Random Variables and probability Distributions-I

Mr. Anup Singh
Department of Mathematics
Mahatma Gandhi Central university
Motihari-845401, Bihar, India
E-mail: anup.singh254@gmail.com

Random Variable: A random variable is a function that associates a real number with each element in the sample space.

In other words, a random variable is a function $X: S \to R$, where S is the sample space of the random experiment under consideration.

Note: Normally a capital letter, say X, is used to denote a random variable and its corresponding small letter, x in this case, for one of its values.

Example: Consider the random experiment of tossing a coin three times and observing the result (a Head or a Tail) for each toss. Let X denote the total number of heads obtained in the three tosses of the coin.

(i) Construct a table that shows the values of the random variable X for each possible outcome of the random experiment.

(ii) Identify the event $\{X \leq 1\}$ in words.

Let Y denote the difference between the number of heads obtained and the number of tails obtained.

(iii) Construct a table showing the value of Y for each possible outcome.

(iv) Identify the event $\{Y = 0\}$ in words.

Discrete Random Variable: If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers (countable), it is called a discrete sample space. A random variable is called a discrete random variable if its set of possible outcomes is countable.

Example: (i) Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. The possible outcomes and the values x of the random variable X, where X is the number of red balls, are

Sample Space	**x**
RR	**2**
RB	**1**
BR	**1**
BB	**0**

Example (ii) Suppose that our experiment consists of tossing 3 fair coins. If we let X denote the number of heads appearing, then X is a random variable taking on one of the values 0, 1, 2, 3 with respective probabilities

S = {(H, H, H), (H, H, T), (H, T, H), (T, H, H),

(T, T, H), (T, H, T), (H, T, T), (T, T, T)}

P{X = 0} = P{(T, T, T)} = 1/8

P{X = 1} = P{(T, T, H), (T, H, T), (H, T, T)} = 3/8

P{X = 2} = P{(T, H, H), (H, T, H),(H, H, T) }= 3/8

P{X = 3} = P{(H, H, H)} = 1/8

Discrete Probability Distributions: A discrete random variable assumes each of its values with a certain probability.

In the case of tossing a coin three times, the variable X, representing the number of heads, assumes the value 2 with probability 3/8, since 3 of the 8 equally likely sample points result in two heads and one tail. The possible values x of X and their probabilities are

X	0	1	2	3
P(X=x)	1/8	3/8	3/8	1/8

Probability Mass Function: Let X be a one dimensional discrete random variable which takes the values $x_1, x_2, x_3, \ldots$. Then $P(X = x_i) = P(x_i)$ satisfies the following conditions

1. $P(x_i) \geq 0$
2. $$\sum_{i=1}^{\infty} P(x_i) = 1$$

Cumulative Distribution Function of Discrete Random Variable X:

The distribution function of a discrete random variable X defined in (–∞,∞) is given by

$$F(x) = P(X \le x) = \sum_{t \le x} f(t), \quad -\infty < x < \infty$$

Properties of the Distribution function

1. $P(a < X \le b) = F(b) - F(a)$
2. $P(a \le X \le b) = P(X = a) + F(b) - F(a)$
3. $P(a < X < b) = F(b) - F(a) - P(X = b)$
4. $P(a \le X < b) = F(b) - F(a) - P(X = b) + P(X = a)$

Example(i) A random variable X has the following probability function

Value of X, x_i	0	1	2	3	4	5	6	7	8
Probability P(x)	a	3a	5a	7a	9a	11a	13a	15a	17a

1. Determine the value of 'a'.
2. Find P(X < 3), P(X ≥ 3). P(0 < X < 5).
3. Find the distribution function of X.

1. Since $\sum_{i=1}^{\infty} P(x_i) = 1$

$$a + 3a + 5a + 7a + 9a + 11a + 13a + 15a + 17a = 1$$

$$a = \frac{1}{81}$$

2. $P(X < 3) = P(0) + P(1) + P(2) = a + 3a + 5a = 9a = 1/9$

$P(X \geq 3) = 1 - P(X < 3) = 8/9$

$P(0 < X < 5) = P(1) + P(2) + P(3) + P(4)$

$= 3a + 5a + 7a + 9a = 24a = 24/81$

3.

x	0	1	2	3	4	5	6	7	8
P(x)	*a*	*3a*	*5a*	*7a*	*9a*	*11a*	*13a*	*15a*	*17a*
F(x)	*a*	*4a*	*9a*	*16a*	*25a*	*36a*	*49a*	*64a*	*81a*

Alternate method for sub-division 2, using the cumulative distribution function F(x).

$P(X < 3) = P(X \leq 2) = F(2) = 9a = 1/9$

$P(X \geq 3) = 1 - P(X < 3) = 1 - (1/9) = 8/9$

$P(0 < X < 5) = F(5) - F(0) - P(X = 5)$

$= 36a - a - 11a$

$= 24a$

$= 24/81$

Example (ii) A random variable X has the following probability function

Value of X, x_i	0	1	2	3	4	5	6	7
Probability P(x)	0	a	$2a$	$2a$	$3a$	a^2	$2a^2$	$7a^2 + a$

1. Determine the value of 'a'.
2. Find P(1.5 < X < 4.5 / X > 2).
3. Find the smallest value of λ for which $P(X \leq \lambda) > 1/2$.

1. Since $\sum_{i=1}^{\infty} P(x_i) = 1$

$10a^2 + 9a = 1$

$a = 1/10$ or -1. As $a = -1$ is meaningless, $a = 1/10$

2. $$P(1.5 < X < 4.5 / X > 2) = \frac{P[(1.5 < X < 4.5) \cap (X > 2)]}{P(X > 2)}$$

$$= \frac{P(X = 3) + P(X = 4)}{1 - [P(X = 0) + P(X = 1) + P(X = 2)} = \frac{5}{7}$$

3. $P(X \le 0) = 0$; $P(X \le 1) = 0.1$; $P(X \le 2) = 0.3$;
$P(X \le 3) = 0.5$ and $P(X \le 4) = 0.8$

$\therefore \lambda = 4$ for which $P(X \le \lambda) > 1/2$.

Continuous Random Variable: If a random variable takes on all values within a certain interval, then the random variable is called Continuous random variable.

E.g., The height, age and weight of individuals, the amount of rainfall on a rainy day.

Probability Density Function: If X is a continuous random variable then $f(x)$ is called the probability density function of X provided $f(x)$ satisfies the following conditions;

1. $f(x) \geq 0, \forall x$

2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Cumulative Probability Distribution of Continuous Random Variable X: The **cumulative distribution function** F(x) of a continuous random variable X with density function f(x) is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt, \; -\infty < x < \infty$$

Results:

(a) 1. $P(a \le x \le b) = \int_a^b f(x)dx = F(b) - F(a)$

2. When X is continuous r.v.

$$P(X = a) = P(a \le X \le a) = \int_a^a f(x)dx = 0$$

$\therefore$ $P(a < X \le b) = P(a \le X \le b) = P(a < X < b) = P(a \le X < b)$.

(b) If F(x) is the distribution function of one dimensional random variables, then

1. $0 \le F(x) \le 1$
2. If $x < y$, then $F(x) \le F(y)$
3. $F(-\infty) = 0$, $F(\infty) = 1$.
4. If X is discrete r.v. taking values $x_1, x_2, x_3, \ldots$ where $x_1 < x_2 < x_3 < \ldots$ then $P(X = x_i) = F(x_i) - F(x_{i-1})$.
5. If X is continuous r.v., then $\frac{dF(x)}{dx} = f(x)$

Example If the density function of a continuous r.v. X is given by

$$f(x) = \begin{cases} ax, & 0 \le x \le 1 \\ a, & 1 \le x \le 2 \\ 3a - ax & 2 \le x \le 3 \\ 0, & elsewhere \end{cases}$$

1. Find the value of *a*
2. Find the cumulative distribution function of X
3. Find $P(1.5 < X \le 3)$
4. Find $P(X > 1.5)$

Solution: 1. Since $f(x)$ is a p.d.f.

$$\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_0^3 f(x)dx = \int_0^1 ax\,dx + \int_1^2 a\,dx + \int_2^3 (3a - ax)dx = 1$$

$a = ½$.

2. $F(x) = P(X \le x) = 0, \quad x < 0$

$$F(x) = \int_0^x ax\,dx = \left[\frac{ax^2}{2}\right]_0^x = \frac{ax^2}{2} = \frac{x^2}{4}, \quad 0 \le x \le 1$$

$$F(x) = \int_0^1 a\,x dx + \int_1^x a\,dx$$

$$= \left[\frac{ax^2}{2}\right]_0^1 + [ax]_1^x$$

$$= \left[\frac{a}{2}\right] + [ax - a] = ax - \frac{a}{2} = \frac{x}{2} - \frac{1}{4}, \quad 1 \le x \le 2$$

$$F(x)=\int_0^1 a\,x\,dx+\int_1^2 a\,dx+\int_2^3 (3a-ax)\,dx$$

$$=\left[\frac{ax^2}{2}\right]_0^1+\left[ax\right]_1^2+\left[3ax-\frac{ax^2}{2}\right]_2^x$$

$$=\left[\frac{a}{2}\right]+a+\left[\left(3ax-\frac{ax^2}{2}\right)-\left(6a-2a\right)\right]$$

$$=3ax-\frac{ax^2}{2}-\frac{5a}{2}$$

$$=\frac{1}{4}\left(6x-x^2-5\right),\ \ 2\le x\le 3$$

$$F(x)=P(X\le x)=1,\ \ x>3$$

2. Cumulative Distribution function

$$F(x) = P(X \le x) = \begin{cases} 0, & x < 0 \\ \dfrac{x^2}{4}, & 0 \le x \le 1 \\ \dfrac{1}{4} + \dfrac{x-1}{2}, & 1 \le x \le 2 \\ \dfrac{-5}{4} + \dfrac{6x - x^2}{4}, & 2 \le x \le 3 \\ 1, & x > 3 \end{cases}$$

3.

$$F(1.5) = \frac{1}{4} + \frac{x-1}{2} = \frac{1}{4} + \frac{1.5-1}{2} = \frac{1}{2}, \quad 1 \le x \le 2$$

$$F(3) = \frac{-5}{4} + \frac{6x - x^2}{4} = -\frac{5}{4} + \frac{6(3)-(3)^2}{4} = 1, \;\; 2 \le x \le 3$$

$$\therefore P(1.5 < x < 3) = F(3) - F(1.5) = 1 - \frac{1}{2} = \frac{1}{2}$$

4.

$$\mathrm{P}(\mathrm{X} > 1.5) = \int_{1.5}^{3} f(x)dx$$

$$= \int_{1.5}^{2} \frac{1}{2}dx + \int_{2}^{3} \left(\frac{3}{2} - \frac{x}{2}\right)dx$$

$$= \frac{1}{2}$$

Reference Books:

1. Erwin Kreyszig, Advance Engineering Mathematics, 9th Edition, John Wiley & Sons, 2006.
2. Sheldon Ross, A first course in Probability, 8th Edition, Pearson Education India.
3. W. Feller An Introduction to Probability Theory and its Applications, Vol. 1, 3rd Edition, Wiley, 1968.
4. S. C. Gupta and V. . Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand & Sons.
5. P. G. Hoel, S. C. Port and C. J. Stone, Introduction to Probability Theory, Universal Book Stall, 2003.
6. A. M. Mood, F. A. Graybill and D. C. Bose, Introduction to Theory of Statistics. 3rd Edition, Tata McGraw-Hill Publication.

THANK YOU