Types of Random Variables: Discrete and Continuous

Let $(\mathcal{S}, \Sigma, P)$ be a probability space with a random variable $X : \mathcal{S} \longrightarrow \mathbb{R}$, and let $(\mathbb{R}, \mathbb{B}_{\mathbb{R}}, P_X)$ be the probability space induced by X. Let F_X be the distribution function of X. It is known that F_X uniquely determine P_X and vice-versa. Thus, to study the induced probability space $(\mathbb{R}, \mathbb{B}_{\mathbb{R}}, P_X)$, it is sufficient to study the d.f. F_X.

In this course we will restrict ourselves to two types of random variables: discrete and continuous. In the first case, the RV assumes at most a countable number of values and hence its d.f is a step function. In the later case, the d.f. F_X is continuous (we will see the definition later).

Definition 1. *A random variable X is said to be of discrete type, or simply discrete, if there exists a finite or a countable set $E_X \subset \mathbb{R}$ such that $P(\{X = x\}) > 0$, $\forall\, x \in E_X$ and $P(\{X \in E_X\}) = 1$. The set E_X is called the support of the RV X.*

Remark 2.

(1) *If X is any RV with the d.f. F_X, then $P(\{X = x\}) = F_X(x) - F_X(x-)$ for every $x \in \mathbb{R}$. (Prove!)*

(2) *From previous lecture, we know that F_X is discontinuous at $x \in \mathbb{R}$ if and only if $F_X(x-) < F_X(x+) = F_X(x)$. Hence, F_X has only jump discontinuities and the size of the jump at any point x of discontinuity is $P(\{X = x\}) = F_X(x) - F_X(x-)$.*

Remark 3.

(1) *If X is a discrete type RV with support E_X, then*

$$P(\{X \in E_X\}) = \sum_{x \in E_X} P(\{X = x\}) = \sum_{x \in E_X} (F_X(x) - F_X(x-)) = 1.$$

(2) *The d.f. F_X is continuous at every point of E_X^c.*

Definition 4. *Let X be a discrete type random variable with support E_X. The function $f_X : \mathbb{R} \to \mathbb{R}$ defined by,*

$$f_X(x) = \begin{cases} P(\{X = x\}), \text{ if } x \in E_X, \\ 0, \text{ otherwise} \end{cases}$$

is called the probability mass function (p.m.f.) of X.

Remark 5. *Let X be a discrete type RV with support E_X, the d.f. F_X and the p.m.f. f_X.*

(1) *$f_X(x) > 0$, $\forall\, x \in E_X$ and $f_X(x) = 0$, $\forall\, x \notin E_X$.*
(2) *$\sum_{x \in E_X} f_X(x) = 1$.*
(3) *For $A \in \mathbb{B}_{\mathbb{R}}$, we have*

$$\begin{aligned} P_X(A) &= P_X(A \cap E_X) + P_X(A \cap E_X^c) \\ &= P_X(A \cap E_X) \\ &= \sum_{x \in A \cap E_X} f_X(x). \end{aligned}$$

(4) *For $x \in \mathbb{R}$, we have*

$$F_X(x) = \sum_{y \in (-\infty, x] \cap E_X} f_X(y).$$

Example 6. *Consider the the random variable defined as $X(w) = c$ for all $w \in \mathcal{S}$, where c is a fixed real number. Then $P(\{X = c\}) = 1$ and $E_X = \{c\}$. Hence, X is of discrete* *type. Its p.m.f. is given by*

$$f_X(x) = \begin{cases} 1, \text{ if } x = c, \\ 0, \text{ otherwise.} \end{cases}$$

Example 7. *Let X be the indicator function of E, where E is an event. Then $E_X = \{0, 1\}$ and $P(\{X \in E_X\}) = 1$. Thus, X is discrete and its p.m.f. is given by*

$$f_X(x) = \begin{cases} P(E^c), \text{ if } x = 0, \\ P(E), \text{ if } x = 1, \\ 0, \text{ otherwise.} \end{cases}$$

Example 8. *Consider a coin that, in any flip, ends up in head with probability $\frac{1}{4}$ and in tail with probability $\frac{3}{4}$. The coin is tossed repeatedly and independently until a total of two heads have been observed. Let X denote the number of flips required to achieve this. Then $P(\{X = x\}) = 0$, if $x \notin \{2, 3, 4, \ldots\}$. For $n \in \{2, 3, 4, \ldots\}$, let $S_n = \{(w_1, w_2, \ldots, w_n) : w_n = H, w_i = H$ for one i between 1 and $n-1$, and $w_j = T$, for $j \neq i\}$. Now,*

$$\begin{aligned} P(\{X = n\}) &= \sum_{(w_1, w_2, \ldots, w_n) \in S_n} P(\{(w_1, w_2, \ldots, w_n)\}) \\ &= P(\{(w_1, w_2, \ldots, w_n)\})|S_n| \\ &= P(\{w_1\})P(\{w_2\}) \cdots P(\{w_n\})|S_n| \quad \textit{(since all events are independent)} \\ &= \frac{1}{4}\left(\frac{3}{4}\right)^{n-2} \frac{1}{4}\binom{n-1}{1} \\ &= \frac{n-1}{16}\left(\frac{3}{4}\right)^{n-2}. \end{aligned}$$

Also, $\sum_{n=2}^{\infty} P(\{X = n\}) = 1$. Thus, X is of discrete type with support $E_X = \{2, 3, 4, \ldots\}$ and p.m.f.

$$f_X(x) = \begin{cases} \frac{x-1}{16}\left(\frac{3}{4}\right)^{x-2}, \text{ if } x \in \{2, 3, 4, \ldots\}, \\ 0, \ \text{ otherwise.} \end{cases}$$

The d.f. of X is

$$\begin{aligned} F_X(x) &= P(\{X \leq x\}) \\ &= \begin{cases} 0, \ \text{ if } x < 2, \\ \frac{1}{16}\sum_{j=2}^{i}(j-1)(\frac{3}{4})^{j-2}, \ \text{ if } i \leq x < i+1, i = 2, 3, 4, \ldots, \end{cases} \\ &= \begin{cases} 0, \ \text{ if } x < 2, \\ 1 - \frac{i+3}{4}(\frac{3}{4})^{i-2}, \ \text{ if } i \leq x < i+1, i = 2, 3, 4, \ldots. \end{cases} \end{aligned}$$

Definition 9. *A random variable X with the d.f. F_X is said to be of continuous type, or simply continuous, if there exists an integrable function $f_X : \mathbb{R} \to \mathbb{R}$ such that $f_X(x) \geq 0$ for every $x \in \mathbb{R}$ and*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \ x \in \mathbb{R}.$$

The function f_X is called the probability density function (p.d.f.) of random variable X and the set $E_X = \{x \in \mathbb{R} : f_X(x) > 0\}$ is called the support of random variable X (or of p.d.f. f_X.

Remark 10. *Let X be a continuous type RV with the support E_X, the d.f. F_X and a p.d.f. f_X.*

(1) $\lim_{x\to\infty} F_X(x) = F_X(\infty) = \int_{-\infty}^{\infty} f_X(t)dt = 1.$

(2) *F_X is continuous on $\mathbb{R}$. (Prove!) Therefore, $P(\{X = x\}) = 0\ \forall\ x \in \mathbb{R}$. In general, for any countable set C, $P(\{X \in X\}) = 0$.*
(3) *Let $a, b \in \mathbb{R}$ with $a < b$. Then*

$$P(\{a < X \leq b\}) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt.$$

In general, for any $B \in \mathbb{B}_{\mathbb{R}}$, we have $P(\{X \in B\}) = \int\limits_{-\infty}^{\infty} f_X(t)I_B(t)dt$, where I_B is the indicator function of B.

Remark 11. (1) *Suppose that the d.f. F_X of an RV X is differentiable at every $x \in \mathbb{R}$. Then*

$$F_X(x) = \int_{-\infty}^{x} F'_X(t)dt,\ x \in \mathbb{R}.$$

This implies that X is of continuous type and we may take its p.d.f to be $f_X(x) = F'_X(x)$, $x \in \mathbb{R}$.
(2) *Suppose that the d.f. F_X of an RV X is differentiable everywhere except on a countable set C. Further suppose that*

$$\int_{-\infty}^{\infty} F'_X(t)I_{C^c}dt = 1.$$

This again will imply that X is of continuous type with p.d.f. (Verify!)

$$f_X(x) = \begin{cases} F'_X(x), \text{if } x \notin C, \\ a_x, \text{if } x \in C, \end{cases}$$

where a_x, $x \in C$ are arbitrary non negative constants.
(3) *From the previous remark, it is clear that p.d.f. of a continuous random variable need not be unique.*
(4) *There are random variables that are neither of discrete type nor of continuous type. Find some examples.*

Example 12. *Let X be an RV having d.f.*

$$F_X(x) = \begin{cases} 0, \text{if } x < 0, \\ 1 - e^{-x}, \text{if } x \geq 0. \end{cases}$$

We observe that F_X is differentiable everywhere except at $x = 0$. Let $C = \{0\}$. Moreover,

$$\int_{-\infty}^{\infty} F'_X(t)I_{C^c}dt = \int_0^{\infty} e^{-t} = 1.$$

Hence, X is of continuous type and its p.d.f. is

$$f_X(x) = \begin{cases} 0, \text{if } x \leq 0, \\ e^{-x}, \text{if } x > 0. \end{cases}$$