Lecture 11: Random Variables

Lecturer: Dr. Krishna Jagannathan   *Scribe: Sudharsan, Gopal, Arjun B, Debayani*

The study of random variables is motivated by the fact that in many scenarios, one might not be interested in the precise elementary outcome of a random experiment, but rather in some numerical function of the outcome. For example, in an experiment involving ten coin tosses, the experimenter may only want to know the total number of heads, rather than the precise sequence of heads and tails.

The term random variable is a misnomer, because a random variable is neither random, nor is it a variable. A random variable X is a function from the sample space Ω to real field $\mathbb{R}$. The term 'random' actually signifies the underlying randomness in picking an element ω from the sample space Ω. Once the elementary outcome ω is fixed, the random variable takes a fixed real value, $X(\omega)$. It is important to remember that the probability measure is associated with subsets (events), whereas a random variable is associated with each elementary outcome ω.

Just as not all subsets of the sample space are not necessarily considered events, not all functions from Ω to $\mathbb{R}$ are considered random variables. In particular, a random variable is an $\mathcal{F}$-*measurable* function, as we define below.

Definition 11.1 *Measurable function:*
Let $(\Omega, \mathcal{F})$ be a measurable space. A function $f : \Omega \to \mathbb{R}$ is said to be an $\mathcal{F}$-measurable function if the pre-image of every Borel set is an $\mathcal{F}$-measurable subset of Ω.

In the above definition, the pre-image of a Borel set B under the function f is given by

$$f^{-1}(B) \triangleq \{\omega \in \Omega \mid f(\omega) \in B\}. \tag{11.1}$$

Thus, according to the above definition, $f : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable function if $f^{-1}(B)$ is an $\mathcal{F}$-measurable subset of Ω for *every* Borel set B.

Definition 11.2 *Random Variable:*
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random variable X is an $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}$.

In other words, for every Borel set B, its pre-image under a random variable X is an *event.* In Figure 11.1, X is a random variable that maps every element ω in the sample space Ω to the real line $\mathbb{R}$. B is a Borel set, i.e., $B \in \mathcal{B}(\mathbb{R})$. The inverse image of B is an event $E \in \mathcal{F}$.

Since the set $\{\omega \in \Omega | X(\omega) \in B\}$ is an event for every Borel set B, it has an associated probability measure. This brings us to the concept of the *probability law* of the random variable X.

Definition 11.3 *Probability law of a random variable X:*
The probability law $\mathbb{P}_X$ of a random variable X is a function $\mathbb{P}_X : \mathcal{B}(\mathbb{R}) \to [0, 1]$, which is defined as $\mathbb{P}_X(B) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \in B\})$.

Thus, the probability law can be seen as the composition of $\mathbb{P}(\cdot)$ with the inverse image $X^{-1}(\cdot)$, i.e., $\mathbb{P}_X(\cdot) = \mathbb{P} \circ X^{-1}(\cdot)$. Indeed, the probability law of a random variable completely specifies the statistical properties of
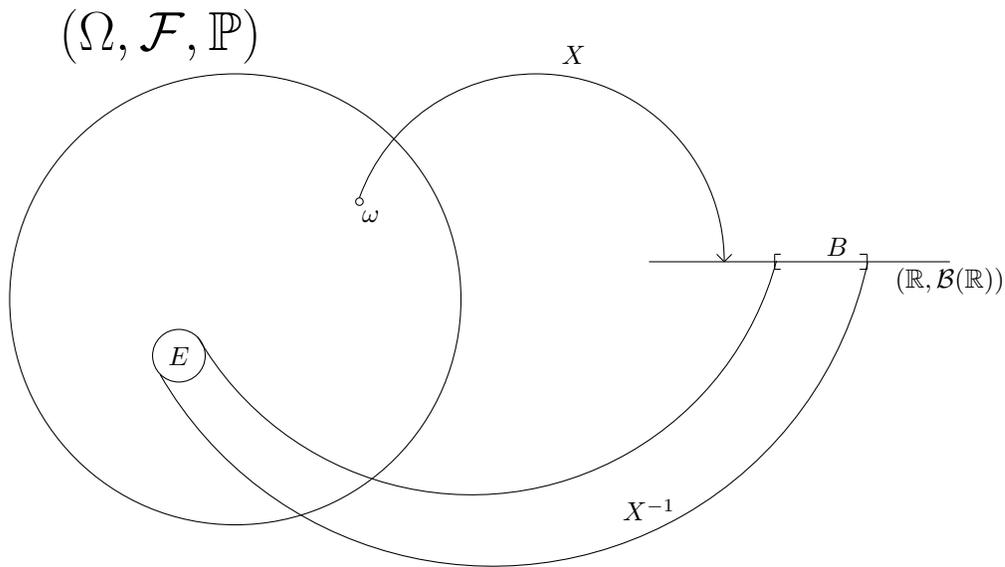
Figure 11.1: A random variable $X : \Omega \to \mathbb{R}$. The pre-image of a Borel set B is an event E.
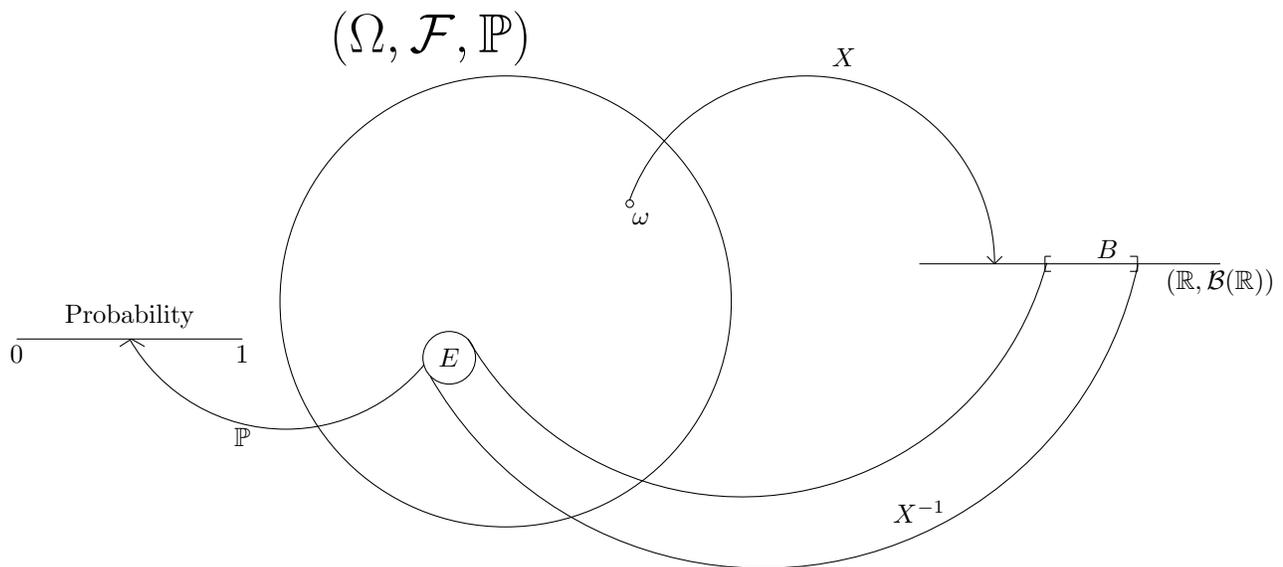

Figure 11.2: The probability law $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ specifies the probability that the random variable X takes a value in some particular Borel set.

the random variable, as it specifies the probability of the random variable taking values in any given Borel set.

In Figure 11.2, $\mathbb{P}$ is the mapping from event E to the probability space. $\mathbb{P}_X$ is the mapping from B to the probability space such that $\mathbb{P}_X$ is a composition of $\mathbb{P}$ with X^{-1}.

Theorem 11.4 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let X be a real-valued random variable. Then, the probability law $\mathbb{P}_X$ of X is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

Next, we make a short digression to introduce a mathematical structure known as a π-*system* (read as pi-system).

Definition 11.5 *Given a set Ω, a π-system on Ω is a non-empty collection of subsets of Ω that is stable under finite intersections. That is, $\mathcal{P}$ is a π-system on Ω, if A, $B \in \mathcal{P}$ implies $A \cap B \in \mathcal{P}$.*

One of the most commonly used π-systems on $\mathbb{R}$ is the class of all closed semi-infinite intervals defined as

$$\pi(\mathbb{R}) \triangleq \{(-\infty, x] : x \in \mathbb{R}\}. \tag{11.2}$$

Lemma 11.6 *The σ-algebra generated by $\pi(\mathbb{R})$ is the Borel σ-algebra, i.e.,*

$$\mathcal{B}(\mathbb{R}) = \sigma(\pi(\mathbb{R})).$$

Now, we turn our attention to a key result from measure theory, which states that if two finite measures agree on a π-system, then they also agree on the σ-algebra generated by that π-system.

Lemma 11.7 *Uniqueness of extension, π-systems:-* *Let Ω be a given set, and let $\mathcal{P}$ be a π-system over Ω. Also, let $\Sigma = \sigma(\mathcal{P})$ be the σ-algebra generated by the π-system $\mathcal{P}$. Suppose μ_1 and μ_2 are measures defined on the measurable space (Ω, Σ) such that $\mu_1(\Omega) = \mu_2(\Omega) < \infty$ and $\mu_1 = \mu_2$ on $\mathcal{P}$. Then,*

$$\mu_1 = \mu_2 \text{ on } \Sigma.$$

Proof: See Section A1.4 of [1]. ∎

In particular, for probability measures, we have the following corollary:

Corollary 11.8 *If two probability measures agree on a π-system, then they agree on the σ-algebra generated by that π-system.*

In particular, if two probability measures agree on $\pi(\mathbb{R})$, then they must agree on $\mathcal{B}(\mathbb{R})$. This result is of importance to us since working with σ-algebras is difficult, whereas working with π-systems is easy!

11.1   Cumulative Distribution Function (CDF) of a Random Variable

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a Probability Space and let $X : \Omega \longrightarrow \mathbb{R}$ be a random variable. Consider the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ induced on the real line by X. Recall that $\mathcal{B}(\mathbb{R}) = \sigma(\pi(\mathbb{R}))$ is the Borel σ-algebra whose

generating class is the collection of semi-infinite intervals (or equivalently, the open intervals). Therefore, for any $x \in \mathbb{R}$,

$$(-\infty, x] \in \mathcal{B}(\mathbb{R}) \Rightarrow X^{-1}\left((-\infty, x]\right) \in \mathcal{F}.$$

It is therefore legitimate to look at the probability law $\mathbb{P}_X$ of these semi-infinite intervals. This is, by definition, the Cumulative Distribution Function (CDF) of X, and is denoted by $F_X(.)$.

Definition 11.9 *The CDF of a random variable X is defined as follows:*

$$F_X(x) \triangleq \mathbb{P}_X\left((-\infty, x]\right) = \mathbb{P}\left(\{\omega | X(\omega) \leq x\}\right), \quad x \in \mathbb{R}. \tag{11.3}$$

Since the notation $\mathbb{P}(\{\omega | X(\omega) \leq x\})$ is a bit tedious, we will use $\mathbb{P}(X \leq x)$ although it is an abuse of notation. Remarkably, it turns out that it is enough to specify the CDF in order to completely characterize the probability law of the random variable! The following theorem asserts this:

Theorem 11.10 *The Probability Law $\mathbb{P}_X$ of a random variable X is uniquely specified by its CDF $F_X(.)$.*

Proof: This is a consequence of the uniqueness result, Lemma 11.7. Another approach is to use the Carathèodory's extension theorem. Here, we present only an overview of the proof.
Let $\mathcal{F}_0$ denote the collection of finite unions of sets of the form $(a, b]$, where $a < b$ and $a, b \in \mathbb{R}$. Define a set function $\mathbb{P}^0 : \mathcal{F}_0 \longrightarrow [0, 1]$ as $\mathbb{P}^0((a, b]) = F_X(a) - F_X(b)$. Having verified countable additivity of $\mathbb{P}^0$ on $\mathcal{F}_0$, we can invoke Carathèodory's Theorem, thereby obtaining a measure $\mathbb{P}_X$ which uniquely extends $\mathbb{P}^0$ on $\mathcal{B}(\mathbb{R})$.

11.2   Properties of CDF

Theorem 11.11 *Let X be a random variable with CDF $F_X(.)$. Then $F_X(.)$ posses the following properties:*

1. *If $x \leq y$, then $F_X(x) \leq F_X(y)$ i.e. the CDF is monotonic non-decreasing in its argument.*

2. $\lim\limits_{x \longrightarrow \infty} F_X(x) = 1$ *and* $\lim\limits_{x \longrightarrow -\infty} F_X(x) = 0$.

3. *$F_X(.)$ is right-continuous i.e. $\forall\ x \in \mathbb{R}$, $\lim\limits_{\epsilon \downarrow 0} F_X(x + \epsilon) = F_X(x)$.*

Proof:

1. Since, for $x \leq y$, $\{\omega | X(\omega) \leq x\} \subseteq \{\omega | X(\omega) \leq y\}$, from monotonicity of the probability measure, it follows that $F_X(x) \leq F_X(y)$.

2. We have

$$
\begin{aligned}
\lim_{x \longrightarrow -\infty} F_X(x) \;&\overset{(a)}{=}\; \lim_{x \longrightarrow -\infty} \mathbb{P}(X \leq x), \\
&\overset{(b)}{=}\; \lim_{n \longrightarrow \infty} \mathbb{P}(X \leq x_n), \\
&\overset{(c)}{=}\; \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} \{\omega : X(\omega) \leq x_n\}\right), \\
&=\; \mathbb{P}(\emptyset) = 0,
\end{aligned}
$$

where (a) follows from the definition of a CDF, (b) follows by considering a sequence $\{x_n\}_{n\in\mathbb{N}}$ that decreases monotonically to $-\infty$, and (c) is a consequence of continuity of probability measures.

Following a very similar derivation, and considering a sequence $\{x_n\}_{n\in\mathbb{N}}$ that monotonically increases to $+\infty$, we get:

$$
\begin{aligned}
\lim_{x\longrightarrow\infty} F_X(x) &= \lim_{x\longrightarrow\infty} \mathbb{P}(X \leq x), \\
&= \lim_{n\longrightarrow\infty} \mathbb{P}(X \leq x_n), \\
&= \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{\omega : X(\omega) \leq x_n\}\right), \\
&= \mathbb{P}(\Omega), \\
&= 1.
\end{aligned}
$$

3. Consider a sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ decreasing to zero. Therefore, for each $x \in \mathbb{R}$,

$$
\begin{aligned}
\lim_{\epsilon\downarrow 0} F_X(x + \epsilon) &\overset{(a)}{=} \lim_{\epsilon\downarrow 0}\mathbb{P}(X \leq x + \epsilon), \\
&= \lim_{n\longrightarrow\infty} \mathbb{P}(X \leq x + \epsilon_n), \\
&\overset{(b)}{=} \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} \{\omega : X(\omega) \leq x + \epsilon_n\}\right), \\
&= \mathbb{P}(X \leq x), \\
&= F_X(x),
\end{aligned}
$$

where (a) follows from the definition of CDF, and (b) follows from continuity of probability measures.

Note that, in general, a CDF need not be continuous. But right-continuity must necessarily be satisfied by any CDF. It turns out that not only are the above three properties satisfied by all CDFs, but any function that satisfies these properties is necessarily a CDF of some random variable!

Theorem 11.12 *Let F be a function satisfying the three properties of a CDF as in theorem (11.11). Consider the Probability Space $\Omega = ([0, 1), \mathcal{B}([0, 1)), \lambda)$. Then, there exists a random variable $X : \Omega \longrightarrow \mathbb{R}$ whose CDF is F.*

A constructive proof can be found in [3].