MODULE 2

RANDOM VARIABLE AND ITS DISTRIBUTION

LECTURE 8

Topics

2.3 DISTRIBUTION FUNCTION AND ITS PROPERTIES

2.3 DISTRIBUTION FUNCTION AND ITS PROPERTIES

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X: \Omega \rightarrow \mathbb{R}$ be a r.v. so that $X^{-1}((-\infty, \mathrm{a}]) = \{\omega \in \mathbb{R}: X(\omega) \leq a\} \in \mathcal{F}, \forall a \in \mathbb{R}$. Throughout we will use the following notation:

$$\{\text{a statement (say S) about } X\} = \{\omega \in \Omega: \text{statement S holds}\};$$

e. g.,

$$\{a < X \leq b\} \stackrel{\text{def}}{=} \{\omega \in \Omega: a < X(\omega) \leq b\} \stackrel{\text{def}}{=} X^{-1}((\mathrm{a}, \mathrm{b}]), -\infty \leq a < b < \infty$$

$$\{X = c\} \stackrel{\text{def}}{=} \{\omega \in \Omega:\ X(\omega) = c\} \stackrel{\text{def}}{=} X^{-1}(\{\mathrm{c}\}),\ \ \mathrm{c} \in \mathbb{R},$$
$$\{X \in B\} \stackrel{\text{def}}{=} \{\omega \in \Omega:\ X(\omega) \in B\} \stackrel{\text{def}}{=} X^{-1}(B),\ \ B \in \mathcal{B}_1,$$

$$\{X \leq \mathrm{c}\} \stackrel{\text{def}}{=} \{\omega \in \Omega: X(\omega) \leq c\} \stackrel{\text{def}}{=} X^{-1}((-\infty, \mathrm{c}]),\ \ \mathrm{c} \in \mathbb{R}.$$

Definition 3.1

The function $F_X: \mathbb{R} \rightarrow \mathbb{R}$, defined by,

$$F_X(x) = P(\{X \leq x\}) = P_X((-\infty, x]),\ \ x \in \mathbb{R},$$

is called the *distribution function* (d.f.) of random variable X. ■

Example 3.1

(i) Let us revisit Example 2.1 (i). The induced probability space is $(\mathbb{R}, \mathcal{B}_1, P_X)$, where $P_X(\{0\}) = P_X(\{3\}) = \frac{1}{8}, P_X(\{1\}) = P_X(\{2\}) = \frac{3}{8}$ and

$$\begin{aligned} P_X(B) &= P(\{X \in B\}) \\ &= \sum_{i \in \{0,1,2,3\} \cap B} P_X(\{i\}),\ \ B \in \mathcal{B}_1. \end{aligned}$$

Clearly, for $x \in \mathbb{R}$,

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

$$= P_X((-\infty, x])$$

$$= \sum_{i \in \{0,1,2,3\} \cap (-\infty, x]} P_X(\{i\})$$

$$= \begin{cases} 0, & \text{if} \quad x < 0 \\ \frac{1}{8}, & \text{if} \ 0 \leq x < 1 \\ \frac{1}{2}, & \text{if} \ 1 \leq x < 2 \\ \frac{7}{8}, & \text{if} \ 2 \leq x < 3 \\ 1, & \text{if} \quad x \geq 3 \end{cases}$$

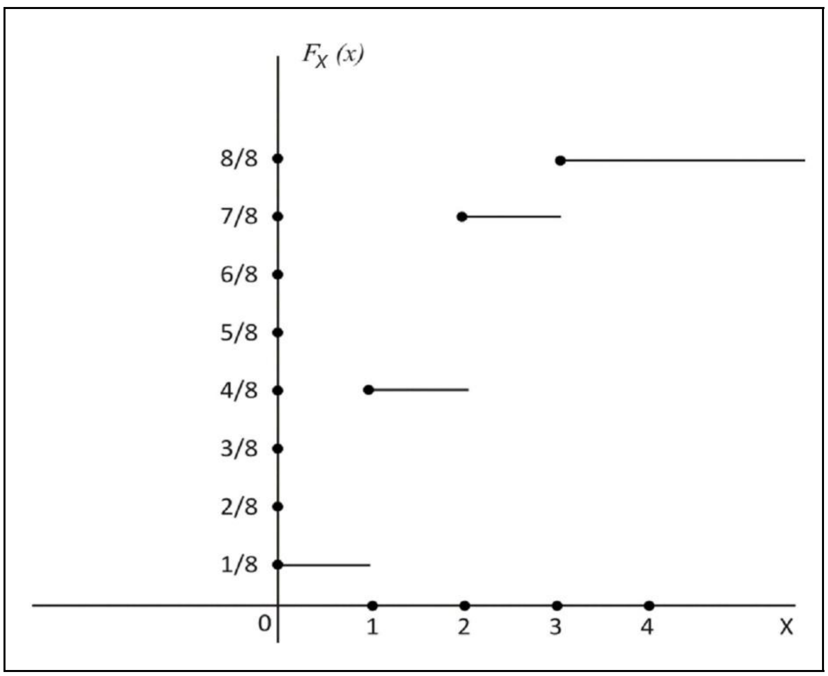

Figure 3.1. Plot of distribution function $F_X(x)$

Note that $F_X(x)$ is non-decreasing, right continuous, $F_X(-\infty) \stackrel{\text{def}}{=} \lim_{x \to -\infty} F_X(x) = 0$ and $F_X(\infty) \stackrel{\text{def}}{=} \lim_{x \to \infty} F_X(x) = 1$. Moreover $F_X(x)$ is a step function having discontinuities at points $0, 1, 2$ and 3.

(ii) Consider Example 2.1 (ii). The probability space induced by r.v. X is $(\mathbb{R}, \mathcal{B}_1, P_X)$, where, for $B \in \mathcal{B}_1$,

$$P_X(B) = 2\int_0^\infty z\, e^{-z^2} I_B(z)dz.$$

Therefore,

$$F_X(x) = P(\{X \leq x\})$$
$$= P_X\big((-\infty, x]\big)$$
$$= 2\int_0^\infty z\, e^{-z^2} I_{(-\infty,x]}(z)dz, \qquad x \in \mathbb{R}.$$

Clearly, for $x < 0$, $F_X(x) = 0$. For $x \geq 0$

$$F_X(x) = 2\int_0^x z\, e^{-z^2} dz = 1 - e^{-x^2}.$$

Thus,

$$F_X(x) = \begin{cases} 0, & \text{if} \quad x < 0 \\ 1 - e^{-x^2}, & \text{if} \quad x \geq 0 \end{cases}.$$

Note that $F_X(x)$ is non-decreasing, continuous, $F_X(-\infty) = \lim_{x\to-\infty} F_X(x) = 0$ and $F_X(\infty) = \lim_{x\to\infty} F_X(x) = 1$. ■

Now we will derive various properties of a distribution function. The following lemma, whose proof is immediate and can be found in any standard text book on calculus, will be useful in studying the properties of a distribution function.

Lemma 3.1

Let $-\infty \leq a < b \leq \infty$ and let $f: (a, b) \to \mathbb{R}$ be a non-decreasing function

$(\text{i.e.}, f(s) \leq f(t), \forall a < s < t < b)$.

Then

(i) for all $x \in (a, b]$ and $y \in [a, b)$, $f(x-)$ and $f(y+)$ exist;
(ii) for all $x \in (a, b)$, $f(x-) \leq f(x) \leq f(x+)$;
(iii) for $a < x < y < b$, $f(x+) \leq f(y-)$;
(iv) f has at most countable number of discontinuities;

where $f(c-)$ and $f(c+)$ denote, respectively, the left hand and right hand limits of the function f at point $c \in (a, b)$. ■

Theorem 3.1

Let F_X be the distribution function of a random variable X. Then

(i) F_X is non-decreasing;
(ii) F_X is right continuous;
(iii) $F_X(-\infty) \stackrel{\text{def}}{=} \lim_{x\to-\infty} F_X(x) = 0$ and $F_X(\infty) \stackrel{\text{def}}{=} \lim_{x\to\infty} F_X(x) = 1$.

Proof.

(i) Let $-\infty < x < y < \infty$. Then $(-\infty, x] \subseteq (-\infty, y]$ and therefore, using monotonicity of probability measures, we get

$$F_X(x) = P_X\big((-\infty, x]\big) \le P_X((-\infty, y]) = F_X(y).$$

(ii) Fix $x \in \mathbb{R}$. Since F_X is non-decreasing, it follows from Lemma 3.1 that $F_X(x+)$ exists. Therefore

$$F_X(x+) = \lim_{n\to\infty} F_X\left(x + \frac{1}{n}\right) = \lim_{n\to\infty} P_X\left((-\infty, x + \frac{1}{n}]\right).$$

Note that $\left(-\infty, x + \frac{1}{n}\right] \downarrow$and $\mathrm{Lim}_{n\to\infty}\left(-\infty, x+\frac{1}{n}\right] = \bigcap_{n=1}^{\infty}\left(-\infty, x + \frac{1}{n}\right] = (-\infty, x]$. Now using continuity of probability measures (Theorem 4.1, Module 1) we have

$$F_X(x+) = \lim_{n\to\infty} P_X\left((-\infty, x + \frac{1}{n}]\right)$$

$$= P_X\left(\mathrm{Lim}_{n\to\infty}\, (-\infty, x + \frac{1}{n}]\right)$$

$$= P_X((-\infty, x])$$

$$= F_X(x).$$

(iii) Using standard arguments of calculus it follows that $F_X(-\infty) = \lim_{n\to\infty} F_X(-n)$ and $F_X(\infty) = \lim_{n\to\infty} F_X(n)$, where limits are taken along the sequence $\{n: n = 1, 2, \cdots\}$. Note that $(-\infty, -n] \downarrow$, $(-\infty, n] \uparrow$, $\mathrm{Lim}_{n\to\infty}(-\infty, -n] = \bigcap_{n=1}^{\infty}(-\infty, -n] = \phi$ and $\mathrm{Lim}_{n\to\infty}(-\infty, n] = \bigcup_{n=1}^{\infty}(-\infty, n] = \mathbb{R}$. Again using the continuity of probability measures, we get

$$F_X(-\infty) = \lim_{n\to\infty} F_X(-n) = \lim_{n\to\infty} P_X\big((-\infty, -n]\big) = P_X\left(\mathrm{Lim}_{n\to\infty}(-\infty, -n]\right) = P_X(\phi) = 0,$$

and

$$F_X(\infty) = \lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} P_X\big((-\infty, n]\big) = P_X\left(\mathrm{Lim}_{n\to\infty}\,(-\infty, n]\right) = P_X(\mathbb{R}) = 1. \blacksquare$$

Remark 3.1

(i) Using Lemma-3.1 (i)-(ii) and Theorem 3.1 (i) it follows that for a d.f. F_X, $F_X(x+)$ and $F_X(x-)$ exist for every $x \in \mathbb{R}$ and F_X is discontinuous at $x \in \mathbb{R}$ if and only if $F_X(x-) < F_X(x+) = F_X(x)$. Consequently a d.f. has only jump discontinuities (a discontinuity point $x \in \mathbb{R}$ of F_X is called a *jump discontinuity* if $F_X(x+)$ and $F_X(x-)$ exist but $F_X(x-) = F_X(x+) = F_X(x)$ does not hold). Moreover the size of the jump at a point $x \in \mathbb{R}$ of discontinuity is $p_x = F_X(x) - F_X(x-)$.

(ii) Using Lemma 3.1 (iv) and Theorem 3.1 (i) it follows that any d.f. F_X has atmost countable number of discontinuities.

(iii) Let $a \in \mathbb{R}$. Since $(-\infty, \mathrm{a} - \frac{1}{\mathrm{n}}] \uparrow$ and $\mathrm{Lim}_{n\to\infty}\ (-\infty, \mathrm{a} - \frac{1}{\mathrm{n}}] = \cup_{n=1}^{\infty}\ (-\infty, \mathrm{a} - \frac{1}{\mathrm{n}}] = (-\infty, a)$, the continuity of probability measures implies

$$
\begin{aligned}
P(\{X < a\}) &= P_X\big((-\infty, a)\big) \\
&= P_X\left(\underset{n\to\infty}{\mathrm{Lim}}(-\infty, \mathrm{a} - \frac{1}{\mathrm{n}}]\right) \\
&= \lim_{n\to\infty} P_X\left((-\infty, \mathrm{a} - \frac{1}{\mathrm{n}}]\right) \\
&= \lim_{n\to\infty} F_X\left(a - \frac{1}{n}\right) \\
&= F_X(a-).
\end{aligned}
$$

Therefore,

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

Also,

$$F_X(x) = F_X(x+) \le F_X(y-), \forall -\infty < x < y < \infty \text{ (using Lemma 3.1 (iii))}$$

and

$$P(\{X = x\}) = P(\{X \le x\}) - P(\{X < x\}) = F_X(x) - F_X(x-), \forall x \in \mathbb{R}.$$

Thus F_X is continuous (discontinuous) at a point $x \in \mathbb{R}$ if, and only if, $P(\{X = x\}) = 0$ $(P(\{X = x\}) > 0)$.

(iv) Let D_X denote the set of discontinuity points (jump points) of d.f. F_X. Then D_X is a countable set and

$$\sum_{x \in D_X} [F_X(x) - F_X(x-)] = \sum_{x \in D_X} P(\{X = x\}) = P_X(D_X) \leq 1,$$

i.e., the sum of sizes of jumps of a d.f. does not exceed 1.

(v) Let $-\infty < a < b < \infty$. Then

$$P(\{a < X \leq b\}) = P(\{X \leq b\}) - P(\{X \leq a\}) = F_X(b) - F_X(a)$$
$$P(\{a < X < b\}) = P(\{X < b\}) - P(\{X \leq a\}) = F_X(b-) - F_X(a)$$
$$P(\{a \leq X < b\}) = P(\{X < b\}) - P(\{X < a\}) = F_X(b-) - F_X(a-)$$
$$P(\{a \leq X \leq b\}) = P(\{X \leq b\}) - P(\{X < a\}) = F_X(b) - F_X(a-),$$

$$P(\{X \geq a\}) = 1 - P(\{X < a\}) = 1 - F_X(a-),$$

and

$$P(\{X > a\}) = 1 - P(\{X \leq a\}) = 1 - F_X(a). \blacksquare$$

We state the following theorem without providing the proof. The theorem essentially states that any function $G: \mathbb{R} \to \mathbb{R}$ that is non-decreasing and right continuous with $G(-\infty) = \lim_{x \to -\infty} G(x) = 0$ and $G(\infty) = \lim_{x \to \infty} G(x) = 1$ can be regarded as d.f. of a random variable.

Theorem 3.2

Let $G: \mathbb{R} \to \mathbb{R}$ be a non-decreasing and right continuous function for which $G(-\infty) = 0$ and $G(\infty) = 1$. Then there exists a random variable X defined on a probability space $(\Omega, \mathcal{F}, P)$ such that the distribution function of X is G. $\blacksquare$

Example 3.2

(i) Consider a function $G: \mathbb{R} \to \mathbb{R}$, defined by,

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

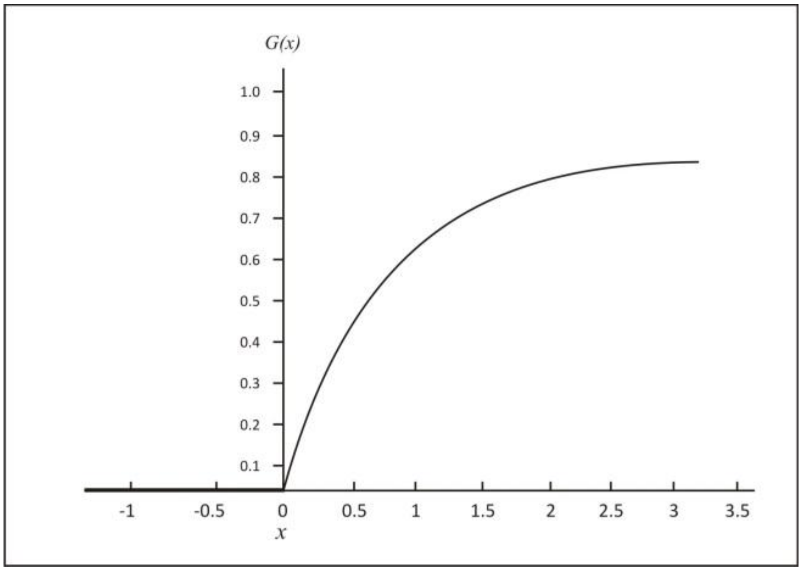

Figure 3.2. Plot of distribution function $G(x)$

Clearly G is non-decreasing, continuous and satisfies $G(-\infty) = 0$ and $G(\infty) = 1$. Therefore G can be treated as d.f. of some r.v., say X. Since G is continuous we have

$$P(\{X = x\}) = G(x) - G(x-) = 0, \forall x \in \mathbb{R},$$

and, for $-\infty < a < b < \infty$,

$$P(\{a < X < b\}) = P(\{a \leq X < b\}) = \mathrm{P}(a \leq X \leq b) = P(\{a < X \leq b\})$$
$$= G(b) - \mathrm{G}(a).$$

Moreover, for $-\infty < a < \infty$,

$$P(\{X \geq a\}) = P(\{X > a\}) = 1 - G(a)$$

and

$$P(\{X < a\}) = P(\{X \leq a\}) = G(a).$$

In particular

$$P(\{2 < X \leq 3\}) = G(3) - G(2) = e^{-2} - e^{-3};$$

$$P(\{-2 < X \leq 3\}) = G(3) - G(-2) = 1 - e^{-3};$$

$$P(\{1 \leq X < 4\}) = G(4) - G(1) = e^{-1} - e^{-4};$$

$$P(\{5 \leq X < 8\}) = G(8) - G(5) = e^{-5} - e^{-8};$$

$$P(\{X \geq 2\}) = 1 - G(2) - e^{-2};$$

and

$$P(\{X > 5\}) = 1 - G(5) = e^{-5}.$$

Note that the sum of sizes of jumps of G is 0.

(ii) Let $H: \mathbb{R} \to \mathbb{R}$ be given by

$$H(x) = \begin{cases} 0, & \text{if} \quad x < 0 \\ \frac{x}{4}, & \text{if} \quad 0 \leq x < 1 \\ \frac{x}{3} & \text{if} \quad 1 \leq x < 2 \\ \frac{3x}{8} & \text{if} \quad 2 \leq x < \frac{5}{2} \\ 1, & \text{if} \quad x \geq \frac{5}{2} \end{cases}.$$

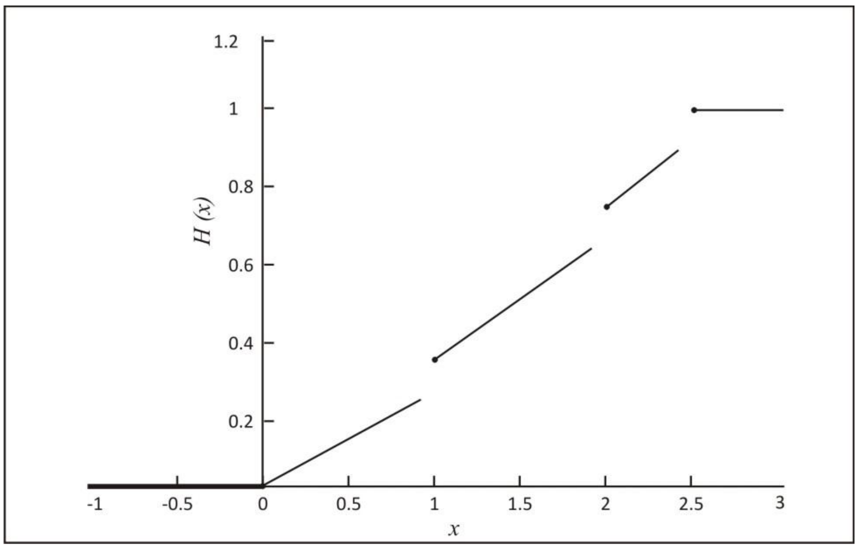

Figure 3.3. Plot of distribution function $H(x)$

Clearly H is non-decreasing, right continuous and satisfies $H(-\infty) = 0$ and $H(\infty) = 1$. Therefore H can be treated as d.f. of some r.v., say Y. $H$is continuous everywhere except at points $1, 2,$ and $5/2$where it has jump discontinuities with jumps of sizes $P(\{Y = 1\}) = H(1) - H(1-) = \frac{1}{12}$, $P(\{Y = 2\}) = H(2) - H(2-) = 1/12$ and $P(\{Y = 5/2\}) = H(5/2) - H(5/2-) = 1/16$. Moreover for $x \in \mathbb{R} - \{1, 2, 5/2\}$, $P(\{Y = x\}) = 0$. We also have

$$P\left(\left\{1 < Y \leq \frac{5}{2}\right\}\right) = H\left(\frac{5}{2}\right) - H(1) = 1 - \frac{1}{3} = \frac{2}{3};$$

$$P\left(\left\{1 < Y < \frac{5}{2}\right\}\right) = H\left(\frac{5}{2} -\right) - H(1) = \frac{15}{16} - \frac{1}{3} = \frac{29}{48};$$

$$P\left(\left\{1 \leq Y < \frac{5}{2}\right\}\right) = H\left(\frac{5}{2} -\right) - H(1 -) = \frac{15}{16} - \frac{1}{4} = \frac{11}{16};$$

$$P(\{-2 \leq Y < 1\}) = H(1 -) - H(-2 -) = \frac{1}{4} - 0 = \frac{1}{4};$$

$$P(\{Y \geq 2\}) = 1 - H(2 -) = 1 - \frac{2}{3} = \frac{1}{3};$$

and $$P(\{Y > 2\}) = 1 - H(2) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Note that sum of sizes of jumps of H is $11/48 \in (0, 1)$.

(iii) Let $F: \mathbb{R} \to \mathbb{R}$ be given by

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{8}, & \text{if } 0 \leq x < 2 \\ \frac{1}{4} & \text{if } 2 \leq x < 3 \\ \frac{1}{2} & \text{if } 3 \leq x < 6 \\ \frac{4}{5}, & \text{if } 6 \leq x < 12 \\ \frac{7}{8}, & \text{if } 12 \leq x < 15 \\ 1, & \text{if } x \geq 15 \end{cases}.$$

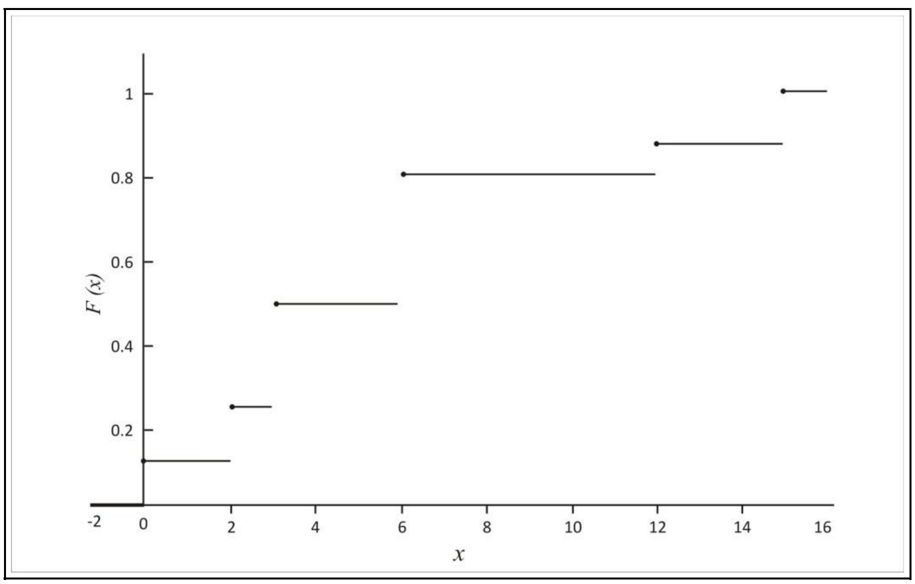

Figure 3.4. Plot of distribution function $F(x)$

As F is non-decreasing and right continuous with $F(-\infty) = 0$ and $F(\infty) = 1$, it can be regarded as d.f. of some r.v., say Z. Clearly, except at points $0, 2, 3, 6, 12$ and 15, F is continuous at all other points and at discontinuity points $0, 2, 3, 6$,12 and 15 it has jump discontinuities with jumps of sizes

$$P(\{Z = 0\}) = F(0) - F(0\,-) = \frac{1}{8},$$
$$P(\{Z = 2\}) = F(2) - F(2\,-) = \frac{1}{8},$$
$$P(\{Z = 3\}) = F(3) - F(3\,-) = \frac{1}{4},$$
$$P(\{Z = 6\}) = F(6) - F(6\,-) = \frac{3}{10},$$
$$P(\{Z = 12\}) = F(12) - F(12\,-) = \frac{3}{40},$$

and

$$P(\{Z = 15\}) = F(15) - F(15\,-) = \frac{1}{8}.$$

Moreover $P(\{Z = x\}) = F(x) = F(x\,-) = 0,\ \forall x \in \mathbb{R} - \{0, 2, 3, 6, 12, 15\}$. Note that in this case sum of sizes of jumps of F is 1. ■

Remark 3.2

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. In advanced courses on probability theory it is shown that the d.f. F_X uniquely determines the induced probability measure P_X and vice-versa. Thus to study the induced probability space $(\mathbb{R}, \mathcal{B}_1, P_X)$ it suffices to study the d.f. F_X. ■