AI 5090: Stochastic Processes

Lecture 04

Scribe: Vedansh Chaturvedi, Raghavan R

17 January 2025

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Contents

This lecture focusses on extending the notion of a random variable to higher dimensions. We start with a simple instance of a bivariate random vector, and generalize that to a n-variate random vector. A brief discussion on sequence of random variables concludes our scribe.

Topics Covered: Bivariate Random Vectors and Its Properties, Demystifying $\mathscr{B}(\mathbb{R}^2)$, Multi-Variate Random Vectors and Its Properties, Joint Cumulative Distribution Function, Probability Law of Multi-Variate Random Vector, Sequence of Random Variables.

Note: This scribe considers $(\Omega, \mathscr{F}, \mathbb{P})$ as the measurable space for all defintions.

1 Bivariate Random Vector

Recall the definition of a random variable:

Definition 1. *A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if*

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F}, \quad \text{for all } B \in \mathscr{B}(\mathbb{R}).$$

1.1 Definition of Bivariate Random Vector

An extension of the aforementioned definition to two functions, $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$, encompasses a **bivariate random vector**.

Definition 2. *Given two functions $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$, we say $(X, Y) : \Omega \to \mathbb{R}^2$ is a bivariate random vector with respect to $\mathscr{F}$ if*

$$(X, Y)^{-1}(B) = \{\omega \in \Omega : (X(\omega), Y(\omega)) \in B\} \in \mathscr{F}, \quad \forall B \in \mathscr{B}(\mathbb{R}^2).$$

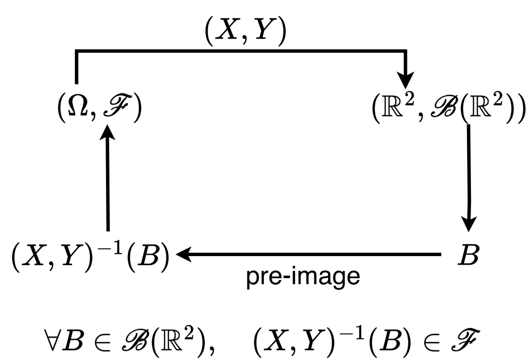


Figure 1: Schematic representation of a bivariate random vector.

Figure 1 simplifies the complex definition of a bivariate random vector. The function (X, Y) takes an ω from the sample space Ω and outputs a vector $\vec{b} \in \mathbb{R}^2$. To ensure that (X, Y) is a bivariate random vector, pick any arbitrary $B \in \mathscr{B}(\mathbb{R}^2)$ and check $(X, Y)^{-1}(B) \in \mathscr{F}$. If that is the case, then (X, Y) is a bivariate random vector with respect to $\mathscr{F}$.

A bivariate random vector (X, Y) is realized as follows:

$$(X, Y)(\omega) = \begin{bmatrix} X(\omega) \\ Y(\omega) \end{bmatrix}$$

Note: When we use the term 'realized' for a bivariate random vector (X, Y), we often mean that both the functions X and Y simultaneously generate outputs for an outcome $\omega \in \Omega$. Figure 2 highlights it. (X, Y) is a function by itself.

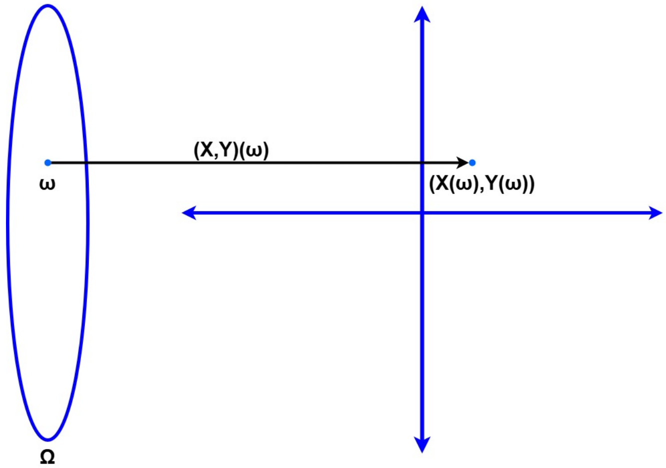


Figure 2: A visual depiction of the realization of (X, Y).

1.2 Demystifying $\mathscr{B}(\mathbb{R}^2)$

Consider a collection of open intervals, $\mathscr{O} = \{(a, b) : a, b \in \mathbb{R}\}$. We define Borel sets in $\mathbb{R}$, $\mathscr{B}(\mathbb{R})$, as $\sigma(\mathscr{O})$. We know that $\mathscr{B}(\mathbb{R})$ can be sufficiently expressed as $\sigma(\mathscr{D})$, where $\mathscr{D} = \{(-\infty, x] : x \in \mathbb{R}\}$. $\mathscr{D}$ is a π-system that also happens to be a generating set. Taking sets of the form $\mathscr{D}$ and constructing the smallest σ-algebra, $\sigma(\mathscr{D})$, will lead us to $\mathscr{B}(\mathbb{R})$.

Do we have a similar generating set for $\mathscr{B}(\mathbb{R}^2)\}$? Yes!

Collection $\mathscr{E} = \{(-\infty, x] \times (-\infty, y] : x, y \in \mathbb{R}\}$, which is a π-system, can be geometrically viewed as a set of semi-infinite rectangles as in figure 3. It also happens to be a generating set, $\sigma(\epsilon) = \mathscr{B}(\mathbb{R}^2)$.

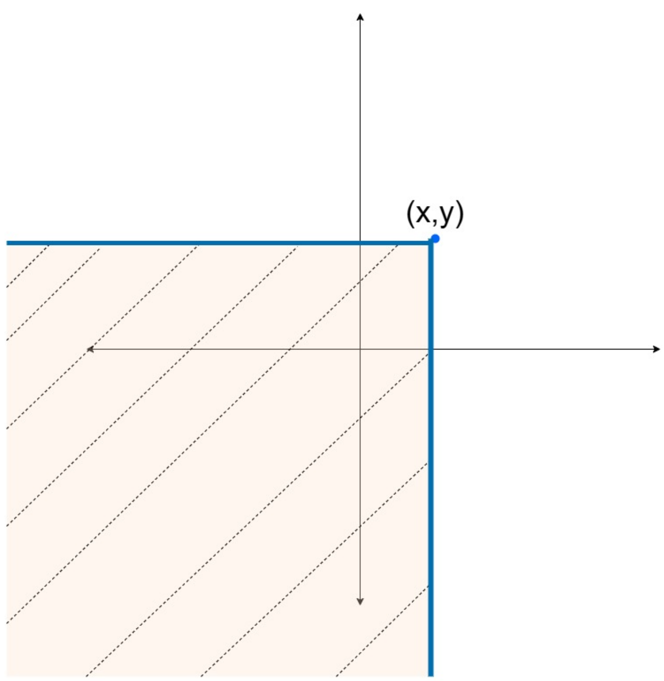


Figure 3: Semi-infinite rectangle closed at X = x and Y = y.

In $\mathscr{B}(\mathbb{R}^2)$, we have other sets like:

1. $(-\infty, x] \times \mathbb{R}$: To get this, let y take any real value. It leads to a form of semi-infinite rectangle in figure 4.
2. $(-\infty, x) \times \mathbb{R}$: To obtain sets of this form, we perform $\bigcap_{n=1}^{\infty}(-\infty, x - 1/n] \times \mathbb{R}$.

3. $(a, b) \times \mathbb{R}$ ($a, b \in \mathbb{R}$): To obtain sets of this form, we perform $((-\infty, b) \times \mathbb{R}) \bigcap ((-\infty, a] \times \mathbb{R})^{\complement}$. Refer figure 5 for a visual reference.

4. $x \in \mathbb{R}$: To obtain sets of this form, we perform $\bigcap_{n=1}^{\infty}(x-1/n, x+1/n) \times (-1/n, 1/n)$. Sets of the form $(a, b) \times (c, d)$ is also a candidate set in $\mathscr{B}(\mathbb{R}^2)$.

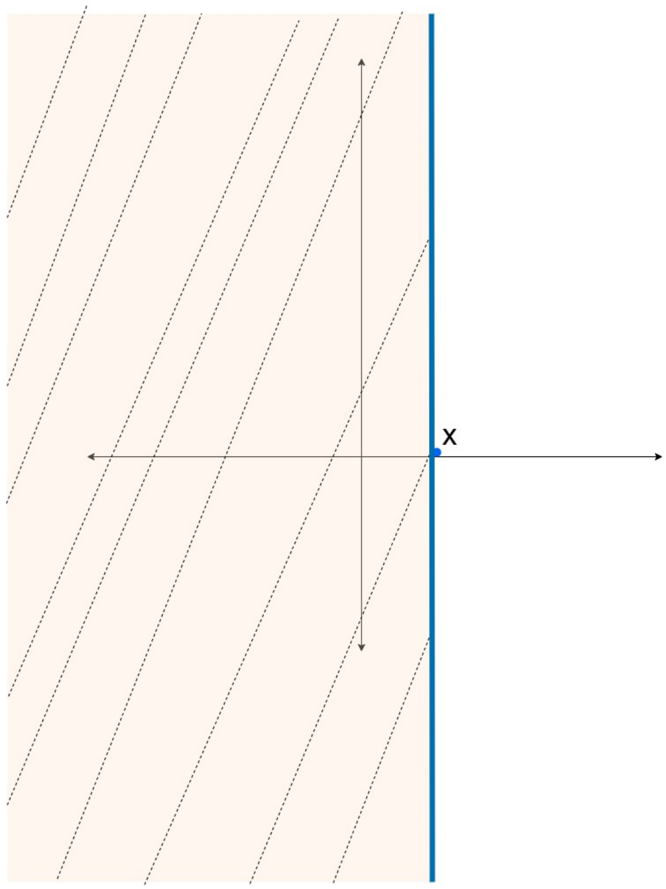


Figure 4: 'Half'-infinite rectangle closed at $X = x$.

Additionally, we could also have circles of radius r centered at the origin ($r > 0$). Some sets have been exempted from the above list as they can be derived from the aforementioned sets.

1.3 Transitioning from $(\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2))$ to $(\Omega, \mathscr{F})$ World

Let $B = (-\infty, x] \times (-\infty, y]$, for some $x, y \in \mathbb{R}$. Assume that (X, Y) is a bivariate random vector with respect to $\mathscr{F}$. Then:

$$
\begin{aligned}
(X, Y)^{-1}((-\infty, x] \times (-\infty, y]) &= \{\omega \in \Omega : (X(\omega), Y(\omega)) \in (-\infty, x] \times (-\infty, y]\} \in \mathscr{F} \\
&= \{\omega \in \Omega : X(\omega) \leq x, Y(\omega) \leq y\} \in \mathscr{F} \\
&= \{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\} \in \mathscr{F}
\end{aligned}
$$

1.4 Identity of X and Y in Bivariate Random Vector (X, Y)

We have been viewing X and Y as functions from Ω to $\mathbb{R}$.
Does their existence move beyond the definition of a normal function? Yes!

Let us take sets of the form $B = (-\infty, x] \times \mathbb{R}$, $x \in \mathbb{R}$. Then:

$$(X, Y)^{-1}(B) = \{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \in \mathbb{R}\}.$$

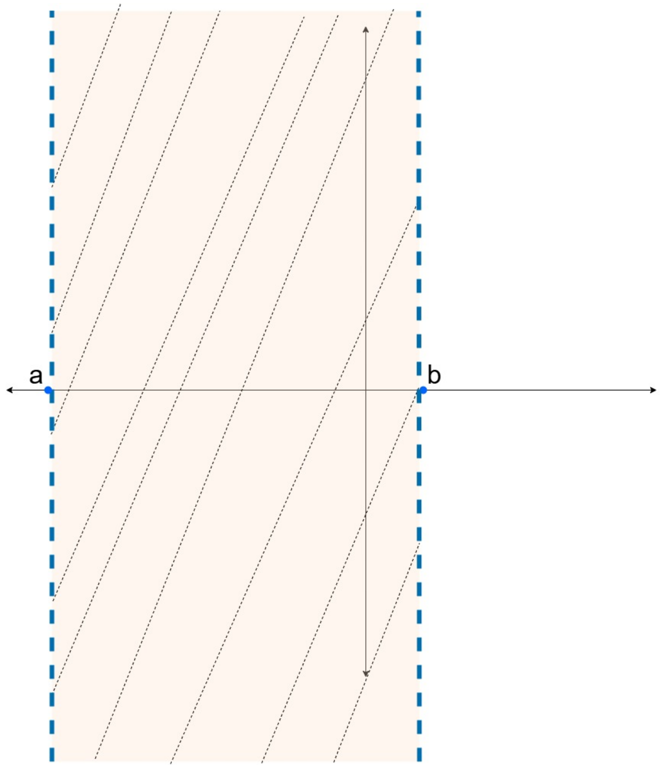


Figure 5: Rectangle that encapsulates $(a, b) \times \mathbb{R}$.

Since $\{\omega \in \Omega : Y(\omega) \in \mathbb{R}\}$ translates to the sample space Ω itself, we are left with:

$$(X, Y)^{-1}(B) = \{\omega \in \Omega : X(\omega) \leq x\}$$

As (X, Y) is a bivariate random vector, $(X, Y)^{-1}(B) \in \mathscr{F}$. This implies that:

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F}$$

By recalling the definition of a univariate random variable, we can see that X **is also a random variable**.

Consider $B = \mathbb{R} \times (-\infty, y], y \in \mathbb{R}$. By following the same steps as before, we can arrive at a conclusion that Y **must also be a random variable**. We have arrived at an important implication:

If (X, Y) is a bivariate random variable, X and Y are univariate random variables.

1.4.1 What About The Converse?

Suppose X and Y are both random variables, then will (X, Y) be a bivariate random vector?

Let $B = (-\infty, x] \times (-\infty, y]$, for an arbitrary $x, y \in \mathbb{R}$

$$(X, Y)^{-1}(B) = \{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\}$$

As X and Y are both random variables, $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F}$ and $\{\omega \in \Omega : Y(\omega) \leq y\} \in \mathscr{F}$. If $A \in \mathscr{F}$ and $B \in \mathscr{F}$, we know that $A \cap B \in \mathscr{F}$ by the theoretical constructs of a σ-algebra. Therefore:

$$\{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\} \in \mathscr{F} = (X, Y)^{-1}(B) \in \mathscr{F}$$

If X and Y are both random variables, then (X, Y) is definitely a bivariate random vector. This gives us an opportunity to restructure the previous implication as follows:

(X, Y) is a bivariate random vector $\iff$ X and Y are individually random variables.

1.5 Practical Examples of Bivariate Random Vectors

Random variables measure a certain aspect of some random experiment. Consider the typical coin toss experiment. When we model the number of heads as a random variable, we are essentially measuring the number of heads that occurs as an outcome of the coin toss. This notion of random variable as a measurement abstracts out the underlying phenomena that may have lead to the measurements in the first place. The $(\Omega, \mathscr{F})$ can be neglected once we have a complete description of the random variable $X : \Omega \to \mathbb{R}$. We can rely on the $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ world to get our job done.

A bivariate random vector is two random variables that get realized simultaneously once an outcome is generated by a random experiment. Consider a machine that generates both temperature and humidity of an environment it is placed in. Temperature and humidity together can be modelled as a bivariate random vector. The underlying phenomena that leads to these measurements is not known. There are instances where we may not know Ω in practice. This example is one such case. We tend to work with the $(\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2))$ world, which happens to be temperature and humidity in the aforementioned example. These measurements of the random experiment (the universe has command over the experiment in our example) can be further analyzed to understand about the unknown.

1.6 Probability Law of a Bivariate Random Vector

Definition 3. *Given two random variables $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ with respect to $\mathscr{F}$, their joint probability law $\mathbb{P}_{X,Y} : \mathscr{B}(\mathbb{R}^2) \to [0, 1]$, is the probability measure defined as*

$$\mathbb{P}_{X,Y}(B) = \mathbb{P}(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in B\}), \qquad B \in \mathscr{B}(\mathbb{R}^2)$$

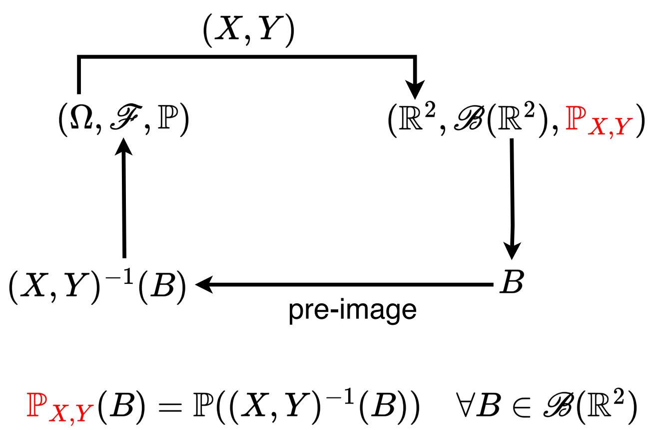


Figure 6: Joint Probability Law of Bivariate Random Vectors.

$\mathbb{P}_{X,Y}$ is called the pushforward of random vector (X, Y). This is also a transition from the probability measure $\mathbb{P}$ in $(\Omega, \mathscr{F})$ world, to $\mathbb{P}_{X,Y}$ in $(\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2))$ world.

1.7 Joint CDF of a Bivariate Random Vector

Joint CDF is the restatement of the joint probability law of bivariate random variables for $B \in \mathscr{E}$, where $\mathscr{E} = \{(-\infty, x] \times (-\infty, y] : x, y \in \mathbb{R}\}$.

Definition 4. *Given random variables $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ with respect to $\mathscr{F}$, their joint CDF $F_{X,Y}(x, y) : \mathbb{R}^2 \to [0, 1]$ is defined as*

$$F_{X,Y}(x, y) = \mathbb{P}_{X,Y}((-\infty, x] \times (-\infty, y]) = \mathbb{P}(\{X \le x\}, \{Y \le y\}), \quad x, y \in \mathbb{R}$$

1.8 Joint CDF $\iff$ Joint Probability Law

If we know the joint probability law $\mathbb{P}_{X,Y} = \{\mathbb{P}_{X,Y}(B) : B \in \mathscr{B}(\mathbb{R}^2)\}$, we can uniquely extract the joint CDF $F_{X,Y} : \mathbb{R}^2 \to [0, 1]$ by using the formula

$$F_{X,Y}(x,y) = \mathbb{P}_{X,Y}((-\infty,x],(-\infty,y]), \quad x,y \in \mathbb{R}$$

If given only the joint CDF $F_{X,Y} : \mathbb{R}^2 \to [0,1]$, we can let

$$\mathbb{P}_{X,Y}((-\infty,x],(-\infty,y]) = F_{X,Y}(x,y), \quad x,y \in \mathbb{R}$$

By Caratheodory's extension theorem, there exists a unique extension of $\mathbb{P}_{X,Y}$ to all Borel subsets of $\mathbb{R}^2$.

2 High Dimensional Random Vectors

2.1 Definition of Random Vector

Fix $n \in \mathbb{N}$.

Definition 5. *Given random variables $X_1,\dots,X_n$ defined with respect to $\mathscr{F}$, we say $(X_1,\dots,X_n) : \Omega \to \mathbb{R}^n$ is a random vector with respect to $\mathscr{F}$ if*

$$(X_1,\dots,X_n)^{-1}(B) = \{\omega \in \Omega : (X_1(\omega),\dots,X_n(\omega)) \in B\} \in \mathscr{F}, \quad \forall B \in \mathscr{B}(\mathbb{R}^n)$$

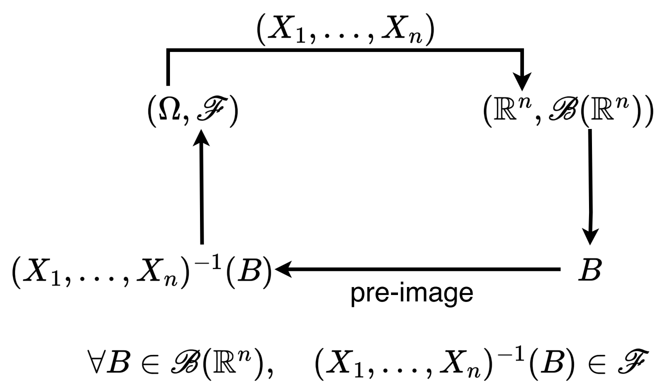


Figure 7: Schematic representation of a High-Dimensional Random Vector.

2.2 Random Vector Property

Equivalently, $(X_1,\dots,X_n)$ is a random vector if:

$$(X_1,\dots,X_n)^{-1}((-\infty,x_1] \times \cdots \times (-\infty,x_n]) = \bigcap_{i=1}^{n} \{\omega \in \Omega : X_i(\omega) \le x_i\} \in \mathscr{F}$$

$$(X_1,\dots,X_n)^{-1}((-\infty,x_1] \times \cdots \times (-\infty,x_n]) = \bigcap_{i=1}^{n} \{X_i \le x_i\} \in \mathscr{F} \quad \forall x_1,\dots,x_n \in \mathbb{R}.$$

$$\boxed{(X_1,\dots,X_n) \text{ random vector} \iff X_1 \text{ RV}, \quad \cdots \quad, X_n \text{ RV}}$$

2.3 Some Important Points

- $(X_1,\dots,X_n)$ is a random vector, if each individual component $X_1,\dots,X_n$ is a random variable.
- The space $\mathbb{R}^n$ represents an n-dimensional Euclidean space.
- In practical scenarios, n can be extremely large (such as $n = 100$, or $n = 1000$). It is commonly encountered in modern data analysis.
- The study of these high-dimensional settings is the focus of the field known as *High Dimensional Statistics*, which explores methods to handle a vast number of variables efficiently.

2.4 Understanding $\mathscr{B}(\mathbb{R}^n)$ for $n > 2$

- Consider the special class of semi-infinite rectangles in $\mathbb{R}^n$, given by:

$$\mathcal{E} = \{(-\infty, x_1] \times \cdots \times (-\infty, x_n] : x_1, \ldots, x_n \in \mathbb{R}\}.$$

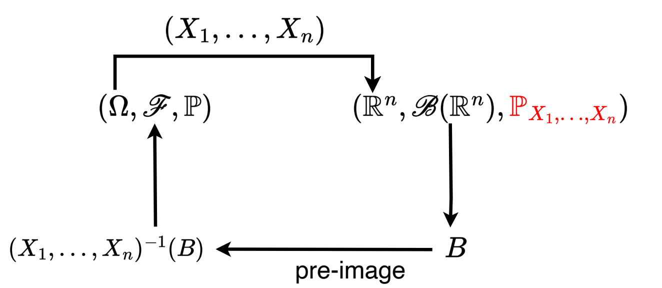
- Extend this set by adding all necessary subsets to form the smallest σ-algebra containing $\mathcal{E}$. This σ-algebra is denoted as $\mathscr{B}(\mathbb{R}^n)$.
- $\mathscr{B}(\mathbb{R}^n) = \sigma(\mathcal{E})$.
- Now, instead of checking whether the preimage of every set in $\mathscr{B}(\mathbb{R}^n)$ with respect to the random vector lies in $\mathscr{F}$, it suffices to check only for sets of the form in $\mathcal{E}$ because it forms the generating class for $\mathscr{B}(\mathbb{R}^n)$.

3 Probability Law and CDF of Random Vectors

3.1 Probability Law of Random Vectors

Definition 6. *Given random variables $X_1, \ldots, X_n$ defined with respect to $\mathscr{F}$, their* **joint probability law** *is the probability measure $\mathbb{P}_{X_1,\ldots,X_n} : \mathscr{B}(\mathbb{R}^n) \to [0,1]$ defined as*

$$\mathbb{P}_{X_1,\ldots,X_n}(B) = \mathbb{P}\left(\{\omega \in \Omega : (X_1(\omega), \ldots, X_n(\omega)) \in B\}\right), \quad B \in \mathscr{B}(\mathbb{R}^n).$$



$$\mathbb{P}_{X_1,\ldots,X_n}(B) = \mathbb{P}((X_1, \ldots, X_n)^{-1}(B)) \quad \forall B \in \mathscr{B}(\mathbb{R}^n)$$

Figure 8: Schematic Representation of Probability Law of Random Vector.

- To assign probabilities to sets in $\mathscr{B}(\mathbb{R}^n)$, it suffices to assign probabilities to the generating class $\mathcal{E}$.
- Then by using Carathéodory's extension theorem, the probability assignment can be uniquely extended to the entire σ-algebra $\mathscr{B}(\mathbb{R}^n)$.

3.2 Joint CDF of Random Variables

Definition 7. *Given random variables $X_1, \ldots, X_n$ defined w.r.t. $\mathscr{F}$, their* **joint CDF** *$F_{X_1,\ldots,X_n} : \mathbb{R}^n \to [0,1]$ is defined as*

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathbb{P}_{X_1,\ldots,X_n}\left((-\infty, x_1] \times \cdots \times (-\infty, x_n]\right)$$

$$= \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \le x_i\}\right) = \mathbb{P}(X_1 \le x_1, \ldots, X_n \le x_n), \quad x_1, \ldots, x_n \in \mathbb{R}.$$

- We can also call it the CDF of the random vector $(X_1, \ldots, X_n)$.

4 Sequence of Random Variables

4.1 Definition of Sequence of Random Variables

Definition 8. *A sequence of random variables is a collection* $\{X_n\}_{n=1}^{\infty}$ *such that*

$$\forall n \in \mathbb{N}, \forall k_1, \ldots, k_n \in \mathbb{N}, \quad (X_{k_1}, \ldots, X_{k_n}) \text{ is a random vector.}$$

- The problem with defining a sequence of random variables as a usual random vector is that the sequence contains countably infinite random variables.
- A sequence of random variables is indexed by the natural numbers, meaning it maps to the space $\mathbb{R} \times \mathbb{R} \times \cdots$ (countably infinite times). This **index set** has the same cardinality as the set of natural numbers ($\mathbb{N}$).
- According to topology, this space is very complicated and challenging to analyze.
- The solution to this problem is to avoid analyzing this space directly. Instead, consider finitely many coordinates of this space, and on these coordinates, one should be able to define a random vector.
- For example, consider $(X_1, X_2, \ldots, X_{10})$—this must be a random vector. Another example which must be a random vector from this space can be (X_1, X_{100}, X_{999}).

Note

When an outcome $\omega \in \Omega$ results from the random experiment, the **entire sequence** $X_1(\omega), X_2(\omega), \ldots$ realizes at once.

- The trajectory of the realizations of random variables from a single outcome $\omega \in \Omega$ is called the **Sample Path** corresponding to the outcome ω.

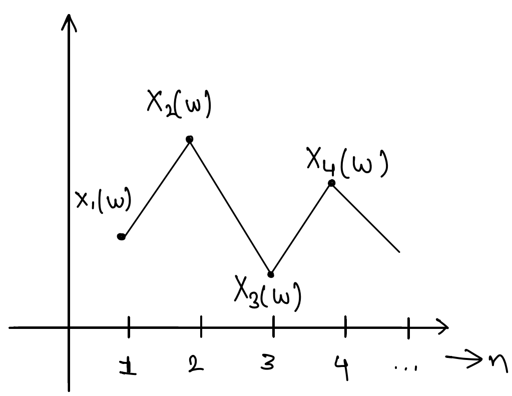


Figure 9: Example of a Sample Path resulting from an outcome $\omega \in \Omega$.

4.2 Finite Dimensional Distributions

Definition 9. *Consider a sequence of random variables* $\{X_n\}_{n=1}^{\infty}$ *w.r.t.* $\mathscr{F}$. *The collection*

$$\left\{F_{X_{k_1}, \ldots, X_{k_n}} : n \in \mathbb{N},\ k_1, \ldots, k_n \in \mathbb{N}\right\}$$

is called the set of all **finite-dimensional distributions** *of the sequence* $\{X_n\}_{n=1}^{\infty}$.