 0 \le y_3 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \quad (8.4)$$

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 8.2 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,

$$f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w}) = f_{Y_1,\ldots,Y_4}(v_1,w_1,w_2,v_2) = \begin{cases} 4 & 0 \le v_1 \le w_1 \le 1; \\ & 0 \le w_2 \le v_2 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8.5}$$

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

$$f_{\mathbf{V}}(\mathbf{v}) = f_{Y_1,Y_4}(v_1,v_2), \qquad f_{\mathbf{W}}(\mathbf{w}) = f_{Y_2,Y_3}(w_1,w_2). \tag{8.6}$$

In Example 5.22, we found the marginal PDFs $f_{Y_1,Y_4}(y_1,y_4)$ and $f_{Y_2,Y_3}(y_2,y_3)$ in Equations (5.78) and (5.80). From these marginal PDFs, we have

$$f_{\mathbf{V}}(\mathbf{v}) = \begin{cases} 4(1-v_1)v_2 & 0 \le v_1, v_2 \le 1, \\ 0 & \text{otherwise,} \end{cases} \tag{8.7}$$

$$f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} 4w_1(1-w_2) & 0 \le w_1, w_2 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8.8}$$

Therefore,

$$f_{\mathbf{V}}(\mathbf{v})\, f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} 16(1-v_1)v_2w_1(1-w_2) & 0 \le v_1, v_2, w_1, w_2 \le 1, \\ 0 & \text{otherwise,} \end{cases} \tag{8.9}$$

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

# Quiz 8.2

Use the components of $\mathbf{Y} = \left[Y_1, \ldots, Y_4\right]'$ in Example 8.2 to construct two independent random vectors $\mathbf{V}$ and $\mathbf{W}$. Prove that $\mathbf{V}$ and $\mathbf{W}$ are independent.

# Quiz 8.2 Solution

In the PDF $f_{\mathbf{Y}}(\mathbf{y})$, the components have dependencies as a result of the ordering constraints $Y_1 \leq Y_2$ and $Y_3 \leq Y_4$. We can separate these constraints by creating the vectors

$$\mathbf{V} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \qquad \mathbf{W} = \begin{bmatrix} Y_3 \\ Y_4 \end{bmatrix}. \tag{1}$$

The joint PDF of $\mathbf{V}$ and $\mathbf{W}$ is

$$f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w}) = \begin{cases} 4 & \begin{array}{l} 0 \leq v_1 \leq v_2 \leq 1; \\ 0 \leq w_1 \leq w_2 \leq 1, \end{array} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

We must verify that $\mathbf{V}$ and $\mathbf{W}$ are independent. For $0 \leq v_1 \leq v_2 \leq 1$,

$$\begin{aligned} f_{\mathbf{V}}(\mathbf{v}) &= \iint f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w})\, dw_1\, dw_2 \\ &= \int_0^1 \left( \int_{w_1}^1 4\, dw_2 \right) dw_1 = \int_0^1 4(1 - w_1)\, dw_1 = 2. \end{aligned} \tag{3}$$

[Continued]

# Quiz 8.2 Solution (Continued 2)

Similarly, for $0 \le w_1 \le w_2 \le 1$,

$$f_{\mathbf{W}}(\mathbf{w}) = \iint f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w})\ dv_1\,dv_2$$

$$= \int_0^1 \left( \int_{v_1}^1 4\,dv_2 \right) dv_1 = 2. \tag{4}$$

It follows that $\mathbf{V}$ and $\mathbf{W}$ have PDFs

$$f_{\mathbf{V}}(\mathbf{v}) = \begin{cases} 2 & 0 \le v_1 \le v_2 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

$$f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} 2 & 0 \le w_1 \le w_2 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

It is easy to verify that $f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w}) = f_{\mathbf{V}}(\mathbf{v}) f_{\mathbf{W}}(\mathbf{w})$, confirming that $\mathbf{V}$ and $\mathbf{W}$ are independent vectors.

# Section 8.3

## Functions of Random Vectors

# Theorem 8.1

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

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

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

# Example 8.3 Problem

Consider an experiment that consists of spinning the pointer on the wheel of circumference 1 meter in Example 4.1 $n$ 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 8.3 Solution

If $X_i$ is the position of the pointer on spin $i$, 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

$$F_{Y_n}(y) = \mathrm{P}\left[Y_n \le y\right] = \mathrm{P}\left[X_1 \le y, X_2 \le y, \ldots X_n \le y\right]. \tag{8.10}$$

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

$$F_{Y_n}(y) = \mathrm{P}\left[X_1 \le y\right] \cdots \mathrm{P}\left[X_n \le y\right] = \left(\mathrm{P}\left[X \le y\right]\right)^n = \left(F_X(y)\right)^n. \tag{8.11}$$

Example 4.2 derives Equation (4.8):

$$F_X(x) = \begin{cases} 0 & x < 0, \\ x & 0 \le x < 1, \\ 1 & x \ge 1. \end{cases} \tag{8.12}$$

Equations (8.11) and (8.12) imply that the CDF and corresponding PDF are

$$F_{Y_n}(y) = \begin{cases} 0 & y < 0, \\ y^n & 0 \le y \le 1, \\ 1 & y > 1, \end{cases} \qquad f_{Y_n}(y) = \begin{cases} ny^{n-1} & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8.13}$$

# Theorem 8.2

Let $\mathbf{X}$ be a vector of $n$ iid continuous 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 8.2

By definition, $F_Y(y) = \mathrm{P}[Y \leq y]$. 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 $F_W(w) = 1 - \mathrm{P}[W > w]$ 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 8.3

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

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

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

# Theorem 8.4

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

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

# Proof: Theorem 8.4

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

$$\mathsf{E}\left[g_1(X_1) \cdots g_n(X_n)\right] = \sum_{x_1 \in S_{X_1}} \cdots \sum_{x_n \in S_{X_n}} g_1(x_1) \cdots g_n(x_n) P_{\mathbf{X}}(\mathbf{x}) \tag{8.14}$$

$$= \left( \sum_{x_1 \in S_{X_1}} g_1(x_1) P_{X_1}(x_1) \right) \cdots \left( \sum_{x_n \in S_{X_n}} g_n(x_n) P_{X_n}(x_n) \right) \tag{8.15}$$

$$= \mathsf{E}\left[g_1(X_1)\right] \mathsf{E}\left[g_2(X_2)\right] \cdots \mathsf{E}\left[g_n(X_n)\right]. \tag{8.16}$$

The derivation is similar for independent continuous random variables.

# Theorem 8.5

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 8.5

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

$$F_{\mathbf{Y}}(\mathbf{y}) = \text{P}\left[X_1 \leq \frac{y_1 - b}{a}, \ldots, X_n \leq \frac{y_n - b}{a}\right] = F_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right). \tag{8.17}$$

Definition 5.13 defines the joint PDF of $\mathbf{Y}$,

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\partial^n F_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)}{\partial y_1 \cdots \partial y_n} = \frac{1}{a^n} f_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right). \tag{8.18}$$

# Theorem 8.6

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

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

# Proof: Theorem 8.6

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

$$F_{\mathbf{Y}}(\tilde{\mathbf{y}}) = \mathrm{P}\left[\mathbf{X} \in T(B)\right] = \int_{T(B)} f_{\mathbf{X}}(\mathbf{x})\ d\mathbf{x} \tag{8.19}$$

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

$$F_{\mathbf{Y}}(\tilde{\mathbf{y}}) = \int_B f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})\right) \left|\det\left(\mathbf{A}^{-1}\right)\right|\ d\mathbf{y} \tag{8.20}$$

where $|\det(\mathbf{A}^{-1})|$ is the absolute value of the determinant of $\mathbf{A}^{-1}$. Definition 8.3 for the CDF and PDF of a random vector combined with Theorem 5.23(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 8.3(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: $\mathrm{P}[L] = 0.3$, $\mathrm{P}[A] = 0.6$, and $\mathrm{P}[R] = 0.1$. 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 8.3(A) Solution

Referring to Theorem 2.9, each test is a subexperiment with three possible outcomes: $L$, $A$ and $R$. In five trials, the vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$ indicating the number of outcomes of each subexperiment has the multinomial PMF

$$P_{\mathbf{X}}(\mathbf{x}) = \binom{5}{x_1, x_2, x_3} 0.3^{x_1} 0.6^{x_2} 0.1^{x_3}.$$

We can find the marginal PMF for each $X_i$ from the joint PMF $P_{\mathbf{X}}(\mathbf{x})$; however it is simpler to just start from first principles and observe that $X_1$ is the number of occurrences of $L$ in five independent tests. If we view each test as a trial with success probability $\mathrm{P}[L] = 0.3$, we see that $X_1$ is a binomial $(n, p) = (5, 0.3)$ random variable. Similarly, $X_2$ is a binomial $(5, 0.6)$ random variable and $X_3$ is a binomial $(5, 0.1)$ random variable. That is, for $p_1 = 0.3$, $p_2 = 0.6$ and $p_3 = 0.1$,

$$P_{X_i}(x) = \binom{5}{x} p_i^x (1 - p_i)^{5-x}. \tag{1}$$

[Continued]

# Quiz 8.3(A) Solution (Continued 2)

From the marginal PMFs, we see that $X_1$, $X_2$ and $X_3$ are not independent. Hence, we must use Theorem 8.1 to find the PMF of $W$. In particular, since $X_1 + X_2 + X_3 = 5$ and since each $X_i$ is non-negative, $P_W(0) = P_W(1) = 0$. Furthermore,

$$\begin{aligned} P_W(2) &= P_{\mathbf{X}}(1,2,2) + P_{\mathbf{X}}(2,1,2) + P_{\mathbf{X}}(2,2,1) \\ &= \frac{5!0.3(0.6)^2(0.1)^2}{2!2!1!} + \frac{5!0.3^2(0.6)(0.1)^2}{2!2!1!} + \frac{5!0.3^2(0.6)^2(0.1)}{2!2!1!} \\ &= 0.1458. \end{aligned} \tag{2}$$

In addition, for $w = 3$, $w = 4$, and $w = 5$, the event $W = w$ occurs if and only if one of the mutually exclusive events $X_1 = w$, $X_2 = w$, or $X_3 = w$ occurs. Thus,

$$P_W(3) = \sum_{i=1}^{3} P_{X_i}(3) = 0.486, \tag{3}$$

$$P_W(4) = \sum_{i=1}^{3} P_{X_i}(4) = 0.288, \tag{4}$$

$$P_W(5) = \sum_{i=1}^{3} P_{X_i}(5) = 0.0802. \tag{5}$$

# Quiz 8.3(B)

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

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} e^{-x_3} & 0 \le x_1 \le x_2 \le x_3, \\ 0 & \text{otherwise.} \end{cases} \tag{8.21}$$

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

# Quiz 8.3(B) Solution

Since each $Y_i = 2X_i + 4$, we can apply Theorem 8.5 to write

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{2^3} f_{\mathbf{X}}\left(\frac{y_1 - 4}{2}, \frac{y_2 - 4}{2}, \frac{y_3 - 4}{2}\right)$$
$$= \begin{cases} (1/8)e^{-(y_3-4)/2} & 4 \leq y_1 \leq y_2 \leq y_3, \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

Note that for other matrices $\mathbf{A}$, the constraints on $\mathbf{y}$ resulting from the constraints $0 \leq X_1 \leq X_2 \leq X_3$ can be much more complicated.

# Expected Value Vector and Correlation Matrix

# **Definition 8.6** Expected Value Vector

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

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

# Example 8.4 Problem

If $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$, what are the components of $\mathbf{X}\mathbf{X}'$?

# Example 8.4 Solution

$$\mathbf{X}\mathbf{X}' = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} = \begin{bmatrix} X_1^2 & X_1X_2 & X_1X_3 \\ X_2X_1 & X_2^2 & X_2X_3 \\ X_3X_1 & X_3X_2 & X_3^2 \end{bmatrix}. \tag{8.22}$$

## Definition 8.7 Expected Value of a Random Matrix

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

# Definition 8.8 Vector Correlation

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

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

# Example 8.5

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

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

# Definition 8.9 Vector Covariance

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

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

# Example 8.6

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

$$\mathbf{C}_{\mathbf{X}} = \begin{bmatrix} \text{Var}[X_1] & \text{Cov}\,[X_1, X_2] & \text{Cov}\,[X_1, X_3] \\ \text{Cov}\,[X_2, X_1] & \text{Var}[X_2] & \text{Cov}\,[X_2, X_3] \\ \text{Cov}\,[X_3, X_1] & \text{Cov}\,[X_3, X_2] & \text{Var}[X_3] \end{bmatrix} \tag{8.23}$$

# Theorem 8.7

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

$$\mathbf{C_X} = \mathbf{R_X} - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}'_\mathbf{X}.$$

# Proof: Theorem 8.7

The proof is essentially the same as the proof of Theorem 5.16(a), with vectors replacing scalars. Cross multiplying inside the expectation of Definition 8.9 yields

$$\begin{aligned} \mathbf{C}_{\mathbf{X}} &= \mathsf{E}\left[\mathbf{X}\mathbf{X}' - \mathbf{X}\boldsymbol{\mu}'_{\mathbf{X}} - \boldsymbol{\mu}_{\mathbf{X}}\mathbf{X}' + \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}\right] \\ &= \mathsf{E}\left[\mathbf{X}\mathbf{X}'\right] - \mathsf{E}\left[\mathbf{X}\boldsymbol{\mu}'_{\mathbf{X}}\right] - \mathsf{E}\left[\boldsymbol{\mu}_{\mathbf{X}}\mathbf{X}'\right] + \mathsf{E}\left[\boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}\right]. \end{aligned} \tag{8.24}$$

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

$$\mathbf{C}_{\mathbf{X}} = \mathbf{R}_{\mathbf{X}} - \mathsf{E}\left[\mathbf{X}\right]\boldsymbol{\mu}'_{\mathbf{X}} - \boldsymbol{\mu}_{\mathbf{X}}\,\mathsf{E}\left[\mathbf{X}'\right] + \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}} = \mathbf{R}_{\mathbf{X}} - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}. \tag{8.25}$$

# Example 8.7 Problem

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

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 2 & 0 \leq x_1 \leq x_2 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8.26}$$

# Example 8.7 Solution

The elements of the expected value vector are

$$\mathrm{E}[X_i] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_i\, dx_1 dx_2, \quad i = 1, 2. \quad (8.27)$$

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

$$\mathrm{E}\left[X_1^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^2 f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_1^2\, dx_1 dx_2, \quad (8.28)$$

$$\mathrm{E}\left[X_2^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2^2 f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_2^2\, dx_1 dx_2, \quad (8.29)$$

$$\mathrm{E}[X_1 X_2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{X}}(\mathbf{x})\, dx_1 dx_2 = \int_0^1 \int_0^{x_2} 2x_1 x_2\, dx_1 dx_2. \quad (8.30)$$

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

[Continued]

# Example 8.7 Solution (Continued 2)

Therefore,

$$\mathbf{R_X} = \begin{bmatrix} 1/6 & 1/4 \\ 1/4 & 1/2 \end{bmatrix}. \tag{8.31}$$

We use Theorem 8.7 to find the elements of the covariance matrix.

$$\mathbf{C_X} = \mathbf{R_X} - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}'_\mathbf{X} = \begin{bmatrix} 1/6 & 1/4 \\ 1/4 & 1/2 \end{bmatrix} - \begin{bmatrix} 1/9 & 2/9 \\ 2/9 & 4/9 \end{bmatrix} = \begin{bmatrix} 1/18 & 1/36 \\ 1/36 & 1/18 \end{bmatrix}. \tag{8.32}$$

# Definition 8.10 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* $\mathbf{R_{XY}}$ *with* $i, j$*th element* $R_{XY}(i,j) = \mathsf{E}[X_i Y_j]$, *or, in vector notation,*

$$\mathbf{R_{XY}} = \mathsf{E}\left[\mathbf{XY}'\right].$$

# Definition 8.11 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* $\mathbf{C_{XY}}$ *with* $i, j$*th element* $C_{\mathbf{XY}}(i,j) = \text{Cov}[X_i, Y_j]$, *or, in vector notation,*

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

# Theorem 8.8

$\mathbf{X}$ 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 8.8

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

$$\boldsymbol{\mu}_{\mathbf{Y}} = \mathsf{E}\left[\mathbf{A}\mathbf{X} + \mathbf{b}\right] = \mathbf{A}\,\mathsf{E}\left[\mathbf{X}\right] + \mathsf{E}\left[\mathbf{b}\right] = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}. \tag{8.33}$$

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

$$\begin{aligned} \mathbf{C}_{\mathbf{Y}} &= \mathsf{E}\left[(\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}))(\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}))'\right] \\ &= \mathsf{E}\left[\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}))(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'\mathbf{A}'\right] = \mathbf{A}\,\mathsf{E}\left[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'\right]\mathbf{A}' = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}'. \end{aligned} \tag{8.34}$$

# Example 8.8 Problem

Given the expected value $\boldsymbol{\mu}_{\mathbf{X}}$, the correlation $\mathbf{R}_{\mathbf{X}}$, and the covariance $\mathbf{C}_{\mathbf{X}}$ of random vector $\mathbf{X}$ in Example 8.7, and $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 6 & 3 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 0 \\ -2 \\ -2 \end{bmatrix}, \tag{8.35}$$

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

# Example 8.8 Solution

From the matrix operations of Theorem 8.8, we obtain $\boldsymbol{\mu}_{\mathbf{Y}} = \begin{bmatrix} 1/3 & 2 & 3 \end{bmatrix}'$ and

$$\mathbf{R}_{\mathbf{Y}} = \begin{bmatrix} 1/6 & 13/12 & 4/3 \\ 13/12 & 7.5 & 9.25 \\ 4/3 & 9.25 & 12.5 \end{bmatrix}; \qquad \mathbf{C}_{\mathbf{Y}} = \begin{bmatrix} 1/18 & 5/12 & 1/3 \\ 5/12 & 3.5 & 3.25 \\ 1/3 & 3.25 & 3.5 \end{bmatrix}. \qquad (8.36)$$

# Theorem 8.9

The vectors $\mathbf{X}$ and $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ have cross-correlation $\mathbf{R}_{\mathbf{XY}}$ and cross-covariance $\mathbf{C}_{\mathbf{XY}}$ given by

$$\mathbf{R}_{\mathbf{XY}} = \mathbf{R}_{\mathbf{X}}\mathbf{A}' + \boldsymbol{\mu}_{\mathbf{X}}\mathbf{b}', \qquad \mathbf{C}_{\mathbf{XY}} = \mathbf{C}_{\mathbf{X}}\mathbf{A}'.$$

# Example 8.9 Problem

Continuing Example 8.8 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 $\rho_{Y_1,Y_3}$ and $\rho_{X_2,Y_1}$.

# Example 8.9 Solution

(a) Direct matrix calculation using Theorem 8.9 yields

$$\mathbf{R_{XY}} = \begin{bmatrix} 1/6 & 13/12 & 4/3 \\ 1/4 & 5/3 & 29/12 \end{bmatrix}; \quad \mathbf{C_{XY}} = \begin{bmatrix} 1/18 & 5/12 & 1/3 \\ 1/36 & 1/3 & 5/12 \end{bmatrix}. \tag{8.37}$$

(b) Referring to Definition 5.6 and recognizing that $\text{Var}[Y_i] = C_{\mathbf{Y}}(i,i)$, we have

$$\rho_{Y_1,Y_3} = \frac{\text{Cov}[Y_1, Y_3]}{\sqrt{\text{Var}[Y_1]\,\text{Var}[Y_3]}} = \frac{C_Y(1,3)}{\sqrt{C_Y(1,1)C_Y(3,3)}} = 0.756 \tag{8.38}$$

Similarly,

$$\rho_{X_2,Y_1} = \frac{\text{Cov}[X_2, Y_1]}{\sqrt{\text{Var}[X_2]\,\text{Var}[Y_1]}} = \frac{C_{\mathbf{XY}}(2,1)}{\sqrt{C_{\mathbf{X}}(2,2)C_{\mathbf{Y}}(1,1)}} = 1/2. \tag{8.39}$$

# Quiz 8.4

The three-dimensional random vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$ has PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 6 & 0 \leq x_1 \leq x_2 \leq x_3 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8.40}$$

Find $E[\mathbf{X}]$ and the correlation and covariance matrices $\mathbf{R_X}$ and $\mathbf{C_X}$.

# Quiz 8.4 Solution

To solve this problem, we need to find the expected values $E[X_i]$ and $E[X_iX_j]$ for each $I$ and $j$. To do this, we need the marginal PDFs $f_{X_i}(x_i)$ and $f_{X_i,X_j}(x_i,x_j)$. First we note that each marginal PDF is nonzero only if any subset of the $x_i$ obeys the ordering contraints $0 \leq x_1 \leq x_2 \leq x_3 \leq 1$. Within these constraints, we have

$$f_{X_1,X_2}(x_1,x_2) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\ dx_3 = \int_{x_2}^{1} 6\,dx_3 = 6(1-x_2), \tag{1}$$

and

$$f_{X_2,X_3}(x_2,x_3) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\ dx_1 = \int_{0}^{x_2} 6\,dx_1 = 6x_2, \tag{2}$$

and

$$f_{X_1,X_3}(x_1,x_3) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\ dx_2 = \int_{x_1}^{x_3} 6\,dx_2 = 6(x_3-x_1). \tag{3}$$

In particular, we must keep in mind that $f_{X_1,X_2}(x_1,x_2) = 0$ unless $0 \leq x_1 \leq x_2 \leq 1$, $f_{X_2,X_3}(x_2,x_3) = 0$ unless $0 \leq x_2 \leq x_3 \leq 1$, and that $f_{X_1,X_3}(x_1,x_3) = 0$ unless $0 \leq x_1 \leq x_3 \leq 1$. The complete expressions are

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} 6(1-x_2) & 0 \leq x_1 \leq x_2 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

$$f_{X_2,X_3}(x_2,x_3) = \begin{cases} 6x_2 & 0 \leq x_2 \leq x_3 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

$$f_{X_1,X_3}(x_1,x_3) = \begin{cases} 6(x_3-x_1) & 0 \leq x_1 \leq x_3 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

[Continued]

Now we can find the marginal PDFs. When $0 \le x_i \le 1$ for each $x_i$,

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\ dx_2$$
$$= \int_{x_1}^{1} 6(1 - x_2)\, dx_2 = 3(1 - x_1)^2. \tag{7}$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_2,X_3}(x_2, x_3)\ dx_3$$
$$= \int_{x_2}^{1} 6x_2\, dx_3 = 6x_2(1 - x_2). \tag{8}$$

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty} f_{X_2,X_3}(x_2, x_3)\ dx_2$$
$$= \int_{0}^{x_3} 6x_2\, dx_2 = 3x_3^2. \tag{9}$$

[Continued]

# Quiz 8.4 Solution (Continued 3)

The complete expressions are

$$f_{X_1}(x_1) = \begin{cases} 3(1-x_1)^2 & 0 \le x_1 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

$$f_{X_2}(x_2) = \begin{cases} 6x_2(1-x_2) & 0 \le x_2 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

$$f_{X_3}(x_3) = \begin{cases} 3x_3^2 & 0 \le x_3 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Now we can find the components $\mathsf{E}[X_i] = \int_{-\infty}^{\infty} x f_{X_i}(x)\,dx$ of $\boldsymbol{\mu}_X$.

$$\mathsf{E}[X_1] = \int_0^1 3x(1-x)^2\,dx = 1/4, \tag{13}$$

$$\mathsf{E}[X_2] = \int_0^1 6x^2(1-x)\,dx = 1/2, \tag{14}$$

$$\mathsf{E}[X_3] = \int_0^1 3x^3\,dx = 3/4. \tag{15}$$

[Continued]

To find the correlation matrix $\mathbf{R}_X$, we need to find $\text{E}[X_i X_j]$ for all $i$ and $j$. We start with the second moments:

$$\text{E}\left[X_1^2\right] = \int_0^1 3x^2(1-x)^2\,dx = \frac{1}{10}. \tag{16}$$

$$\text{E}\left[X_2^2\right] = \int_0^1 6x^3(1-x)\,dx = \frac{3}{10}. \tag{17}$$

$$\text{E}\left[X_3^2\right] = \int_0^1 3x^4\,dx = \frac{3}{5}. \tag{18}$$

Using marginal PDFs, the cross terms are

$$\begin{aligned} \text{E}[X_1X_2] &= \iint x_1x_2 f_{X_1,X_2}(x_1,x_2)\,,\,dx_1\,dx_2 \\ &= \int_0^1 \left(\int_{x_1}^1 6x_1x_2(1-x_2)\,dx_2\right)\,dx_1 = \int_0^1 [x_1 - 3x_1^3 + 2x_1^4]\,dx_1 = \frac{3}{20}. \end{aligned} \tag{19}$$

$$\text{E}[X_2X_3] = \int_0^1 \int_{x_2}^1 6x_2^2x_3\,dx_3\,dx_2 = \int_0^1 [3x_2^2 - 3x_2^4]\,dx_2 = \frac{2}{5}.$$

[Continued]

# Quiz 8.4 Solution (Continued 5)

$$\mathrm{E}[X_1X_3] = \int_0^1 \int_{x_1}^1 6x_1x_3(x_3 - x_1)\,dx_3\,dx_1$$
$$= \int_0^1 \left( (2x_1x_3^3 - 3x_1^2x_3^2)\Big|_{x_3=x_1}^{x_3=1} \right) dx_1$$
$$= \int_0^1 [2x_1 - 3x_1^2 + x_1^4]\,dx_1 = 1/5. \tag{20}$$

Summarizing the results, $\mathbf{X}$ has correlation matrix

$$\mathbf{R}_X = \begin{bmatrix} 1/10 & 3/20 & 1/5 \\ 3/20 & 3/10 & 2/5 \\ 1/5 & 2/5 & 3/5 \end{bmatrix}. \tag{21}$$

Vector $\mathbf{X}$ has covariance matrix

$$\mathbf{C}_X = \mathbf{R}_X - \mathrm{E}[\mathbf{X}]\,\mathrm{E}[\mathbf{X}]'$$
$$= \begin{bmatrix} \frac{1}{10} & \frac{3}{20} & \frac{1}{5} \\ \frac{3}{20} & \frac{3}{10} & \frac{2}{5} \\ \frac{1}{5} & \frac{2}{5} & \frac{3}{5} \end{bmatrix} - \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{3}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix} = \frac{1}{80} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}. \tag{22}$$

This problem shows that even for fairly simple joint PDFs, computing the covariance matrix can be time consuming.

## Section 8.5

# Gaussian Random Vectors

# Definition 8.12 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*

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

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

# Theorem 8.10

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

# **Proof:** Theorem 8.10

First, if the components of $\mathbf{X}$ are independent, then for $i \neq j$, $X_i$ and $X_j$ are independent. By Theorem 5.17(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

$$\mathbf{C_X} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix} \quad \text{and} \quad \mathbf{C_X}^{-1} = \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_n^2 \end{bmatrix}. \tag{8.41}$$

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

$$(\mathbf{x} - \boldsymbol{\mu}_\mathbf{X})'\mathbf{C_X}^{-1}(\mathbf{x} - \boldsymbol{\mu}_\mathbf{X}) = \sum_{i=1}^{n} \frac{(X_i - \mu_i)^2}{\sigma_i^2}. \tag{8.42}$$

From Definition 8.12, we see that

$$f_\mathbf{X}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i^2} \exp\left(-\sum_{i=1}^{n} (x_i - \mu_i)/2\sigma_i^2\right) \tag{8.43}$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-(x_i - \mu_i)^2/2\sigma_i^2\right). \tag{8.44}$$

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

# Example 8.10 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

$$\mathbf{C_X} = \begin{bmatrix} 16.0 & 12.8 & 11.2 \\ 12.8 & 16.0 & 12.8 \\ 11.2 & 12.8 & 16.0 \end{bmatrix}. \qquad (8.45)$$

(a) Write the joint PDF of $X_1, X_2$ using the algebraic notation of Definition 5.10.

(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 8.10 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 5.6 that the correlation coefficient is

$$\rho_{X_1,X_2} = \frac{\text{Cov}\,[X_1, X_2]}{\sigma_1\sigma_2} = \frac{12.8}{16} = 0.8. \tag{8.46}$$

From Definition 5.10, 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 $2 \times 2$ submatrix in the upper left corner is the covariance matrix of the random vector $\mathbf{W}$. Thus

[Continued]

# Example 8.10 Solution (Continued 2)

$$\boldsymbol{\mu}_{\mathbf{W}} = \begin{bmatrix} 50 \\ 62 \end{bmatrix}, \qquad \mathbf{C}_{\mathbf{W}} = \begin{bmatrix} 16.0 & 12.8 \\ 12.8 & 16.0 \end{bmatrix}. \tag{8.47}$$

We observe that $\det(\mathbf{C}_{\mathbf{W}}) = 92.16$ and $\det(\mathbf{C}_{\mathbf{W}})^{1/2} = 9.6$. From Definition 8.12, the joint PDF of $\mathbf{W}$ is

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{1}{60.3} \exp\left(-\frac{1}{2}(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{W}})^T \mathbf{C}_{\mathbf{W}}^{-1} (\mathbf{w} - \boldsymbol{\mu}_{\mathbf{W}})\right). \tag{8.48}$$

(c) Since $\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} 50 & 62 & 58 \end{bmatrix}'$ and $\det(\mathbf{C}_{\mathbf{X}})^{1/2} = 22.717$, $\mathbf{X}$ has PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{357.8} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})\right). \tag{8.49}$$

# Theorem 8.11

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 8.11

The proof of Theorem 8.8 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 8.5.14. When $m = n$, we use Theorem 8.6 to write

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\Big(\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b})\Big) \tag{8.50}$$

$$= \frac{\exp\left(-\frac{1}{2}[\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b}) - \boldsymbol{\mu}_{\mathbf{X}}]' \mathbf{C}_{\mathbf{X}}^{-1} [\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b}) - \boldsymbol{\mu}_{\mathbf{X}}]\right)}{(2\pi)^{n/2} |\det(\mathbf{A})| \, |\det(\mathbf{C}_{\mathbf{X}})|^{1/2}}. \tag{8.51}$$

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

$$\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b}) - \boldsymbol{\mu}_{\mathbf{X}} = \mathbf{A}^{-1}[\mathbf{y} - (\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b})] = \mathbf{A}^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{Y}}), \tag{8.52}$$

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

[Continued]

# **Proof:** Theorem 8.11 (Continued 2)

Applying (8.52) to (8.51) yields

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}[\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})]'\mathbf{C}_{\mathbf{X}}^{-1}[\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})]\right)}{(2\pi)^{n/2}\,|\det(\mathbf{A})|\,|\det(\mathbf{C}_{\mathbf{X}})|^{1/2}}. \tag{8.53}$$

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

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})'(\mathbf{A}')^{-1}\mathbf{C}_{\mathbf{X}}^{-1}\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{Y}})\right)}{(2\pi)^{n/2}\,|\det(\mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}')|^{1/2}}. \tag{8.54}$$

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 (8.54) 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 8.11 Problem

Continuing Example 8.10, 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 8.11 Solution

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

$$\mathbf{A} = \begin{bmatrix} 5/9 & 0 & 0 \\ 0 & 5/9 & 0 \\ 0 & 0 & 5/9 \end{bmatrix}, \qquad \mathbf{b} = -\frac{160}{9}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{8.55}$$

(b) Since $\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} 50 & 62 & 58 \end{bmatrix}'$, from Theorem 8.11,

$$\boldsymbol{\mu}_{\mathbf{Y}} = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b} = \begin{bmatrix} 10 \\ 50/3 \\ 130/9 \end{bmatrix}. \tag{8.56}$$

(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 $3 \times 3$ 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

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{24.47}\exp\left(-\frac{81}{50}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{Y}})^T\mathbf{C}_{\mathbf{X}}^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{Y}})\right). \tag{8.57}$$

## **Definition 8.13** Standard Normal Random Vector

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

# Theorem 8.12

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 8.12

Applying Theorem 8.11 with $\mathbf{A}$ replaced by $\mathbf{A}^{-1}$, and $\mathbf{b} = \mathbf{A}^{-1}\boldsymbol{\mu}_{\mathbf{X}}$, we have that $\mathbf{Z}$ is a Gaussian random vector with expected value

$$\mathsf{E}[\mathbf{Z}] = \mathsf{E}\left[\mathbf{A}^{-1}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})\right] = \mathbf{A}^{-1}\,\mathsf{E}[\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}] = \mathbf{0} \tag{8.58}$$

and covariance

$$\mathbf{C}_{\mathbf{Z}} = \mathbf{A}^{-1}\mathbf{C}_{\mathbf{X}}(\mathbf{A}^{-1})' = \mathbf{A}^{-1}\mathbf{A}\mathbf{A}'(\mathbf{A}')^{-1} = \mathbf{I}. \tag{8.59}$$

# Theorem 8.13

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 8.13

By Theorem 8.11, $\mathbf{X}$ is a Gaussian random vector with expected value

$$\mu_{\mathbf{X}} = \mathsf{E}[\mathbf{X}] = \mathsf{E}[\mathbf{AZ} + \mu_{\mathbf{X}}] = \mathbf{A}\,\mathsf{E}[\mathbf{Z}] + \mathbf{b} = \mathbf{b}. \tag{8.60}$$

The covariance of $\mathbf{X}$ is

$$\mathbf{C_X} = \mathbf{AC_ZA}' = \mathbf{AIA}' = \mathbf{AA}'. \tag{8.61}$$

# Theorem 8.14

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

# **Proof:** Theorem 8.14

To verify this fact, we connect some simple facts:

- In Problem 8.4.12, we ask you 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

$$\mathbf{C_X} = \mathbf{UD}^{1/2}\mathbf{D}^{1/2}\mathbf{U}' = \left(\mathbf{UD}^{1/2}\right)\left(\mathbf{UD}^{1/2}\right)'. \tag{8.62}$$

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

# Quiz 8.5

$\mathbf{Z}$ is the two-dimensional standard normal random vector. The Gaussian random vector $\mathbf{X}$ has components

$$X_1 = 2Z_1 + Z_2 + 2 \quad \text{and} \quad X_2 = Z_1 - Z_2. \tag{8.65}$$

Calculate the expected value vector $\boldsymbol{\mu}_{\mathbf{X}}$ and the covariance matrix $\mathbf{C}_{\mathbf{X}}$.

# Quiz 8.5 Solution

We observe that $\mathbf{X} = \mathbf{A}\mathbf{Z} + \mathbf{b}$ where

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{1}$$

It follows from Theorem 8.13 that $\boldsymbol{\mu}_X = \mathbf{b}$ and that

$$\mathbf{C}_X = \mathbf{A}\mathbf{A}' = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 1 & 2 \end{bmatrix}.$$

# Section 8.6

Matlab

# Example 8.12 Problem

Finite random vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2, \cdots & X_5 \end{bmatrix}'$ has PMF

$$P_{\mathbf{X}}(\mathbf{x}) = \begin{cases} k\sqrt{\mathbf{x}'\mathbf{x}} & x_i \in \{-10, -9, \ldots, 10\}\,; \\ & i = 1, 2, \ldots, 5, \\ 0 & \text{otherwise.} \end{cases} \tag{8.66}$$

What is the constant $k$? Find the expected value and standard deviation of $X_3$.

# Example 8.12 Solution

Summing $P_{\mathbf{X}}(\mathbf{x})$ over all possible values of $\mathbf{x}$ is the sort of tedious task that Matlab handles easily. Here are the code and corresponding output:

```
%x5.m
sx=-10:10;
[SX1,SX2,SX3,SX4,SX5]...
    =ndgrid(sx,sx,sx,sx,sx);
P=sqrt(SX1.^2 +SX2.^2+SX3.^2+SX4.^2+SX5.^2);
k=1.0/(sum(sum(sum(sum(sum(P))))))
P=k*P;
EX3=sum(sum(sum(sum(sum(P.*SX3)))))
EX32=sum(sum(sum(sum(sum(P.*(SX3.^2))))));
sigma3=sqrt(EX32-(EX3)^2)
```

```
>> x5
k =
  1.8491e-008
EX3 =
 -3.2960e-017
sigma3 =
    6.3047
>>
```

In fact, by symmetry arguments, it should be clear that $\mathrm{E}[X_3] = 0$. In adding $11^5$ terms, Matlab's finite precision led to a small error on the order of $10^{-17}$.

# Example 8.13 Problem

Write a Matlab function `f=gaussvectorpdf(mu,C,x)` that calculates $f_{\mathbf{X}}(\mathbf{x})$ for a Gaussian $(\boldsymbol{\mu}, \mathbf{C})$ random vector.

# Example 8.13 Solution

```
function f=gaussvectorpdf(mu,C,x)
n=length(x);
z=x(:)-mu(:);
f=exp(-z'*inv(C)*z)/...
    sqrt((2*pi)^n*det(C));
```

`gaussvectorpdf` computes the Gaussian vector PDF $f_{\mathbf{X}}(\mathbf{x})$ of Definition 8.12. Of course, Matlab makes the calculation simple by providing operators for matrix inverses and determinants.

# Quiz 8.6

The daily noon temperature, measured in degrees Fahrenheit, in New Jersey in July can be modeled as a Gaussian random vector $\mathbf{T} = \begin{bmatrix} T_1 & \cdots & T_{31} \end{bmatrix}'$ where $T_i$ is the temperature on the $i$th day of the month. Suppose that $E[T_i] = 80$ for all $i$, and that $T_i$ and $T_j$ have covariance

$$\text{Cov}\left[T_i, T_j\right] = \frac{36}{1 + |i - j|} \tag{8.67}$$

Define the daily average temperature as

$$Y = \frac{T_1 + T_2 + \cdots + T_{31}}{31}. \tag{8.68}$$

Based on this model, write a Matlab program `p=julytemps(T)` that calculates $P[Y \geq T]$, the probability that the daily average temperature is at least $T$ degrees.

# Quiz 8.6 Solution

First, we observe that $Y = \mathbf{AT}$ where $\mathbf{A} = \begin{bmatrix} 1/31 & 1/31 & \cdots & 1/31 \end{bmatrix}'$. Since $\mathbf{T}$ is a Gaussian random vector, Theorem 8.11 tells us that $Y$ is a 1 dimensional Gaussian vector, i.e., just a Gaussian random variable. The expected value of $Y$ is $\mu_Y = \mu_T = 80$. The covariance matrix of $Y$ is $1 \times 1$ and is just equal to $\text{Var}[Y]$. Thus, by Theorem 8.11, $\text{Var}[Y] = \mathbf{AC}_T\mathbf{A}'$.

In `julytemps.m` shown below, the first two lines generate the $31 \times 31$ covariance matrix `CT`, or $\mathbf{C}_T$. Next we calculate $\text{Var}[Y]$. The final step is to use the $\Phi(\cdot)$ function to calculate $\text{P}[Y < T]$.

```
function p=julytemps(T);
[D1 D2]=ndgrid((1:31),(1:31));
CT=36./(1+abs(D1-D2));
A=ones(31,1)/31.0;
CY=(A')*CT*A;
p=phi((T-80)/sqrt(CY));
```

[Continued]

Here is the output of `julytemps.m`:

```
>> julytemps([70 75 80 85 90])
ans =
 0.0000 0.0221 0.5000 0.9779 1.0000
```

Note that P$[T \leq 70]$ is not actually zero and that P$[T \leq 90]$ is not actually 1.0000. Its just that the Matlab's short format output, invoked with the command `format short`, rounds off those probabilities. The long format output resembles:

```
>> format long
>> julytemps([70 75])
ans =
  0.000028442631  0.022073830676
>> julytemps([85 90])
ans =
  0.977926169323  0.999971557368
```

The `ndgrid` function is a useful to way calculate many covariance matrices. However, in this problem, $\mathbf{C}_X$ has a special structure; the $i,j$th element is

[Continued]

$$C_{\mathbf{T}}(i,j) = c_{|i-j|} = \frac{36}{1+|i-j|}. \tag{1}$$

If we write out the elements of the covariance matrix, we see that

$$\mathbf{C}_{\mathbf{T}} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{30} \\ c_1 & c_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1 \\ c_{30} & \cdots & c_1 & c_0 \end{bmatrix}. \tag{2}$$

This covariance matrix is known as a symmetric Toeplitz matrix. Because Toeplitz covariance matrices are quite common, Matlab has a `toeplitz` function for generating them. The function `julytemps2` use the `toeplitz` to generate the correlation matrix $\mathbf{C}_{\mathbf{T}}$.

```
function p=julytemps2(T);
c=36./(1+abs(0:30));
CT=toeplitz(c);
A=ones(31,1)/31.0;
CY=(A')*CT*A;
p=phi((T-80)/sqrt(CY));
```