MODULE 2

RANDOM VARIABLE AND ITS DISTRIBUTION

LECTURE 9

Topics

2.4 TYPES OF RANDOM VARIABLES: DISCRETE, CONTINUOUS AND ABSOLUTELY CONTINUOUS

2.4 TYPES OF RANDOM VARIABLES: DISCRETE, CONTINUOUS AND ABSOLUTELY CONTINUOUS

Let X be a r.v. defined on a probability space $(\Omega, \mathcal{F}, P)$ and let $(\mathbb{R}, \mathcal{B}_1, P_X)$ be the probability space induced by X. Let F_X be the d.f. of X. Then F_X will either be continuous everywhere or it will have countable number of discontinuities. Moreover the sum of sizes of jumps at the point of discontinuities of F_X will be either 1 or less than 1. These properties can be used to classify a r.v. into three broad categories.

Definition 4.1

A random variable X is said to be of *discrete type* if there exists a non-empty and countable set S_X such that $P(\{X = x\}) = F_X(x) - F_X(x-) > 0, \forall x \in S_X$ and $P_X(S_X) = \sum_{x \in S_X} P(\{X = x\}) = \sum_{x \in S_X} [F_X(x) - F_X(x-)] = 1$. The set S_X is called the *support* of the discrete random variable X. ■

Remark 4.1

If a r.v. X is of discrete type then $P_X(S_X^C) = 1 - P_X(S_X) = 0$ and, consequently $P(\{X = x\}) = 0, \forall x \in S_X^C$, i.e., $F_X(x) - F_X(x-) = 0,\ \forall x \in S_X^C$ and F_X is continuous at every point of S_X^C. Moreover, $F_X(x) - F_X(x-) = P(\{X = x\}) > 0,\ \forall x \in S_X$. It follows that the support S_X of a discrete type r.v. X is nothing but the set of discontinuity points of the d.f. F_X. Moreover the sum of sizes of jumps at the point of discontinuities is

$$\sum_{x \in S_x} [F_X(x) - F_X(x-)] = \sum_{x \in S_x} P(\{X = x\}) = P_X(S_X) = 1. ■$$

Thus we have the following theorem.

Theorem 4.1

Let X be a random variable with distribution function F_X and let D_X be the set of discontinuity points of F_X. Then X is of discrete type if, and only if, $P(\{X \in D_X\}) = 1$. ■

Definition 4.2

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

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

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

Example 4.1

Let us consider a r.v. Z having the d.f. F considered in Example 3.2 (iii). The set of discontinuity points of F is $D_Z = \{0, 2, 3, 6, 12, 15\}$ and

$$P(\{Z \in D_Z\}) = \sum_{z \in D_Z} [F(z) - F(z-)] = 1.$$

Therefore the r.v. Z is of discrete type with support $S_Z = D_Z = \{0, 2, 3, 6, 12, 15\}$ and p.m.f.

$$f_Z(z) = \begin{cases} [F(z) - F(z-)], & \text{if } \; z \in S_Z \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{8}, & \text{if } \; z \in \{0, 2, 15\} \\ \frac{1}{4}, & \text{if } \; z = 3 \\ \frac{3}{10}, & \text{if } \; z = 6 \\ \frac{3}{40}, & \text{if } \; z = 12 \\ 0, & \text{otherwise} \end{cases} .$$ ■

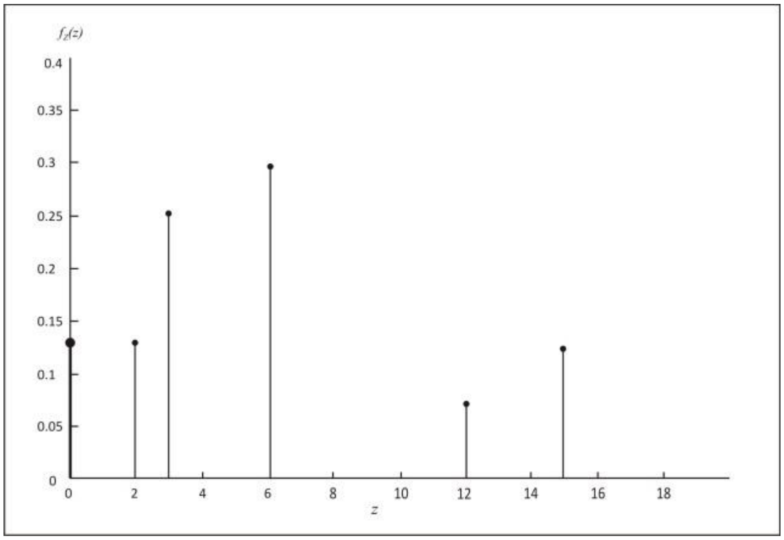

Figure 4.1. Plot of p.m.f. $f_Z(z)$

Note that the p.m.f. f_X of a discrete type r.v. X, having support S_X, satisfies the following properties:

(i) $f_X(x) > 0,\ \forall x \in S_X$ and $f_X(x) = 0, \forall x \notin S_X$, (4.1)

(ii) $\sum_{x \in S_X} f_X(x) = \sum_{x \in S_X} P(\{X = x\}) = 1.$ (4.2)

Moreover, for $B \in \mathcal{B}_1$,

$$P_X(B) = P_X(B \cap S_X) + P_X(B \cap S_X^C)$$

$$= P_X(B \cap S_X) \qquad \text{(since } B \cap S_X^C \subseteq S_X^C \text{ and } P_X(S_X^C) = 0\text{)}$$

$$= \sum_{x \in B \cap S_X} f_X(x).$$

This suggests that we can study probability space $(\mathbb{R}, \mathcal{B}_1, P_X)$, induced by a discrete type r.v. X, through the study of its p.m.f. f_X. Also

$$F_X(x) = \sum_{y \in (-\infty, x] \cap S_X} f_X(y), \quad x \in \mathbb{R}$$

and

$$f_X(x) = P(\{X = x\}) = F_X(x) - F_X(x-), \quad x \in \mathbb{R}.$$

Thus, given a p.m.f. of a discrete type of r.v., we can get its d.f. and vice-versa. In other words, there is one-one correspondence between p.m.f.s and distribution functions of discrete type random variables.

The following theorem establishes that any function $g: \mathbb{R} \to \mathbb{R}$ satisfying (4.1) and (4.2) is p.m.f. of some discrete type random variable.

Theorem 4.2

Suppose that there exists a non-empty and countable set $S \subseteq \mathbb{R}$ and a function $g: \mathbb{R} \to \mathbb{R}$ satisfying: (i) $g(x) > 0, \forall x \in S$; (ii)$g(x) = 0,\ \forall x \notin S$, and (iii) $\sum_{x \in S} g(x) = 1$. Then there exists a discrete type random variable on some probability space $(\mathbb{R}, \mathcal{B}_1, P)$ such that the p.m.f. of X is g.

Proof. Define the set function $P: \mathcal{B}_1 \to \mathbb{R}$ by

$$P(B) = \sum_{x \in B \cap S} g(x), \qquad B \in \mathcal{B}_1.$$

It is easy to verify that P is a probability measure on $\mathcal{B}_1$, i.e., $(\mathbb{R}, \mathcal{B}_1, P)$is a probability space. Define $X: \mathbb{R} \to \mathbb{R}$ by $X(\omega) = \omega, \omega \in \mathbb{R}$. Clearly X is a r.v. on the probability space $(\mathbb{R}, \mathcal{B}_1, P)$ and it induces the same probability space $(\mathbb{R}, \mathcal{B}_1, P)$. Clearly $P(\{X = x\}) = g(x),\ x \in \mathbb{R}$, and $\sum_{x \in S} g(x) = 1$. Therefore the r.v. X is of discrete type with support S and p.m.f.g. ■

Example 4.2

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 $x \notin \{2, 3, 4, \cdots\}$. For $i \in \{2, ,3 ,4 \ldots\}$

$$P(\{X = i\}) = \left(\binom{i-1}{1} \frac{1}{4} \left(\frac{3}{4}\right)^{i-2}\right) \frac{1}{4}$$

$$= \frac{i-1}{16} \left(\frac{3}{4}\right)^{i-2}.$$

Moreover, $\sum_{i=2}^{\infty} P(\{X = i\}) = 1$. It follows that X is a discrete type r.v. with support $S_X = \{2, 3, 4, \ldots\}$ and p.m.f.

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

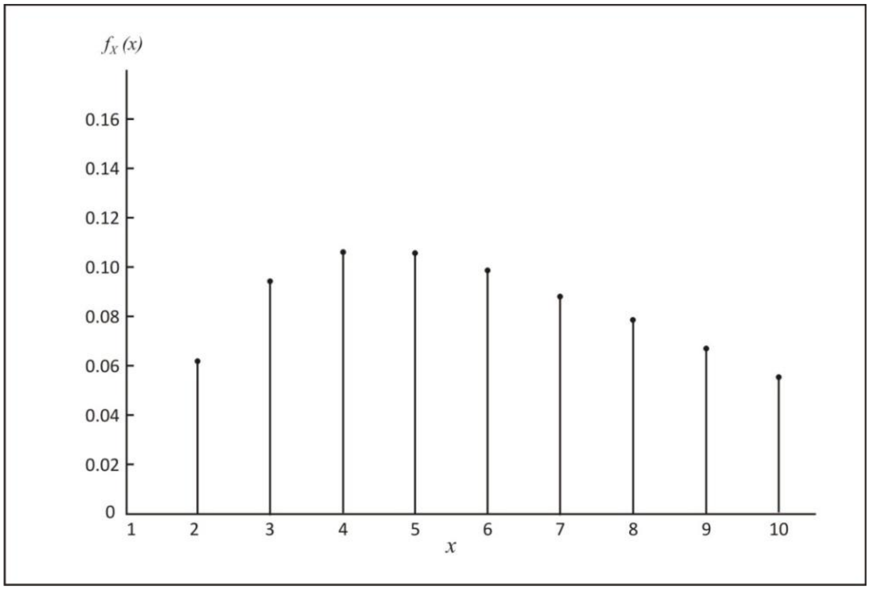

Figure 4.2. Plot of p.m.f. $f_X(x)$

The d.f. of X is

$$F_X(x) = P(\{X \le x\})$$

$$= \begin{cases} 0, & \text{if } x < 2 \\ \frac{1}{16}\sum_{j=2}^{i}(j-1)\left(\frac{3}{4}\right)^{j-2}, & \text{if } i \le x < i+1, \quad i = 2,3,4,\cdots \end{cases}$$

$$= \begin{cases} 0, & \text{if } x < 2 \\ 1 - \frac{i+3}{4}\left(\frac{3}{4}\right)^{i-1}, & \text{if } i \le x < i+1, \quad i = 2,3,4,\cdots \end{cases} \blacksquare$$

Example 4.3

A r.v. X has the d.f.

$$F_X(x) = \begin{cases} 0, & \text{if} \quad x < 2 \\ \frac{2}{3}, & \text{if} \quad 2 \le x < 5 \\ \frac{7-6k}{6}, & \text{if} \quad 5 \le x < 9 \\ \frac{3k^2-6k+7}{6}, & \text{if} \quad 9 \le x < 14 \\ \frac{16k^2-16k+19}{16}, & \text{if} \quad 14 \le x \le 20 \\ 1, & \text{if} \quad x > 20 \end{cases},$$

where $k \in \mathbb{R}$.

(i) Find the value of constant k;
(ii) Show that the r.v. X is of discrete type and find its support;
(iii) Find the p.m.f. of X.

Solution. (i) Since F_X is right continuous, we have

$$F_X(20) = F_X(20+)$$

$$\Rightarrow \quad 16k^2 - 16k + 3 = 0$$

$$\Rightarrow k = \frac{1}{4} \text{ or } k = \frac{3}{4}. \tag{4.3}$$

Also F_X is non-decreasing. Therefore

$$F_X(5-) \le F_X(5)$$

$$\Rightarrow k \le \frac{1}{2}. \tag{4.4}$$

On combining (4.3) and (4.4) we get $k = 1/4$. Therefore

$$F_X(x) = \begin{cases} 0, & \text{if} \quad x < 2 \\ \frac{2}{3}, & \text{if} \quad 2 \le x < 5 \\ \frac{11}{12}, & \text{if} \quad 5 \le x < 9 \\ \frac{91}{96}, & \text{if} \quad 9 \le x < 14 \\ 1, & \text{if} \quad x \ge 14 \end{cases}.$$

(ii) The set of discontinuity points of F_X is $D_X = \{2, 5, 9, 14\}$. Moreover

$$P(\{X = 2\}) = F_X(2) - F_X(2-) = \frac{2}{3},$$
$$P(\{X = 5\}) = F_X(5) - F_X(5-) = \frac{1}{4},$$
$$P(\{X = 9\}) = F_X(9) - F_X(9-) = \frac{1}{32},$$
$$P(\{X = 14\}) = F_X(14) - F_X(14-) = \frac{5}{96},$$

and

$$P(\{X \in D_X\}) = P(\{X = 2\}) + P(\{X = 5\}) + P(\{X = 9\}) + P(\{X = 14\})$$
$$= 1.$$

Therefore the r.v. X is of discrete type with support $S_X = \{2, 5, 9, 14\}$.

(iii) Clearly the p.m.f. of X is given by

$$f_X(x) = P(\{X = x\}) = \begin{cases} \frac{2}{3}, & \text{if } x = 2 \\ \frac{1}{4}, & \text{if } x = 5 \\ \frac{1}{32}, & \text{if } x = 9 \\ \frac{5}{96}, & \text{if } x = 14 \\ 0, & \text{otherwise} \end{cases}. \blacksquare$$

Example 4.4

A r.v. X has the p.m.f.

$$f_X(x) = \begin{cases} \frac{c}{(2x-1)(2x+1)}, & \text{if} \quad x \in \{1, 2, 3, \cdots\} \\ 0, & \text{otherwise} \end{cases},$$

where $c \in \mathbb{R}$.

(i) Find the value of constant c;

(ii) For positive integers m and n, such that $m < n$, evaluate $P(\{X < m + 1\}), P(\{X \geq m\}), P(\{m \leq X < n\})$ and $P(\{m < X \leq n\})$;

(iii) Determine the d.f. of X.

Solution.

(i) Let S_X be the support of X so that $S_X = \{x \in \mathbb{R}: f_X(x) > 0\}$ and $\sum_{x \in S_X} f_X(x) = 1$. Clearly, S_X=1, 2, 3, … and

$$\sum_{i=1}^{\infty} \frac{c}{(2i-1)(2i+1)} = 1$$

$$\Rightarrow \lim_{n\to\infty} \sum_{i=1}^{n} \frac{c}{(2i-1)(2i+1)} = 1$$

$$\Rightarrow \frac{c}{2} \lim_{n\to\infty} \sum_{i=1}^{n} \left[\frac{1}{2i-1} - \frac{1}{2i+1}\right] = 1$$

$$\Rightarrow \frac{c}{2}\lim_{n\to\infty}\left[\sum_{i=1}^{n}\frac{1}{2i-1}-\sum_{i=1}^{n}\frac{1}{2i+1}\right]=1$$

$$\Rightarrow \frac{c}{2}\lim_{n\to\infty}\left[1-\frac{1}{2n+1}\right]=1$$

$$\Rightarrow \quad \mathrm{c}=2.$$

(ii) For positive integers m and n such that $m<n$, we have

$$P(\{X<m+1\})=P(\{X\le m\})$$

$$=\sum_{i=1}^{m}\frac{2}{(2i-1)(2i+1)}$$

$$=\sum_{i=1}^{m}\left[\frac{1}{2i-1}-\frac{1}{2i+1}\right]$$

$$=1-\frac{1}{2m+1}$$

$$=\frac{2m}{2m+1},$$

$$P(\{X\ge m\})=1-P(\{X<m\})$$

$$=1-\frac{2(m-1)}{2(m-1)+1}$$

$$=\frac{1}{2m-1},$$

$$P(\{m\le X<n\})=P(\{X<n\})-P(\{X<m\})$$

$$=\frac{2(n-1)}{2n-1}-\frac{2(m-1)}{2m-1}$$

$$=\frac{2(n-m)}{(2n-1)(2m-1)},$$

and
$$P(\{m<X\le n\})=P(\{m+1\le X<n+1\})$$

$$=\frac{2(n-m)}{(2n+1)(2m+1)}.$$

(iii) Clearly, for $x<1, F_X(x)=0$. For $i\le x<i+1, i=1,2,3,\cdots$

$$F_X(x) = P(\{X < i + 1\}) = \frac{2i}{2i + 1}. \qquad \text{(using (ii))}$$

Therefore,

$$F_X(x) = \begin{cases} 0, & \text{if } \; x < 1 \\ \dfrac{2i}{2i+1}, & \text{if } \; i \le x < i + 1, \;\; i = 1, 2, 3, \cdots \end{cases}. \quad \blacksquare$$