UNIT 1 (i)

Random variables and its Characterization

Random Variable

Definition

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

Example1

For example, the sample space giving a detailed description of each possible outcome when three electronic components are tested may be written

$$S = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\},$$

where N denotes nondefective and D denotes defective.

the *random variable* X, the number of defective items *assigned a numerical value* of 0, 1, 2, or 3.

Example2

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 y of the random variable Y, where Y is the number of red balls, are

Sample Space	y
RR	2
RB	1
BR	1
BB	0

Example3

Statisticians use **sampling plans** to either accept or reject batches or lots of material. Suppose one of these sampling plans involves sampling independently 10 items from a lot of 100 items in which 12 are defective.

Let X be the random variable defined as the number of items found defective in the sample of 10. In this case, the random variable takes on the values $0, 1, 2, \ldots, 9, 10$.

Example4

Interest centers around the proportion of people who respond to a certain mail order solicitation. Let X be that proportion. X is a random variable that takes on all values x for which $0 \leq x \leq 1$.

Example5

Let X be the random variable defined by the waiting time, in hours, between successive speeders spotted by a radar unit. The random variable X takes on all values x for which $x \geq 0$.

Definition

If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers, it is called a **discrete sample space**.

i.e. X may be assumed as $x_1, x_{2,\ldots\ldots,}\ x_{n},\ldots\ldots$

Definition

If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a **continuous sample space**.

X is random variable that can take all values in an interval

Example: the length of time during which a vacuum tube installed in a circuit function is continuous RV

Discrete Probability Distributions

a random variable X $f(x) = P(X = x)$:

The set of ordered pairs $(x, f(x))$ is called the **probability function**, **probability mass function**, or **probability distribution** of the discrete random variable X.

It can also be expressed as collection of pair $\{(x_i, p_i)\}$ i=1,2,3……

Example6

the case of tossing a coin three times, the variable X representing the number of heads

S={ HHH,HHT,HTH,THH,HTT,THT,TTH,TTT}

Then

x:	0	1	2	3
P(x):	1/8	3/8	3/8	1/8

Definition 3

The set of ordered pairs $(x, f(x))$ is a **probability function**, **probability mass function**, or **probability distribution** of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

Example7

A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Solution: Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x can only take the numbers 0, 1, and 2

$$f(0) = P(X = 0) = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = \frac{68}{95}, \quad f(1) = P(X = 1) = \frac{\binom{3}{1}\binom{17}{1}}{\binom{20}{2}} = \frac{51}{190},$$

$$f(2) = P(X = 2) = \frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}} = \frac{3}{190}.$$

Thus, the probability distribution of X is

x	0	1	2
$f(x)$	$\frac{68}{95}$	$\frac{51}{190}$	$\frac{3}{190}$

Definition

The **cumulative distribution function** $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty.$$

Example8

Find the cumulative distribution function of the random variable X in Example 3.9. Using $F(x)$, verify that $f(2) = 3/8$.

***Solution*:** Direct calculations of the probability distribution of Example 3.9 give $f(0)= 1/16$, $f(1) = 1/4$, $f(2)= 3/8$, $f(3)= 1/4$, and $f(4)= 1/16$. Therefore,

$$\begin{aligned}
F(0) &= f(0) = \frac{1}{16}, \\
F(1) &= f(0) + f(1) = \frac{5}{16}, \\
F(2) &= f(0) + f(1) + f(2) = \frac{11}{16}, \\
F(3) &= f(0) + f(1) + f(2) + f(3) = \frac{15}{16}, \\
F(4) &= f(0) + f(1) + f(2) + f(3) + f(4) = 1.
\end{aligned}$$

Hence,

$$F(x) = \begin{cases} 0, & \text{for } x < 0, \\ \frac{1}{16}, & \text{for } 0 \leq x < 1, \\ \frac{5}{16}, & \text{for } 1 \leq x < 2, \\ \frac{11}{16}, & \text{for } 2 \leq x < 3, \\ \frac{15}{16}, & \text{for } 3 \leq x < 4, \\ 1 & \text{for } x \geq 4. \end{cases}$$

$$f(2) = F(2) - F(1) = \frac{11}{16} - \frac{5}{16} = \frac{3}{8}.$$

Continuous Probability Distributions

Definition

The function $f(x)$ is a **probability density function** (pdf) for the continuous random variable X, defined over the set of real numbers, if

1. $f(x) \geq 0$, for all $x \in R$.
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$.
3. $P(a < X < b) = \int_a^b f(x)\, dx$.

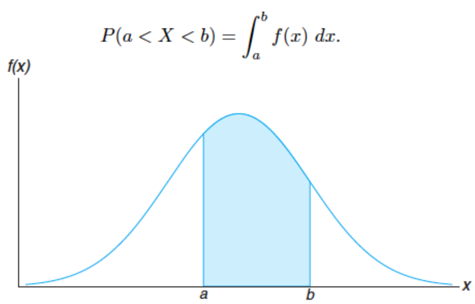

Figure 3.5: $P(a < X < b)$.

Definition

The **cumulative distribution function** $F(x)$ of a continuous random variable X with density function $f(x)$ is

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

Example9

Suppose that the error in the reaction temperature, in °C, for a controlled laboratory experiment is a continuous random variable X having the probability density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

.

(a) Verify that $f(x)$ is a density function.

(b) Find $P(0 < X \leq 1)$.

(c) find $F(x)$ and use it to evaluate $P(0 < X \leq 1)$

Sol.

a)

$f(x) \geq 0.$

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-1}^{2} \frac{x^2}{3} dx = \frac{x^3}{9} |_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1.$$

b)

$$P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \left. \frac{x^3}{9} \right|_0^1 = \frac{1}{9}.$$

c)

For $-1 < x < 2$,

$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-1}^{x} \frac{t^2}{3} dt = \left. \frac{t^3}{9} \right|_{-1}^{x} = \frac{x^3+1}{9}.$$

$$F(x) = \begin{cases} 0, & x < -1, \\ \frac{x^3+1}{9}, & -1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

Using above F(x)

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9},$$

Example10

Determine the value c so that a probability distribution of the discrete random variable X

$f(x) = c(x^2 + 4)$, for $x = 0, 1, 2, 3$

Soln:-

$c = 1/30$ since $1 = \sum_{x=0}^{3} c(x^2 + 4) = 30c.$

Example11

The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{20,000}{(x+100)^3}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that a bottle of this medicine will have a shell life of

(a) at least 200 days;

(b) anywhere from 80 to 120 days.

Soln:

(a) $P(X > 200) = \int_{200}^{\infty} \frac{20000}{(x+100)^3}\, dx = - \frac{10000}{(x+100)^2}\Big|_{200}^{\infty} = \frac{1}{9}.$

(b) $P(80 < X < 200) = \int_{80}^{120} \frac{20000}{(x+100)^3}\, dx = - \frac{10000}{(x+100)^2}\Big|_{80}^{120} = \frac{1000}{9801} = 0.1020.$

Example12

The cdf is given as

$$F(t) = \begin{cases} 0, & t < 1, \\ \frac{1}{4}, & 1 \le t < 3, \\ \frac{1}{2}, & 3 \le t < 5, \\ \frac{3}{4}, & 5 \le t < 7, \\ 1, & t \ge 7, \end{cases}$$

find

(a) $P(T = 5)$;

(b) $P(T > 3)$;

(c) $P(1.4 < T < 6)$;

(d) $P(T \le 5 \mid T \ge 2)$.

Soln:

(a) $P(T=5)=F(5)-F(4)=3/4-1/2=1/4$.

(b) $P(T>3)=1-F(3)=1-1/2=1/2$.

(c) $P(1.4<T<6)=F(6)-F(1.4)=3/4-1/4=1/2$.

(d) $P(T\leq 5|T\geq 2)=\frac{P(2\leq T\leq 5)}{P(T\geq 2)}=\frac{3/4-1/4}{1-1/4}=\frac{2}{3}$.

Example13

The waiting time, in hours, between successive speeders spotted by a radar unit is a continuous random variable with cumulative distribution function

$$F(x)=\begin{cases}0, & x<0,\\ 1-e^{-8x}, & x\geq 0.\end{cases}$$

Find the probability of waiting less than 12 minutes between successive speeders

(a) using the cumulative distribution function of X;

(b) using the probability density function of X.

Soln:

(a) $P(X<0.2)=F(0.2)=1-e^{-1.6}=0.7981$;

(b) $f(x)=F'(x)=8e^{-8x}$. Therefore, $P(X<0.2)=8\int_0^{0.2}e^{-8x}\,dx=-e^{-8x}\big|_0^{0.2}=0.7981$.