variable ... formal definition may be given as follows.

Definition: A random variable (abbreviatively RV) is a function that assigns a real number $X(s)$ to every element $s \in S$, where S is the sample space corresponding to a random experiment E.)

> **Note** *Although we are expected to perform the random experiment E, we observe the outcome $s \in S$ and then evaluate $X(s)$ [i.e., assign a real number x to $X(s)$], the number $x = X(s)$ itself can be thought of as the outcome of the experiment and R_x as the sample space of the experiment. In this sense, we will hereafter talk about a random variable X taking the value x and $P(X = x)$. Actually, $P(X = x) = P\{s: X(s) = x\}$.*

Hereafter, R_x will be referred to as ***Range space***.

Similarly $\{X \leq x\}$ represents the subset $\{s: X(s) \leq x\}$ and hence an event associated with the experiment.

Discrete Random Variable

If X is a random variable (RV) which can take a finite number or countably infinite number of values, X is called a discrete RV. When the RV is discrete, the possible values of X may be assumed as $x_1, x_2, \ldots, x_n, \ldots$. In the finite case, the list of values terminates and in the countably infinite case, the list goes upto infinity.

For example, the number shown when a die is thrown and the number of alpha particles emitted by a radioactive source are discrete RVs.

Probability Function

If X is a discrete RV which can take the values $x_1, x_2, x_3, \ldots$ such that $P(X = x_i) = p_i$, then p_i is called the *probability function or probability mass function or point probability function*, provided p_i ($i = 1, 2, 3, \ldots$) satisfy the following conditions:

(i) $p_i \geq 0$, for all i, and

(ii) $\sum_i p_i = 1$

The collection of pairs $\{x_i, p_i\}$, $i = 1, 2, 3, \ldots$, is called *the probability distribution of the* RV X, which is sometimes displayed in the form of a table as given below:

$X = x_i$	$P(X = x_i)$
x_1	p_1
x_2	p_2
$\vdots$	$\vdots$
x_r	p_r
$\vdots$	$\vdots$

Continuous Random Variable

If X is an RV which can take all values (i.e., *infinite number* of values) in an interval, then X is called a *continuous* RV.

For example, the length of time during which a vacuum tube installed in a circuit functions is a continuous RV.

Probability Density Function

If X is a continuous RV such that

$$P\left\{x - \frac{1}{2}\mathrm{d}x \leq X \leq x + \frac{1}{2}\mathrm{d}x\right\} = f(x)\mathrm{d}x$$

then $f(x)$ is called the *probability density function* (shortly denoted as pdf) of X, provided $f(x)$ satisfies the following conditions:

(i) $f(x) \geq 0$, for all $x \in R_x$, and

(ii) $\int_{R_X} f(x)\mathrm{d}x = 1$

Moreover, $P(a \leq X \leq b)$ or $P(a < X < b)$ is defined as

$$P(a \le X \le b) = \int_a^b f(x)\mathrm{d}x.$$

The curve $y = f(x)$ is called the *probability curve of the* RV X.

> **Note** *When X is a continuous RV*
>
> $$P(X = a) = P(a \le X \le a) = \int_a^a f(x)\,\mathrm{d}x = 0$$
>
> *This means that it is almost impossible that a continuous RV assumes a specific value. Hence $P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b)$.*

Cumulative Distribution Function (cdf)

If X is an R V, discrete or continuous, then $P(X \le x)$ is called the *cumulative distribution function* of X or *distribution function* of X and denoted as $F(x)$.

If X is discrete,

$$F(x) = \sum_{\substack{j \\ X_j \le x}} p_j$$

If X is continuous,

$$F(x) = P(-\infty < X \le x) = \int_{-\infty}^{x} f(x)\,\mathrm{d}x$$

Properties of the *cdf* $F(x)$

1. $F(x)$ is a non-decreasing function of x, i.e., if $x_1 < x_2$, then $F(x_1) \le F(x_2)$.
2. $F(-\infty) = 0$ and $F(\infty) = 1$.
3. If X is a discrete R V taking values $x_1, x_2, \ldots$, where $x_1 < x_2 < x_3 < \ldots < x_{i-1} < x_i < \ldots$, then $P(X = x_i) = F(x_i) - F(x_{i-1})$.
4. If X is a continuous R V, then $\dfrac{\mathrm{d}}{\mathrm{d}x} F(x) = f(x)$, at all points where $F(x)$ is differentiable.

> **Note** *Although we may talk of probability distribution of a continuous RV, it cannot be represented by a table as in the case of a discrete RV. The probability distribution of a continuous RV is said to be known, if either its pdf or cdf is given.*

Special Distributions

The probability mass functions of some discrete RVs and the probability density functions of some continuous RVs, which are of frequent applications, are as follows:

Discrete Distributions

1. If the discrete RV X can take the values 0, 1, 2, ..., n, such that $P(X = i) = nC_i\, p^i\, q^{n-i}$, $i = 0, 1, \ldots, n$, where $p + q = 1$, then X is said to follow binomial distribution with parameters n and p, which is denoted a $B(n, p)$
2. If the discrete RV X can take the values 0, 1, 2, ..., such that $P(x = i) = \dfrac{e^{-\lambda}\, \lambda^i}{i!}$, $i = 0, 1, 2, \ldots$, then X is said to follow a Poisson distribution with parameter λ.
3. If the discrete RV X can take the values 0, 1, 2, ..., such that $P(X = i) = (n + i - 1)C_i\, p^n q^i$, $i = 0, 1, 2, \ldots$, where $p + q = 1$, then X is said to follow a *Pascal (or negative binomial) distribution* with parameter n.
4. A Pascal distribution with parameter 1 [i.e., $P(X = i) = pq^i$, $i = 0, 1, 2, \ldots$ and $p + q = 1$] is called a *geometric distribution*.

Continuous Distributions

5. If the pdf of a continuous RV X is $f(x) = \dfrac{1}{b - a}$ (a constant), $a \le x \le b$, then X follows a *uniform distribution* (*or rectangular distribution*).
6. If the pdf of a continuous RV X is $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2}$, $-\infty < x < \infty$, then X is said to follow a *normal distribution* (or *Gaussian distribution*) with parameters μ and σ, which will be hereafter denoted as $N(\mu, \sigma)$.
7. If the pdf of a continuous RV X is $f(x) = \dfrac{1}{\Gamma(n)}\, e^{-x}\, x^{n-1}$, $0 < x < \infty$ and $n > 0$, then X follows a *gamma distribution* with parameter n. Gamma distribution is a particular case of *Erlang distribution*, the pdf of which is $f(x) = \dfrac{c^n}{\Gamma(n)}\, x^{n-1} e^{-cx}$, $0 < x < \infty$, $n > 0$, $c > 0$.
8. An Erlang distribution with $n = 1$ [i.e., $f(x) = ce^{-cx}$, $0 < x < \infty$, $c > 0$] is called *an exponential* (or *negative exponential*) *distribution* with parameter c.
9. If the pdf of a continuous RV X is $f(x) = \dfrac{x}{\alpha^2}\, e^{-x^2/2\alpha^2}$, $0 < x < \infty$, then X follows a *Rayleigh distribution* with parameter α.
10. If the pdf of a continuous RV X is $f(x) = \dfrac{\sqrt{2}}{\alpha^3\sqrt{\pi}}\, x^2 e^{-x^2/2\alpha^2}$, $0 < x < \infty$, then X follows a *Maxwell distribution* with parameter α.

Example 5

A random variable X has the following probability distribution.

x:	-2	-1	0	1	2	3
$p(x)$:	0.1	K	0.2	$2K$	0.3	$3K$

(a) Find K, (b) Evaluate $P(X < 2)$ and $P(-2 < X < 2)$, (c) find the cdf of X and (d) evaluate the mean of X. (BU — Apr. 96)

(a) Since $\Sigma P(x) = 1$, $6K + 0.6 = 1$

$$\therefore \qquad K = \frac{1}{15}$$

$\therefore$ the probability distribution becomes

x :	-2	-1	0	1	2	3
$p(x)$:	$1/10$	$1/15$	$1/5$	$2/15$	$3/10$	$1/5$

(b) $P(X < 2) \qquad = P(X = -2, -1, 0 \text{ or } 1)$

$$= P(X = -2) + P(X = -1) + P(X = 0) + P(X = 1)$$

[since the events $(X = -2)$, $(X = -1)$ etc. are mutually exclusive]

$$= \frac{1}{10} + \frac{1}{15} + \frac{1}{5} + \frac{2}{15} = \frac{1}{2}$$

$$P(-2 < X < 2) = P(X = -1, 0 \text{ or } 1)$$

$$= P(X = -1) + P(X = 0) + P(X = 1)$$

$$= \frac{1}{15} + \frac{1}{5} + \frac{2}{15} = \frac{2}{5}$$

$\frac{1}{10} + \frac{1}{15} = \frac{3+2}{30} = \frac{5}{30} = \frac{1}{6}$

(c) $F(x) = 0$, when $x < -2$

$$= \frac{1}{10}, \text{ when } -2 \le x < -1$$

$$= \frac{1}{6}, \text{ when } -1 \le x < 0$$

$$= \frac{11}{30}, \text{ when } 0 \le x < 1$$

$$= \frac{1}{2}, \text{ when } 1 \le x < 2$$

$$= \frac{4}{5}, \text{ when } 2 \le x < 3$$

$$= 1, \text{ when } 3 \le x$$

(d) The mean of X is defined as $E(X) = \Sigma\, xp(x)$ (refer to Chapter 4)

$$\therefore \text{Mean of } X = \left(-2 \times \frac{1}{10}\right) + \left(-1 \times \frac{1}{15}\right) + \left(0 \times \frac{1}{5}\right)$$

$$+ \left(1 \times \frac{2}{15}\right) + \left(2 \times \frac{3}{10}\right) + \left(3 \times \frac{1}{5}\right)$$

$$= -\frac{1}{5} - \frac{1}{15} + \frac{2}{15} + \frac{3}{5} + \frac{3}{5} = \frac{16}{15}$$

Example 6

The probability function of an infinite discrete distribution is given by $P(X = j) = 1/2^j$ ($j = 1, 2, \ldots, \infty$). Verify that the total probability is 1 and find the mean and variance of the distribution. Find also $P(X$ is even), $P(X \ge 5)$ and $P(X$ is divisible by 3).

Let $P(X = j) = p_j$

$$\sum_{j=1}^{\infty} p_j = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots \infty, \text{ that is a geometric series.}$$

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

The mean of X is defined as $E(X) = \sum_{j=1}^{\infty} jp_j$ (refer to Chapter 4).

$$\therefore \quad E(X) = a + 2a^2 + 3a^3 + \ldots \infty, \text{ where } a = \frac{1}{2}$$

$$= a(1 + 2a + 3a^2 + \ldots \infty)$$

$$= a\,(1-a)^{-2} = \frac{\frac{1}{2}}{\left(\frac{1}{2}\right)^2} = 2$$

he variance of X is defined as $V(X) = E(X^2) - [E(X)]^2$,

here $E(X^2) = \sum_{j=1}^{\infty} j^2 p^j$ (refer to Chapter 4).

$$E(X^2) = \sum_{j=1}^{\infty} j^2 a^j, \text{ where } a = \frac{1}{2}$$

$$= \sum_{j=1}^{\infty} [j(j+1) - j]a^j = \sum_{j=1}^{\infty} j(j+1)a^j - \sum_{j=1}^{\infty} ja^j$$

$$= a\,(1.2 + 2.3a + 3.4a^2 + \ldots \infty) - a(1 + 2a + 3a^3 + \ldots \infty)$$

$$= a \times 2\,(1-a)^3 - a \times (1-a)^{-2}$$

$$= \frac{2\,a}{(1-a)^3} - \frac{a}{(1-a)^2} = 8 - 2 = 6$$

$$\therefore \qquad V(X) = E(X^2) - \{E(X)\}^2 = 6 - 4 = 2$$

$P(X \text{ is even}) = P(X = 2 \text{ or } X = 4 \text{ or } X = 6 \text{ or etc.})$

$P\,(X = 2) + P(X = 4) + \ldots + \infty$

(since the events are mutually exclusive)

$$= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^6 + \cdots + \infty$$

$$= \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{1}{3}$$

$$P(X \geq 5) = P(X = 5 \text{ or } X = 6 \text{ or } X = 7 \text{ or etc.})$$

$$= P(X = 5) + P(X = 6) + \ldots + \infty$$

$$= \frac{\frac{1}{2^5}}{1-\frac{1}{2}} = \frac{1}{16}$$

" " "

Conditional Probability

The conditional probability of an event B, assuming that the event A has happened, is denoted by $P(B/A)$ and defined as

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \text{ provided } P(A) \neq 0$$

For example, when a fair die is tossed, the conditional probability of getting '1', given that an odd number has been obtained, is equal to 1/3 as explained below:

$$S = \{1, 2, 3, 4, 5, 6\};\ A = \{1, 3, 5\};\ B = \{1\}$$

$$\therefore \quad P(B/A) = \frac{n(A \cap B)}{n(A)} = \frac{1}{3}$$

As per the definition given above, $P(B/A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{1/6}{1/2} = \dfrac{1}{3}$.

Rewriting the definition of conditional probability, we get $P(A \cap B) = P(A) \times P(B/A)$. This is sometimes referred to as *Product theorem of probability*, which is proved as follows.

Let n_A, n_{AB} be the number of cases favourable to the events A and $A \cap B$, out of the total number n of cases.

$$\therefore \quad P(A \cap B) = \frac{n_{AB}}{n} = \frac{n_A}{n} \times \frac{n_{AB}}{n_A} = P(A) \times P(B/A)$$

The product theorem can be extended to 3 events A, B and C as follows:

$$P(A \cap B \cap C) = P(A) \times P(B/A) \times P(C/A \text{ and } B)$$

The following properties are easily deduced from the definition of conditional probability:

1. If $A \subset B$, $P(B/A) = 1$, since $A \cap B = A$
2. If $B \subset A$, $P(B/A) \geq P(B)$, since $A \cap B = B$, and $\dfrac{P(B)}{P(A)} \geq P(B)$, as $P(A) \leq P(S) = 1$
3. If A and B are mutually exclusive events, $P(B/A) = 0$, since $P(A \cap B) = 0$
4. If $P(A) > P(B)$, $P(A/B) > P(B/A)$ (MKU — Apr. 96)
5. If $A_1 \subset A_2$, $P(A_1/B) \leq P(A_2/B)$ (BU — Apr. 96))

Independent Events

A set of events is said to be independent if the occurrence of any one of them does not depend on the occurrence or non-occurrence of the others.

When 2 events A and B are independent, it is obvious from the definition that $P(B/A) = P(B)$. If the events A and B are independent, the product theorem takes the form $P(A \cap B) = P(A) \times P(B)$. Conversely, if $P(A \cap B) = P(A) \times P(B)$, the events A and B are said to be independent (pairwise independent). The product theorem can be extended to any number of independent events: If $A_1, A_2, \ldots, A_n$ are n independent events.

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \times P(A_2) \times \ldots \times P(A_n)$$

When this condition is satisfied, the events $A_1, A_2, \ldots, A_n$ are also said to be *totally independent*. A set of events $A_1, A_2, \ldots, A_n$ is said to be *mutually independent* if the events are totally independent when considered in sets of 2, 3, ..., n events.

In the case of more than 2 events, the term 'independence' is taken as 'total independence' unless specified otherwise.

draws the white ball.

Theorem of Total Probability

If $B_1, B_2, \ldots, B_n$ be a set of exhaustive and mutually exclusive events, and A is another event associated with (or caused by) B_i, then

$$P(A) = \sum_{i=1}^{n} P(B_i)\, P(A/B_i)$$

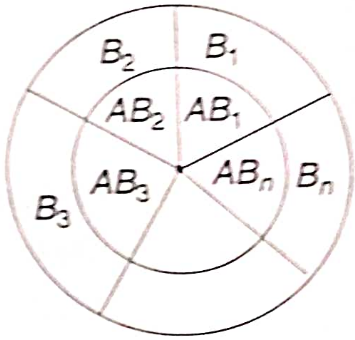

Fig. 1.3

Proof

The inner circle represents the event A. A can occur along with (or due to) $B_1, B_2, \ldots, B_n$ that are exhaustive and mutually exclusive.

$\therefore$ $AB_1, AB_2, \ldots, AB_n$ are also mutually exclusive, such that

$$A = AB_1 + AB_2 + \ldots + AB_n$$

$$\therefore \quad P(A) = P(\Sigma\, AB_i)$$

$$= \Sigma\, P(AB_i) \text{ (since } AB_1, AB_2, \ldots, AB_n \text{ are mutually exclusive)}$$

(by addition theorem)

$$= \sum_{i=1}^{n} P(B_i) \times P(A/B_i)$$

Baye's Theorem or Theorem of Probability of Causes

If $B_1, B_2, \ldots, B_n$ be a set of exhaustive and mutually exclusive events associated with a random experiment and A is another event associated with (or caused by) B_i, then

$$P(B_i/A) = \frac{P(B_i) \times P(A/B_i)}{\sum_{i=1}^{n} P(B_i) \times P(A/B_i)}, \; i = 1, 2, \ldots, n$$

Example 1

A bolt is manufactured by 3 machines A, B and C. A turns out twice as many items as B, and machines B and C produce equal number of items. 2% of bolts produced by A and B are defective and 4% of bolts produced by C are defective. All bolts are put into 1 stock pile and 1 is chosen from this pile. What is the probability that it is defective?

Let A = the event in which the item has been produced by machine A, and so on.

Let D = the event of the item being defective.

$$P(A) = \frac{1}{2},\ P(B) = P(C) = \frac{1}{4}$$

$$P(D/A) = P(\text{an item is defective, given that } A \text{ has produced it})$$

$$= \frac{2}{100} = P(D/B)$$

$$P(D/C) = \frac{4}{100}$$

By theorem of total probability,

$$P(D) = P(A) \times P(D/A) + P(B) \times P(D/B) + P(C) \times P(D/C)$$

$$= \frac{1}{2} \times \frac{2}{100} + \frac{1}{4} \times \frac{2}{100} + \frac{1}{4} \times \frac{4}{100}$$

$$= \frac{1}{40}$$

Example 4

A bag contains 5 balls and it is not known how many of them are white. Two balls are drawn at random from the bag and they are noted to be white. What is the chance that all the balls in the bag are white?

Since 2 white balls have been drawn out, the bag must have contained 2, 3, 4 or 5 white balls.

Let B_1 = Event of the bag containing 2 white balls, B_2 = Events of the bag containing 3 white balls, B_3 = Event of the bag containing 4 white balls and B_4 = Event of the bag containing 5 white balls.

Let A = Event of drawing 2 white balls.

$$P(A/B_1) = \frac{2C_2}{5C_2} = \frac{1}{10},\ P(A/B_2) = \frac{3C_2}{5C_2} = \frac{3}{10}$$

$$P(A/B_3) = \frac{4C_2}{5C_2} = \frac{3}{5},\ P(A/B_4) = \frac{5C_2}{5C_2} = 1$$

Since the number of white balls in the bag is not known, B_i's are equally likely.

$$\therefore \qquad P(B_1) = P(B_2) = P(B_3) = P(B_4) = \frac{1}{4}$$

By Baye's theorem,

$$P(B_4/A) = \frac{P(B_4)\times P(A/B_4)}{\sum_{i=1}^{4} P(B_i)\times P(A/B_i)} = \frac{\frac{1}{4}\times 1}{\frac{1}{4}\times\left(\frac{1}{10}+\frac{3}{10}+\frac{3}{5}+1\right)} = \frac{1}{2}$$

Example 5

There are 3 true coins and 1 false coin with 'head' on both sides. A coin is chosen at random and tossed 4 times. If 'head' occurs all the 4 times, what is the probability that the false coin has been chosen and used?

$$P(T) = P(\text{the coin is a true coin}) = \frac{3}{4}$$

$$P(F) = P(\text{the coin is a false coin}) = \frac{1}{4}$$

Let A = Event of getting all heads in 4 tosses

Then $P(A/T) = \frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} = \frac{1}{16}$ and $P(A/F) = 1$.

By Baye's theorem,

$$P(F/A) = \frac{P(F)\times P(A/F)}{P(F)\times P(A/F)+P(T)\times P(A/T)}$$

$$= \frac{\frac{1}{4}\times 1}{\frac{1}{4}\times 1+\frac{3}{4}\times\frac{1}{16}} = \frac{16}{19}$$