Indian Institute of Technology, Kanpur
Probability Theory, 18–19

Instructors: Purushottam Kar, Neeraj Misra
Authors: Gurpreet Singh
Email: guggu@iitk.ac.in

SCRIBE
7

Joint Distribution of Random Variables and Independence

1. Random Vector

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we may be interested in two or more numerical characteristics of the sample space simultaneously.

Example 7.1. $\mathcal{E}$: Casting two dice simultaneously

$$\Omega \left\{(i,j) \;\middle|\; i,j \in \{1,2\ldots 6\}\right\}$$

Then, we can define two random variables $X_1 : \Omega \to \mathbb{R}$ and $X_2 : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} X_1(\,(i,j)\,) &= i+j \\ X_2(\,(i,j)\,) &= |\,i-j\,| \end{aligned}$$

Since we may be interested in studying X_1 and X_2 simultaneously, we study the function $\underline{X} : \Omega \to \mathbb{R}^2$, where $\underline{X} = [\,X_1, X_2\,]'$, and $\underline{X}(\,(i,j)\,) = [\,X_1(\,(i,j)\,), X_2(\,(i,j)\,)\,]'$ where $(i,j) \in \Omega$. Here, $\underline{X}$ is a random vector.

Definition 7.1 (Random Vector)**.** A function

$$\underline{X} \;=\; [\,X_1, X_2 \ldots X_p\,]' : \Omega \to \mathbb{R}^p$$

is called a p-dimensional random vector ($\mathbb{R}^p$ denotes the p-dimensional Euclidean Space)

2. Probability Distribution for a Random Vector

We can also define a probability measure of the random vector $\underline{X}$ for the sample space Ω and the event space $\mathcal{F}$ as $\mathbb{P}_{\underline{X}} : \mathcal{F} \to [\,0,1\,]$ such that

$$\begin{aligned} \mathbb{P}_{\underline{X}}[\,A\,] &= \mathbb{P}\left[\,\underline{X}^{-1}(A)\,\right] \\ &= \mathbb{P}\left[\,\left\{w \in \Omega \;\middle|\; \underline{X}(w) \in A\right\}\,\right] \end{aligned}$$

It can be proved that this indeed is a probability measure, and hence $\left(\Omega, \mathcal{F}, \mathbb{P}_{\underline{X}}\right)$ is a probability space.

We can also define the Joint Cumulative Distribution Function (Joint CDF) of $\underline{X}$.

Definition 7.2. The joint distribution function (d.f.) of a random vector $\underline{X}$ is the function $F_X : \mathbb{R}_p \to \mathbb{R}$ defined by

$$F_{\underline{X}}(\mathbf{x}) \;=\; \mathbb{P}_{\underline{X}}[\,(-\infty, \mathbf{x}\,]\,]$$

where $(\mathbf{a}, \mathbf{b}\,] = \left\{[\,x_1, x_2 \ldots x_p\,]' \;\middle|\; \forall\, n \in [\,p\,]\,, a_n < x_n \le b_n\right\}$

Definition 7.3 (Marginal Distributive Function)**.** The Joint CDF of any supset of the r.v.s $\{X_1, X_2 \ldots X_p\}$ is called a marginal CDF of $F_{\underline{X}}$. Suppose if $\underline{X} = [\,X_1, X_2 \ldots X_p\,]'$ is a random vector with the joint CDF $F_{\underline{X}}$, then the marginal CDF

$$F_{X_1, X_2 \ldots X_{p-1}}\,(x_1, x_2 \ldots x_{p-1}) \;=\; \lim_{t \to \infty} F_{\underline{X}}\,(x_1, x_2 \ldots x_{p-1}, t)$$

Exercise 7.1. Prove the derived term for the marginal CDF

The above result suggests that to get a marginal CDF, we need to take (in limit) the arguments of unwanted variables in the joint CDF to ∞.

Result 7.3.1. Let $\underline{X} = [\,X_1, X_2 \dots X_p\,]'$ be a p-dimensional random vector with joint CDF F_X. Then, for any p-dimensional "rectangle" $(\mathbf{a}, \mathbf{b}\,]$

$$\begin{aligned} \mathbb{P}\left[\,\underline{X} \in (\mathbf{a}, \mathbf{b}\,]\,\right] &= \mathbb{P}\left[\,\forall n \in [\,p\,]\,, a_n < X_n \leq b_n\,\right] \\ &= \sum_{k=0}^{p} (-1)^n \sum_{\mathbf{z} \in \Delta_{k,p}((\mathbf{a},\mathbf{b}\,])} F_{\underline{X}}(\mathbf{z}) \end{aligned}$$

where for $k \in \{0, 1 \dots p\}$

$$\Delta_{k,p} = \left\{\mathbf{z} = [\,z_1, z_2 \dots z_p\,] \;\middle|\; k \text{ of } z_n s \text{ are } a_n s \text{ and rest are } b_n s\right\}$$

Exercise 7.2. Prove the above result using induction.

2.1 Properties of a Joint CDF

Just like CDF for a random variable, the joint CDF of a random vector must satisfy the following properties

(i)

$$\lim_{\substack{x_n \to \infty \\ n \in [\,p\,]}} F_{\underline{X}}\,(x_1, x_2 \dots x_p) = 1$$

(ii)

$$\lim_{\substack{x_n \to -\infty \\ n \in [\,p\,]}} F_{\underline{X}}\,(x_1, x_2 \dots x_p) = 0$$

(iii) $F_{\underline{X}}$ is right continuous in each dimension, keeping other dimensions fixed

(iv) For each "rectangle" $[\,\mathbf{a}, \mathbf{b}\,] \subseteq \mathbb{R}^p$

$$\mathbb{P}\left[\,\mathbf{a} \leq \underline{X} \leq \mathbf{b}\,\right] \geq 0$$

If a function $G : \mathbb{R} \to [\,0, 1\,]$ satisfies properties mentioned aboce, then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a r.v. $\underline{X} = [\,X_1, X_2 \dots X_p\,]'$ on Ω such that G is the joint CDF of $\underline{X}$

3. Independence of Random Variables

<++>
<++>