Systems and Signals 344

2025

Lecture 21

Random Vectors: Notation, Independence, Functions of Random Vectors

Sections 8.1-8.3

Recap: Multivariate Probability Models

Definition 5.11 Multivariate Joint CDF

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

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

Definition 5.12 Multivariate Joint PMF

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

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

Definition 5.13 Multivariate Joint PDF

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

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

Recap: Multivariate Probability Models

Theorem 5.22

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

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

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

Theorem 5.23

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

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

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

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

To calculate **marginals**

Theorem 5.25

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

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

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

Theorem 5.26

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

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

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

Recap: Multivariate Probability Models

Definition 5.14 N Independent Random Variables

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

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

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

Random Vectors

Previously we looked at the CDF and PDF of n random variables $X_1, \ldots, X_n$, now we will focus on random vectors

$$\mathbf{X} = [X_1 \quad \ldots \quad X_n]'.$$

A random vector treats a collection of n random variables as a single entity. Thus, vector notation provides a concise representation of relationships that would otherwise be extremely difficult to represent.

NOTE: When writing a vector on paper it is typically more convenient to use $\overline{\mathrm{X}}$ instead of $\mathbf{X}$

Vector Notation

- We will use boldface notation $\mathbf{x}$ for a column vector.
- Row vectors are *transposed* column vectors; $\mathbf{x}'$ is a row vector.
- The components of a column vector are, by definition, written in a column. However, to save space, we will often use the transpose of a row vector to display a column vector: $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}'$ is a column vector.
- 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.

Vector Notation

A **random vector** is a column vector $\mathbf{X} = [X_1 \quad \ldots \quad X_n]'$.
Each X_i is a random variable.

A **sample value of a random vector** is a column vector
$$\mathbf{x} = [x_1 \quad \ldots \quad x_n]'.$$
The i^{th} component, x_i, of the vector $\mathbf{x}$ is a sample value of a random variable, X_i.

We use similar notation for a function $g(\mathbf{X}) = g(X_1, \ldots, X_n)$ of n random variables and a function $g(\mathbf{x}) = g(x_1, \ldots, x_n)$ of n numbers.

Vector Notation

The **CDF** of a random vector $\mathbf{X}$ is

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

The **PMF** of a *discrete* random vector $\mathbf{X}$ is

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

The **PDF** of a *continuous* random vector $\mathbf{X}$ is

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

Vector Notation

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

The **joint CDF** of random vectors $\mathbf{X}$ and $\mathbf{Y}$ is

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

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

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

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

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

Vector Notation

Note that we call the pair of random variables $\mathbf{X}$ and $\mathbf{Y}$ simply $\mathbf{W} = [\mathbf{X}' \ \ \mathbf{Y}']' = [X_1 \ \ \ldots \ \ X_n \ \ Y_1 \ \ \ldots \ \ Y_m]'$ which is a concatenation of $\mathbf{X}$ and $\mathbf{Y}$ to form a bigger column vector.

This can simplify the notation, for example: $F_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})$ is the then the same as $F_{\mathbf{W}}(\mathbf{w})$.

Furthermore, one can calculate the **marginal probability models**, for example $f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})$ can marginalise to $f_{\mathbf{X}}(\mathbf{x})$ and $f_{\mathbf{Y}}(\mathbf{y})$.

Example 8.1

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 $\mathbf{X}$?

Independent RVs and Random Vectors

The probability model of the pair of independent random vectors $\mathbf{X}$ and $\mathbf{Y}$ is the product of the probability model of $\mathbf{X}$ and the probability model of $\mathbf{Y}$.

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 8.2

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

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

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?

Functions of Random Vectors

For random variable (not random vector) $W = g(\mathbf{X})$,

Discrete

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

Continuous

$$F_W(w) = P(W \leq w) = \int \ldots \int_{g(\mathbf{x}) \leq w} f_{\mathbf{X}}(\mathbf{x})\, dx_1 \ldots dx_n$$

Example 8.3

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.

Functions of Random Vectors

Let $\mathbf{X}$ be a vector of n **independent and identically distributed (iid)** *continuous* random variables, each with CDF $F_X(x)$ and PDF $f_X(x)$

- The CDF and PDF of $Y = \max\{X_1, \ldots, X_n\}$ are
  - $F_Y(y) = (F_X(y))^n$
  - $f_Y(y) = n(F_X(y))^{n-1} f_X(y)$
- The CDF and the PDF of $W = \min\{X_1, \ldots, X_n\}$ are
  - $F_W(w) = 1 - (1 - F_X(y))^n$
  - $f_W(w) = n(1 - F_X(w))^{n-1} f_X(w)$

Functions of Random Vectors

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

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

Functions of Random Vectors

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

$$\mathrm{E}[g_1(X_1)g_2(X_2)\cdots g_n(X_n)] = \mathrm{E}[g_1(X_1)]\mathrm{E}[g_2(X_2)]\cdots\mathrm{E}[g_n(X_m)]$$

Functions of Random Vectors

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

$$F_{\mathbf{Y}}(y) = F_{\mathbf{X}}\left(\frac{y_1 - b}{a}, \ldots, \frac{y_n - b}{a}\right)$$

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

This is a special case of a transformation of the form $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ which is a consequence of the change-of-variable theorem in multivariable calculus.

Functions of Random Vectors

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

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

Note that this is similar to the previous slide, but now there are multiple values for $\boldsymbol{a}$ and $\boldsymbol{b}$, one for each random variable, so that they can be "weighed" in different ratios.