

Instructors: Purushottam Kar, Neeraj Misra

Authors: Gurpreet Singh

Date: Decemenber 1, 2017

Joint Distribution of Random Variables and Independence

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

$$\Omega \{(i, j) \mid i, j \in \{1, 2 \dots 6\}\}$$

Then, we can define two random variables $X_1 : \Omega \rightarrow \mathbb{R}$ and $X_2 : \Omega \rightarrow \mathbb{R}$ 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 $\underline{X} : \Omega \rightarrow \mathbb{R}^2$, 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

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

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

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 $\mathbb{P}_{\underline{X}} : \mathcal{F} \rightarrow [0, 1]$ such that

$$\begin{aligned} \mathbb{P}_{\underline{X}}[A] &= \mathbb{P}[\underline{X}^{-1}(A)] \\ &= \mathbb{P}[\{w \in \Omega \mid \underline{X}(w) \in A\}] \end{aligned}$$

It can be proved that this indeed is a probability measure, and hence $(\Omega, \mathcal{F}, \mathbb{P}_{\underline{X}})$ 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 $F_{\underline{X}} : \mathbb{R}_p \rightarrow \mathbb{R}$ defined by

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

where $(\mathbf{a}, \mathbf{b}] = \{[x_1, x_2 \dots x_p]' \mid \forall n \in [p], a_n < x_n \leq b_n\}$

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

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

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 $F_{\underline{X}}$. Then, for any p-dimensional “rectangle” $(\mathbf{a}, \mathbf{b}]$

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

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

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

Exercise 7.2. Prove the above result using induction.

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)

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

(ii)

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

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

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

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

If a function $G : \mathbb{R} \rightarrow [0, 1]$ satisfies properties mentioned above, 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}

Independence of Random Variables

<++>

<++>