Lecture-05: Random Vectors

1 Random vectors

Definition 1.1 (Projection). For a vector $x \in \mathbb{R}^n$, we can define $\pi_i : \mathbb{R}^n \to \mathbb{R}$ is the **projection** of an n-length vector onto its i-th component, such that $\pi_i(x) = x_i$.

Remark 1. For a subset $A \subseteq \mathbb{R}$ and projection $\pi_i : \mathbb{R}^n \to \mathbb{R}$, we can write

$$\pi_i^{-1}(A) \triangleq \{x \in \mathbb{R}^n : x_i \in A\} = \mathbb{R} \times \dots A \cdots \times \mathbb{R}.$$

Example 1.2. Consider a function $f : \Omega \to \mathbb{R}^n$, then $f(\omega) \in \mathbb{R}^n$ and can be expressed in terms of its components $(f_1(\omega), \dots, f_n(\omega))$, where $f_i = \pi_i \circ f$. For this function f, we can write the inverse image of a set $B \subseteq \mathbb{R}^n$ as

$$f^{-1}(B) \triangleq \cap_{i=1}^n \{\omega \in \Omega : (\pi_i \circ f)(\omega) \in \pi_i(B)\} = \cap_{i=1}^n f_i^{-1}(\pi_i(B)).$$

Definition 1.3 (Random vectors). Consider a probability space $(\Omega, \mathcal{F}, P)$ and a finite $n \in \mathbb{N}$. A **random vector** $X : \Omega \to \mathbb{R}^n$ is a mapping from sample space to an n-length real-valued vector, such that for $x \in \mathbb{R}^n$, the event

$$A(x) \triangleq \{\omega \in \Omega : X_1(\omega) \leqslant x_1, \dots, X_n(\omega) \leqslant x_n\} = \cap_{i=1}^n X_i^{-1}(-\infty, x_i] \in \mathcal{F}.$$

We say that the random vector X is $\mathcal{F}$-measurable and the probability of this event is denoted by

$$F_{X_1,\dots,X_n}(x_1, \dots, x_n) \equiv F_X(x) \triangleq P(A(x)) = P(\{X_1 \leqslant x_1, \dots, X_n \leqslant x_n\}) = P(\cap_{i=1}^n X_i^{-1}(-\infty, x_i]).$$

The function $F_X : \mathbb{R}^n \to [0,1]$ is called the **joint distribution function** of a random vector X. The **event space generated by the random vector** X is the smallest σ-algebra generated by the collection of events $(A(x) : x \in \mathbb{R}^n)$ and denoted by $\sigma(X) \triangleq \sigma(A(x) : x \in \mathbb{R}^n)$.

Theorem 1.4. *Consider a probability space $(\Omega, \mathcal{F}, P)$. For a finite $n \in \mathbb{N}$, $X : \Omega \to \mathbb{R}^n$ is a random vector if and only if $X_i \triangleq \pi_i \circ X : \Omega \to \mathbb{R}$ random variables for all $i \in [n]$. In particular, $\sigma(X) = \sigma(X_1, \dots, X_n)$.*

Proof. We will first show that $X : \Omega \to \mathbb{R}^n$ implies that $\pi_i \circ X$ is a random variable for any $i \in [n]$. For any $i \in [n]$ and $x_i \in \mathbb{R}$, we take $x = (\infty, \dots, x_i, \dots, \infty)$. This implies that $B \triangleq \pi_i^{-1}(-\infty, x_i] = \mathbb{R} \times \dots (-\infty, x_i] \cdots \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^n)$. Further, defining $A_i(x_i) \triangleq X_i^{-1}(-\infty, x_i]$, we observe from the definition of random vectors that

$$A_i(x_i) = X^{-1} \circ \pi_i^{-1}(-\infty, x_i] = X^{-1}(B) = A(x) \in \mathcal{F}. \tag{1}$$

We will next show that if $X_i : \Omega \to \mathbb{R}$ is a random variable for all $i \in [n]$, then $X \triangleq (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ is a random vector. For any $x \in \mathbb{R}^n$, we have $A_i(x_i) = X_i^{-1}(-\infty, x_i] \in \mathcal{F}$ for all $i \in [n]$, from the definition of random variables. From the closure of event set under countable intersections, we have

$$A(x) = \cap_{i=1}^n A_i(x_i) \in \mathcal{F}. \tag{2}$$

□

Definition 1.5. For a random vector $X : \Omega \to \mathbb{R}^n$ defined on the probability space $(\Omega, \mathcal{F}, P)$, the distribution of the ith random variable $X_i \triangleq \pi_i \circ X : \Omega \to \mathbb{R}$ is called the ith **marginal distribution**, and denoted by $F_{X_i} : \Omega \to [0,1]$ for all $i \in [n]$.

Corollary 1.6 (Marginal distribution). *Consider a random vector $X : \Omega \to \mathbb{R}^n$ defined on a probability space $(\Omega, \mathcal{F}, P)$ with the joint distribution $F_X : \mathbb{R}^n \to [0,1]$. The i-th **marginal distribution** and can be obtained from the joint distribution of X as*

$$F_{X_i}(x_i) = \lim_{x_j \to \infty, \text{ for all } j \neq i} F_X(x).$$

Proof. For any $i \in [n]$ and $x_i \in \mathbb{R}$, we have $X_i^{-1}(-\infty, x_i] = A(x)$ for $x = (\infty, \dots, x_i, \dots, \infty)$ from (1). □

1.1 Independence of random variables

Definition 1.7 (Independent and identically distributed). A random vector $X : \Omega \to \mathbb{R}^n$ defined on the probability space $(\Omega, \mathcal{F}, P)$ is called **independent** if

$$F_X(x) = \prod_{i=1}^{n} F_{X_i}(x_i), \text{ for all } x \in \mathbb{R}^n.$$

The random vector X is called **identically distributed** if each of its components have the identical marginal distribution, i.e.

$$F_{X_i} = F_{X_1}, \text{ for all } i \in [n].$$

Remark 2. Independence of a random vector implies that events $(A_i(x_i) : i \in [n])$ are independent for any $x \in \mathbb{R}^n$.

Definition 1.8. A family of collections of events $(\mathcal{A}_i \subseteq \mathcal{F} : i \in I)$ is called independent, if for any finite set $F \subseteq I$ and $A_i \in \mathcal{A}_i$ for all $i \in F$, we have

$$P(\cap_{i \in F} A_i) = \prod_{i \in F} P(A_i).$$

Remark 3. In general, if two collection of events are mutually independent, then the event space generated by them are independent. This can be proved using Dynkin's π-λ Theorem.

Theorem 1.9. *For an independent random vector $X : \Omega \to \mathbb{R}^n$ defined on a probability space $(\Omega, \mathcal{F}, P)$, the event spaces generated by its components $(\sigma(X_i) : i \in [n])$ are independent.*

Proof. For an we define a family of events $\mathcal{A}_i \triangleq (X_i^{-1}(-\infty, x] : x \in \mathbb{R})$ for each $i \in [n]$. From the definition of independence of random vectors, the families $(\mathcal{A}_i \subseteq \mathcal{F} : i \in [n])$ are mutually independent. Since $\sigma(\mathcal{A}_i) = \sigma(X_i)$, the result follows from the previous remark. □

Definition 1.10 (Independent random vectors). To random vectors $X, Y : \Omega \to \mathbb{R}^n$ defined on the same probability space $(\Omega, \mathcal{F}, P)$ are independent, if the collection of events $(A_X(x) : x \in \mathbb{R}^n)$ and $(A_Y(y) : y \in \mathbb{R}^n)$ are independent, where $A_X(x) \triangleq \cap_{i=1}^n X_i^{-1}(-\infty, x_i]$ and $A_Y(y) \triangleq \cap_{i=1}^n Y_i^{-1}(-\infty, y_i]$.

Example 1.11 (Independent random vectors). Consider a set of vectors $\mathcal{X} = \{(0,0,1), (1,0,0)\} \subseteq \mathbb{R}^3$. Consider two independent coin tosses, such that $\Omega = \{H, T\}^2$, $\mathcal{F} = 2^\Omega$ and $P(\omega) = p^{k_2(\omega)}(1-p)^{2-k_2(\omega)}$, where $k_2(\omega) = \sum_{i=1}^2 \mathbb{1}_{\{\omega_i = H\}}$. We define random vectors

$$X = (0,0,1)\mathbb{1}_{\{\omega_1 = H\}} + (1,0,0)\mathbb{1}_{\{\omega_1 = T\}}, \qquad Y = (0,0,1)\mathbb{1}_{\{\omega_2 = H\}} + (1,0,0)\mathbb{1}_{\{\omega_2 = T\}}.$$

We can verify that $X, Y : \Omega \to \mathbb{R}^3$ are mutually independent random vectors, though we can also check that X_1, X_3 are dependent random variables and so are Y_1, Y_3.

1.2 Discrete random vectors

Definition 1.12 (Discrete random vectors). If a random vector $X : \Omega \to \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \subseteq \mathbb{R}^n$ takes countable values in $\mathbb{R}^n$, then it is called a **discrete random vector**. That is, the range of random vector X is countable, and the random vector is completely specified by the **probability mass function**

$$P_X(x) = P(\cap_{i=1}^n \{X_i = x_i\}) \text{ for all } x \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n.$$

Remark 4. For an independent discrete random vector $X : \Omega \to \mathbb{R}^n$, we have $P_X(x) = \prod_{i=1}^n P_{X_i}(x_i)$ for each $x \in \mathbb{R}^n$.

Example 1.13 (Multiple coin tosses). For a probability space $(\Omega, \mathcal{F}, P)$, such that $\Omega = \{H, T\}^n, \mathcal{F} = 2^\Omega, P(\omega) = \frac{1}{2^n}$ for all $\omega \in \Omega$.

Consider the random vector $X : \Omega \to \mathbb{R}$ such that $X_i(\omega) = \mathbb{1}_{\{\omega_i = H\}}$ for each $i \in [n]$. Observe that X is a bijection from the sample space to the set $\{0, 1\}^n$. In particular, X is a discrete random variable.

For any $x \in [0, 1]^n$, we can write $N(x) = \sum_{i=1}^n \mathbb{1}_{[0,1)}(x_i)$. Further, we can write the joint distribution as

$$F_X(x) = \begin{cases} 1, & x_i \geqslant 1 \text{ for all } i \in [n], \\ \frac{1}{2^{N(x)}}, & x_i \in [0, 1] \text{ for all } i \in [n], \\ 0, & x_i < 0 \text{ for some } i \in [n]. \end{cases}$$

We can derive the marginal distribution for i-th component as

$$F_{X_i}(x_i) = \begin{cases} 1, & x_i \geqslant 1, \\ \frac{1}{2}, & x_i \in [0, 1), \\ 0, & x_i < 0. \end{cases}$$

Therefore, it follows that X is an **i.i.d.** vector.

1.3 Continuous random vectors

Definition 1.14 (Joint density function). For jointly continuous random vector $X : \Omega \to \mathbb{R}^n$ with joint distribution function $F_X : \mathbb{R}^n \to [0, 1]$, there exists a **joint density function** $f_X : \mathbb{R}^n \to [0, \infty)$ such that $f_X(x) = \frac{d^n}{dx_1 \dots dx_n} F_X(x)$, and

$$F_X(x) = \int_{u_1 \leqslant x_1} du_1 \cdots \int_{u_n \leqslant x_n} du_n f_X(u_1, \dots, u_n).$$

Remark 5. For an independent continuous random vector $X : \Omega \to \mathbb{R}^n$, we have $f_X(x) = \prod_{i=1}^n f_{X_i}(x_i)$ for all $x \in \mathbb{R}^n$.

Example 1.15 (Gaussian random vectors). For a probability space $(\Omega, \mathcal{F}, P)$, Gaussian random vector is a continuous random vector $X : \Omega \to \mathbb{R}^n$ defined by its density function

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right) \text{ for all } x \in \mathbb{R}^n,$$

where the mean vector $\mu \in \mathbb{R}^n$ and the positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. The components of the Gaussian random vector are Gaussian random variables, which are independent when Σ is diagonal matrix and are identically distributed when $\Sigma = \sigma^2 I$.

1.4 Properties of the joint distribution function

Lemma 1.16 (Properties of the joint distribution function). *Then the joint distribution function* $F_X : \mathbb{R}^n \to [0,1]$ *satisfies the following properties.*

(i) For $x, y \in \mathbb{R}^n$ *such that* $x_i \leqslant y_i$ *for each* $i \in [n]$*, we have* $F_X(x) \leqslant F_X(y)$*.*

(ii) The function $F_X(x)$ *is right continuous at all points* $x \in \mathbb{R}^n$*.*

(iii) The lower limit is $\lim_{x_i \to -\infty} F_X(x) = 0$*, and the upper limit is* $\lim_{x_i \to \infty, i \in [n]} F_X(x) = 1$*.*

Proof. Consider a random vector $X : \Omega \to \mathbb{R}^n$ defined on the probability space $(\Omega, \mathcal{F}, P)$ and any $x \in \mathbb{R}^n$.

(i) We can verify that $A(x) = \cap_{i=1}^n A_i(x_i) \subseteq \cap_{i=1}^n A_i(y_i) = A(y)$. The result follows from the monotonicity of probability measure.

(ii) The proof is similar to the proof for single random variable.

(iii) The event $A(x) = \emptyset$ when $x_i = -\infty$ for some $i \in [n]$ and $A(x) = \Omega$ when $x_i = \infty$ for all $i \in [n]$, hence the result follow.

□

Example 1.17 (Probability of rectangular events). Consider a probability space $(\Omega, \mathcal{F}, P)$ and a random vector $X : \Omega \to \mathbb{R}^2$. Let events $B_1 \triangleq \{x_1 < X_1 \leqslant y_1\} = A_1(y_1) \setminus A_1(x_1) \in \mathcal{F}$ and $B_2 \triangleq \{x_2 < X_2 \leqslant y_2\} = A_2(y_2) \setminus A_2(x_2) \in \mathcal{F}$. The marginal probabilities are given by

$$P(B_1) = P(A_1(y_1)) - P(A_1(x_1)) = F_{X_1}(y_1) - F_{X_1}(x_1), \quad P(B_2) = P(A_2(y_2)) - P(A_2(x_2)) = F_{X_2}(y_2) - F_{X_2}(x_2).$$

Then the probability of the rectangular event $B_1 \cap B_2 = (A(y_1, y_2) \setminus A(x_1, y_2)) \setminus (A(y_1, x_2) \setminus A(x_1, x_2)) \in \mathcal{F}$ is

$$P(B_1 \cap B_2) = (F_X(y_1, y_2) - F_X(x_1, y_2)) - (F_X(y_1, x_2) - F_X(x_1, x_2)).$$