LECTURE 4
DISCRETE RANDOM VARIABLES, PROBABILITY MASS FUNCTIONS (PMFs)

We already defined in Lecture 2 discrete random variables through the CDF: their CDFs are step functions. In this lecture, we will take a closer look at how discrete distributions arise.

Suppose that a random variable X takes values in a finite or countable subset of $\mathbb{R}$:

$$X : \Omega \to S \subset \mathbb{R},$$

where

$$S = \{x_1, x_2, \ldots, x_n\}$$

or

$$S = \{x_1, x_2, \ldots, x_n, x_{n+1}, \ldots\}.$$

Suppose further that

$$P(X = x) > 0 \text{ for all } x \in S.$$

The set S is called the *support of the distribution* of X. In the first case, the support is finite and consists only of n elements. In the second case, the support is infinite but *countable*: its elements can be put in one-to-one correspondence with positive integers (natural numbers). In either case, such random variables are called *discrete* due to the discrete nature of their support.

Definition 1. A discrete random variable is a random variable that takes a countable number (finite or infinite) of values with *strictly* positive probabilities.

Remark. We will see later that this definition is consistent with defining random variables as those that have a step function for CDFs (Lecture 3, Definition 11).

Example. (a) Suppose that $\Omega = \{H, T\}$, $P(\{H\}) = P(\{T\}) = 1/2$, and $X(\omega) = 1(\omega = H)$, i.e.

$$X = \begin{cases} 1, & \omega = H, \\ 0, & \omega = T. \end{cases}$$

In this case, the support is $S = \{0, 1\}$, since for any other number $x \notin S$, $P(X = x) = 0$.

(b) In rolling a dice experiment, $S = \{1, 2, 3, 4, 5, 6\}$, and $P(X = x) = 1/6$ for $x \in S$.

(c) Consider the following experiment. Roll a dice. If the outcome is 6, the experiment stops. If the outcome is different from six, one has to roll again. Let X denote the number of rolls needed to obtain 6. The sample space is infinite but countable:

$$S = \{1, 2, 3, \ldots\}.$$

Also,

$$\begin{aligned}
P(X=1) &= \frac{1}{6}, \\
P(X=2) &= \frac{5}{6} \times \frac{1}{6}, \\
P(X=3) &= \left(\frac{5}{6}\right)^2 \times \frac{1}{6}, \\
&\cdots \\
P(X=n) &= \left(\frac{5}{6}\right)^{n-1} \times \frac{1}{6}.
\end{aligned}$$

Note that in Examples (a)-(b) the support of the distribution is finite, while in Example (c) the support is infinite (but countable)

One can describe the distribution of X using CDFs as discussed in Lecture 3. However, since the support of the distribution of a discrete random variable is quite simple, a more convenient description can be obtained with a Probability Mass Function (PMF).

Definition 2. (PMF) A probability mass function of a discrete random variable with a support S is defined as

$$p_X(x) = P(X = x) \text{ for } x \in S,$$

and satisfies

$$p_X(x) > 0 \text{ for all } x \in S,$$

and

$$\sum_{x \in S} p_X(x) = 1. \tag{1}$$

Remark. When S is finite, $S = \{x_1, x_2, \ldots, x_n\}$,

$$\sum_{x \in S} p_X(x) = \sum_{i=1}^{n} p_X(x_i) = p_X(x_1) + p_X(x_2) + \ldots + p_X(x_n).$$

When S is countably infinite, $S = \{x_1, x_2, \ldots\}$,

$$\sum_{x \in S} p_X(x) = \sum_{i=1}^{\infty} p_X(x_i) = p_X(x_1) + p_X(x_2) + \ldots.$$

A PMF can be extended to the entire real line by defining $p_X(x) = 0$ if $x \notin S$.

Example. It is straightforward to find the PMFs in the above examples:

(a) $p_X(x) = 1/2$, $x \in \{0, 1\}$.

(b) $p_X(x) = 1/6$, $x \in \{1, 2, 3, 4, 5, 6\}$.

(c) $p_X(x) = \left(\frac{5}{6}\right)^{x-1} \times \frac{1}{6}$, $x \in \{1, 2, 3, \dots\}$. To verify that condition (1) holds,

$$\begin{aligned}
\sum_{i=1}^{\infty} p_X(x_i) &= \frac{1}{6} + \left(\frac{5}{6}\right) \times \frac{1}{6} + \left(\frac{5}{6}\right)^2 \times \frac{1}{6} + \dots \\
&= \frac{1}{6} \times \left(1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \dots\right) \\
&= \frac{1}{6} \times \frac{1}{1 - \frac{5}{6}} \\
&= 1.
\end{aligned}$$

A PMF can be used to compute probabilities of the form $P(X \in B)$ where $B \subset \mathbb{R}$:

$$\begin{aligned}
P(X \in B) &= \sum_{s \in S} (p_X(x) \times 1(x \in B)) \qquad (2) \\
&= p_X(x_1) \times 1(x_1 \in B) + p_X(x_2) \times 1(x_2 \in B) + \dots,
\end{aligned}$$

where $1(x \in B)$ is the indicator function for the set B:

$$1(x \in B) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B. \end{cases}$$

Example. Consider again example (c) above. Suppose that $B = \{1, 2\}$, i.e. we consider an event that it took less than three attempts to roll 6:

$$\begin{aligned}
P(X \in B) &= P(X \in \{1, 2\}) \\
&= p_X(1) \times 1(1 \in \{1, 2\}) + p_X(2) \times 1(2 \in \{1, 2\}) + p_X(3) \times 1(3 \in \{1, 2\}) + \dots \\
&= \frac{1}{6} \times 1 + \frac{5}{6} \times \frac{1}{6} \times 1 + \left(\frac{5}{6}\right)^2 \times \frac{1}{6} \times 0 + \dots \\
&= \frac{1}{6} + \frac{5}{6} \times \frac{1}{6} + 0 + \dots \\
&= \frac{11}{36}.
\end{aligned}$$

Note that we would obtain exactly the same result if we take, for example, $B = [1, 2]$ or $[0, 2.4]$.

Using the definition in (2), we can also see the relationship between CDFs and PMFs. Recall that in the case of CDFs, we are concerned with $P(X \in (-\infty, u]) = P(X \leq u)$ for

$u \in \mathbb{R}$. Hence, for a discrete random variable X,

$$\begin{aligned} F_X(u) &= P(X \leq u) \\ &= \sum_{s \in S} (p_X(x) \times 1(x \leq u)) \\ &= p_X(x_1) \times 1(x_1 \leq u) + p_X(x_2) \times 1(x_2 \leq u) + \ldots, \end{aligned}$$

where

$$1(x \leq u) = \begin{cases} 1, & x \leq u \\ 0, & x > u \end{cases}$$

Thus, to compute $F_X(u)$ we add all the values $p_X(x)$ for $x \in S$ such that $x \leq u$. For example, suppose that the elements in S are ordered in increasing order:

$$x_1 < x_2 < x_3 < \ldots.$$

Then,

$$F_X(u) = \begin{cases} 0, & u < x_1, \\ p_X(x_1), & x_1 \leq u < x_2, \\ p_X(x_1) + p_X(x_2), & x_2 \leq u < x_3, \\ \ldots & \ldots \\ p_X(x_1) + p_X(x_2) + \ldots + p_X(x_n), & x_n \leq u < x_{n+1}. \end{cases}$$

As one can see, the CDF is a step function: it changes only in jumps and only u crosses one of the support points. Furthermore, if x_n and x_{n+1} are two subsequent support points, then

$$F_X(x_{n+1}) - F_X(x_n) = p_X(x_{n+1}).$$

The last equation simply reflects the cumulative nature of a CDF.