Chapter 1. Multivariate Random Variables

Note. We address multivariate random variables in terms of vectors of individual random variables. The case of two random variables is covered in Foundations of Probability and Statistics-Calculus Based (MATH 2050) in Section 2.6. Jointly Distributed Random Variables. These ideas are also covered in Mathematical Statistics 1 (STAT 4047/5047) in Chapter 2, "Multivariate Distributions," where emphasis is on two random variables, but also several random variables are considered in Section 2.6. Extension to Several Random Variables. ETSU's class Applied Multivariate Statistical Analysis (STAT 5730) is devoted to applications of this material; see my online notes (in preparation) for Applied Multivariate Statistical Analysis.

1.1. Introduction

Note. In this section, we define a random vector, concentrating on the case of two-dimensional random vectors. We define joint distribution functions, joint probability functions, independent random variables, variance, and covariance.

Definition. An n-dimensional *random variable* or *random vector* $\mathbf{X}$ is a (measurable) function from the probability space Ω to $\mathbb{R}^n$, $\mathbf{X} : \omega \to \mathbb{R}^n$.

Note. When computing, we treat a random vector as a column vector: $\mathbf{X} = (X_1, x_2, \ldots, X_n)'$ where the prime indicates the transpose of the row vector. Throughout, we assume that all of the components of a random vector are of the same kind, either all discrete or all continuous.

Definition. For random variable $\mathbf{X}$, the *joint distribution function* is

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n),$$

or equivalently $F_{\mathbf{X}}(\mathbf{x}) = P(\mathbf{X} \leq \mathbf{x})$ where $\mathbf{x} \in \mathbb{R}^n$, where the event $\{\mathbf{X} \leq \mathbf{x}\}$ is

$$\{\mathbf{X} \leq \mathbf{x}\} = \{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\} = \cap_{k=1}^{n}\{_k \leq x_k\}.$$

For discrete random variable $\mathbf{X}$, the *joint probability function* is $p_{\mathbf{x}}(\mathbf{x}) = P(\mathbf{X} = \mathbf{x})$ where $\mathbf{x} \in \mathbb{R}^n$, or equivalently

$$p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P(X_1 = x_1, X_2 = x_1, \ldots, X_n = x_n)$$

$$\text{for } x_k \in \mathbb{R}, \text{ where } k = 1,2,\ldots,n.$$

Note. In the discrete case the joint distribution function is $F_{\mathbf{X}}(\mathbf{x}) = \sum_{\mathbf{y} \leq \mathbf{x}} p_{\mathbf{X}}(\mathbf{y})$ or

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = \sum_{y_1 \leq x_1} \sum_{y_2 \leq x_2} \cdots \sum_{y_n \leq x_n} p_{X_1,X_2,\ldots,X_n}(y_1,y_2,\ldots,y_n).$$

Definition. For (absolutely) continuous random variable $\mathbf{X}$, the *joint density function* is $f_{\mathbf{X}}(\mathbf{x}) = \dfrac{d^n F_{\mathbf{X}}(\mathbf{x})}{d\mathbf{x}^n}$ where $\mathbf{x} \in \mathbb{R}^n$, or

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots x_n) = \frac{\partial^n F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n},$$

where $x_k \in \mathbb{R}$ for $k = 1,2,\ldots,n$.

Note. The reasons for requiring $\mathbf{X}$ to be absolutely continuous is explained in Real Analysis 1 (MATH 5210). In my online notes for Real Analysis 1, an absolutely continuous function (of a single variable) is defined in Section 6.4. Absolutely Continuous Functions and it is shown in Section 6.5. Integrating Derivatives: Differentiating Indefinite Integrals that a Fundamental Theorem of Lebesgue Calculus holds for absolutely continuous functions (see Theorem 6.10 and Theorem 6.14).

Definition. Let discrete random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ have joint probability function $p_{\mathbf{X}}(\mathbf{x})$. The *marginal probability function* $p_{X_i}(x_i)$ is

$$p_{X_i}(x_i) = \sum_{X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n} p_{\mathbf{X}}(\mathbf{x}),$$

where i is some fixed index between 1 and n. Let continuous random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ have joint density function $f_{\mathbf{X}}(\mathbf{x})$. The *marginal density function* $f_{X_i}(x_i)$ is

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\,dx_1\,dx_2\cdots dx_{i-1}\,dx_{i+1}\cdots dx_n,$$

where i is some fixed index between 1 and n.

Example 1.1. Let $\mathbf{X} = (X, Y)$ be a two-dimensional continuous random vector with joint distribution

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

We find the marginal density function $f_X(x)$. We have

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy = \frac{2}{\pi}\sqrt{1-x^2}$$

where $-1 < x < 1$, and $f_X(x) = 0$ for $|x| \geq 1$.

Definition. The components of a random vector $\mathbf{X}$ are *independent* if the joint distribution satisfies

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{k=1}^{n} F_{X_k}(x_k) \text{ where } x_k \in \mathbb{R} \text{ for } k = 1, 2, \ldots, n.$$

Note 1.1.A. In the discrete case, independence of the components is equivalent to

$$p_{\mathbf{X}}(\mathbf{x}) = \prod_{k=1}^{n} p_{X_k}(x_k) \text{ where } x_k \in \mathbb{R} \text{ for } k = 1, 2, \ldots, n.$$

In the continuous case, independence of the components is equivalent to

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{k=1}^{n} f_{X_k}(x_k) \text{ where } x_k \in \mathbb{R} \text{ for } k = 1, 2, \ldots, n.$$

Definition. Let X and Y be random variables. Their *covariance* is

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

X and Y are *uncorrelated* if $\mathrm{Cov}(X, Y) = 0$. The collection of random variables $X_1, X_2, \ldots, X_n$ are *pairwise uncorrelated* if every pair of distinct random variables are uncorrelated. The *correlation coefficient* of X and Y is

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

Note. If the variances of X and Y are nonzero then when the correlation coefficient $\rho_{X,Y} = 0$ we have that X and Y are uncorrelated. As commented in Section 0.7.

Sums of Random Variables, Covariance, Correlation, it is shown in Mathematical Statistics 1 (STAT 4047/5047) in Section 2.5. The Correlation Coefficient, it is proved that $-1 \leq \rho_{X,Y} \leq 1$ (see Theorem 2.5.1) and that for independent X and Y we have $\rho_{X,Y} = 0$ (see Theorem 2.5.2). However, the converse of this last result may not hold. That is, we may have $\rho_{X,Y} = 0$ yet X and Y may be dependent; see Example 2.5.3 in the Mathematical Statistics 1 notes.

Revised: 10/18/2022