.4 THE LAW OF TOTAL PROBABILITY

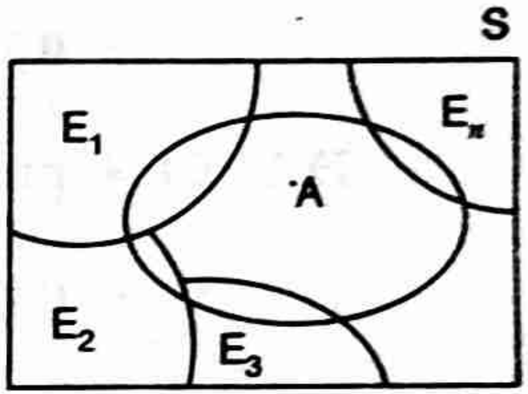
If $E_1, E_2, E_3, \ldots, E_n$ are mutually exclusive and exhaustive events associated with a sample space S of a random experiment and A is any event associated with S, then

$$P(A) = P(E_1)\,P(A \mid E_1) + P(E_2)\,P(A \mid E_2) + \ldots + P(E_n)\,P(A \mid E_n).$$

Proof. Since $E_1, E_2, \ldots, E_n$ are mutually exclusive and exhaustive events of S, therefore,

$$S = E_1 \cup E_2 \cup E_3 \ldots \cup E_n,$$

where $E_i \cap E_j = \phi$ for $i \neq j$.



$$\text{Now } A = A \cap S = A \cap (E_1 \cup E_2 \cup E_3 \ldots \cup E_n)$$

$$= (A \cap E_1) \cup (A \cap E_2) \cup (A \cap E_3) \ldots (A \cap E_n) \quad \ldots(i)$$

Also $A \cap E_i$ and $A \cap E_j$ are subsets of E_i and E_j respectively and it is known that $E_i \cap E_j = \phi$ for $i \neq j$, it follows that $A \cap E_i$ and $A \cap E_j$ are disjoint *i.e.* mutually exclusive for $i \neq j$, $i, j = 1, 2, 3, \ldots, n$.

From (*i*), we get

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + P(A \cap E_3) + \ldots + P(A \cap E_n).$$

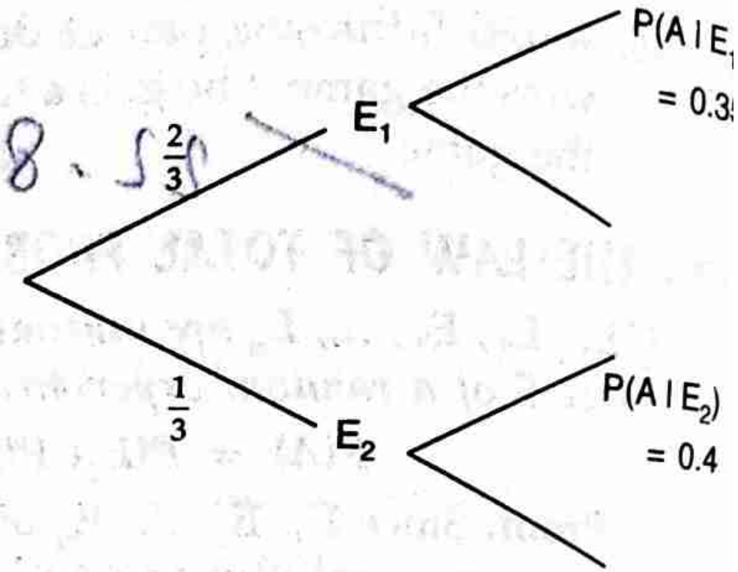
By using multiplication rule of probability, we get

$$P(A) = P(E_1)\,P(A \mid E_1) + P(E_2)\,P(A \mid E_2) + \ldots + P(E_n)\,P(A \mid E_n).$$

Example 2. *Two-thirds of the students of a class are boys and the rest are girls. It is known that th probability of a girl getting a first class marks in Council's Exam is 0.4 and a boy getting first clas marks is 0.35. Find the probability that a student chosen at random will get first class marks in Exam*

Solution. Let E_1, E_2 and A be the events defined as follows :

$$E_1 = \text{a boy is chosen,}$$
$$E_2 = \text{a girl is chosen and}$$
$$A = \text{the student gets first class marks.}$$



Then $P(E_1) = \frac{2}{3}$ and $P(E_2) = \frac{1}{3}$.

Note that E_1 and E_2 are mutually exclusive and exhaustive events.

$P(A | E_1)$ = probability of a boy getting first class marks

$$= 0.35 = \frac{35}{100} = \frac{7}{20} \text{ and}$$

$P(A | E_2)$ = probability of a girl getting first class marks

$$= 0.4 = \frac{4}{10} = \frac{2}{5}.$$

By using law of total probability, we get

$$P(A) = P(E_1)\, P(A | E_1) + P(E_2)\, P(A | E_2)$$

$$= \frac{2}{3} \cdot \frac{7}{20} + \frac{1}{3} \cdot \frac{2}{5} = \frac{7}{30} + \frac{2}{15} = \frac{11}{30}.$$

Example 6. *A fair die is rolled. If 1 turns up, a ball is picked up at random from bag A. If 2 or 3 turns up, a ball is picked up from bag B. If 4, 5 or 6 turns up, a ball is picked up from bag C. Bag A contains 3 red and 2 white balls; bag B contains 3 red and 4 white balls; bag C contains 4 red and 5 white balls. The die is rolled, a bag is selected and a ball is drawn. Find the probability that a red ball is drawn.*

Solution. Let E_1, E_2, E_3 and A be the events defined as follows :

E_1 = bag A is picked up,

E_2 = bag B is picked up,

E_3 = bag C is picked up and

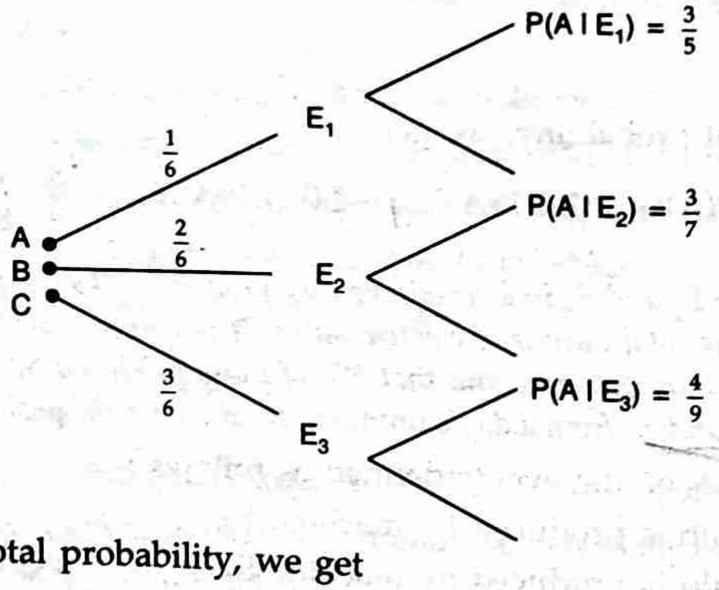
A = a red ball is drawn from the selected bag.

Then $P(E_1) = \frac{1}{6}$, $P(E_2) = \frac{2}{6}$ and $P(E_3) = \frac{3}{6}$.

Note that E_1, E_2 and E_3 are mutually exclusive and exhaustive events.

$P(A|E_1)$ = probability of drawing a red ball from bag A = $\frac{3}{5}$,

$P(A|E_2) = \frac{3}{7}$ and $P(A|E_3) = \frac{4}{9}$ (as shown in tree diagram)



By using law of total probability, we get

$$P(A) = P(E_1)\,P(A|E_1) + P(E_2)\,P(A|E_2) + P(E_3)\,P(A|E_3)$$

$$= \frac{1}{6}\cdot\frac{3}{5} + \frac{2}{6}\cdot\frac{3}{7} + \frac{3}{6}\cdot\frac{4}{9} = \frac{1}{10} + \frac{1}{7} + \frac{2}{9}$$

$$= \frac{63 + 90 + 140}{630} = \frac{293}{630}.$$

10.5 BAYE'S THEOREM

If $E_1, E_2, E_3, \ldots, E_n$ are mutually exclusive and exhaustive events associated with a random experiment and A is any event associated with the experiment, then

$$P(E_i \mid A) = \frac{P(E_i)\, P(A \mid E_i)}{\Sigma P(E_i)\, P(A \mid E_i)}, \text{ where } i = 1, 2, 3, \ldots, n.$$

Proof. By the law of total probability, we have

$$P(A) = P(E_1)\, P(A \mid E_1) + P(E_2)\, P(A \mid E_2) + \ldots + P(E_n)\, P(A \mid E_n)$$
$$= \Sigma P(E_i)\, P(A \mid E_i) \qquad \ldots(i)$$

Also by multiplication law of probability, we have

$$P(A \cap E_i) = P(A)\, P(E_i \mid A) = P(E_i)\, P(A \mid E_i),\ i = 1, 2, 3, \ldots, n$$

$$\Rightarrow \quad P(E_i \mid A) = \frac{P(E_i)\, P(A \mid E_i)}{P(A)},\ i = 1, 2, 3, \ldots, n$$

$$\Rightarrow \quad P(E_i \mid A) = \frac{P(E_i)\, P(A \mid E_i)}{\Sigma P(E_i)\, P(A \mid E_i)},\ i = 1, 2, 3, \ldots, n \qquad \text{(by using (i))}$$

The probability $P(E_i \mid A)$ means finding the probability of event E_i given that event A has occurred. Probability $P(E_i)$ was already known — so it was ***a priori probability***. $P(E_i \mid A)$ is to be calculated after the knowledge that event A has happened — so it is called ***posteriori probability***.

For example, suppose that in a factory, 60% product are manufactured by machine M_1 and 40% by machine M_2. Machine M_1 produces 1% defective items and machine M_2 produces 2% defective items, and let

E_1 = event that product is manufactured by machine M_1,

E_2 = event that product is manufactured by machine M_2 and

A = event that product is defective.

Example. *A man is known to speak truth 3 out of 4 times. He throws a die and reports that it is a six. Find the probability that it is actually a six.*

Solution. Let E_1, E_2 and A be the events defined as follows :

E_1 = die shows six *i.e.* six has occurred,

E_2 = die does not show six *i.e.* six has not occurred and

A = the man reports that six has occurred.

We wish to calculate the probability that six has actually occurred given that the man reports that six occurs *i.e.* $P(E_1 | A)$.

Now, $P(E_1) = \frac{1}{6}$, $P(E_2) = \frac{5}{6}$,

$P(A | E_1)$ = probability that the man reports that six occurs given that six has occurred

= probability that the man is speaking the truth = $\frac{3}{4}$ and

$P(A | E_2)$ = probability that the man reports that six occurs given that six has not occurred

= probability that the man does not speak truth = $\frac{1}{4}$.

By Bayes' theorem, we have :

$$P(E_1 | A) = \frac{P(E_1)\,P(A | E_1)}{P(E_1)\,P(A | E_1) + P(E_2)\,P(A | E_2)}$$

$$= \frac{\frac{1}{6}\cdot\frac{3}{4}}{\frac{1}{6}\cdot\frac{3}{4} + \frac{5}{6}\cdot\frac{1}{4}} = \frac{3}{8}.$$

Coloured balls are distributed in four boxes as follows :

Box	Colour			
	Black	White	Red	Blue
I	3	4	5	6
II	2	2	2	2
III	1	2	3	1
IV	4	3	1	5

A box is selected at random and a ball is drawn. If the colour of the ball is black, what is the probability that ball drawn is from the box III?

Solution. Let E_1, E_2, E_3, E_4 and A be the events defined as follows :

E_1 = box I is chosen,
E_2 = box II is chosen,
E_3 = box III is chosen,
E_4 = box IV is chosen and
A = ball draw is black.

As a box is selected at random,

$$\therefore \quad P(E_1) = P(E_2) = P(E_3) = P(E_4) = \frac{1}{4}.$$

Box I contains 3 black, 4 white, 5 red and 6 blue balls, so the total number of balls in box I = 3 + 4 + 5 + 6 = 18.

$P(A \mid E_1)$ = probability of drawing a black ball when E_1 has occurred *i.e.* drawing a black ball from box I

$$= \frac{3}{18} = \frac{1}{6}.$$

Similarly, $P(A \mid E_2) = \frac{2}{8} = \frac{1}{4}$, $P(A \mid E_3) = \frac{1}{7}$ and $P(A \mid E_4) = \frac{4}{13}$.

By using law of total probability,

$$P(A) = \Sigma P(E_i)\, P(A \mid E_i) = \frac{1}{4}\cdot\frac{1}{6} + \frac{1}{4}\cdot\frac{1}{4} + \frac{1}{4}\cdot\frac{1}{7} + \frac{1}{4}\cdot\frac{4}{13}$$

$$= \frac{1}{4}\left(\frac{1}{6} + \frac{1}{4} + \frac{1}{7} + \frac{4}{13}\right) = \frac{1}{4}\cdot\frac{182 + 273 + 156 + 336}{12 \times 91} = \frac{947}{48 \times 91}.$$

We want to find $P(E_3 \mid A)$.

By Baye's theorem, we have :

$$P(E_3 \mid A) = \frac{P(E_3)\, P(A \mid E_3)}{P(A)} = \frac{\frac{1}{4}\cdot\frac{1}{7}}{\frac{947}{48 \times 91}} = \frac{1}{28} \times \frac{48 \times 91}{947} = \frac{156}{947}.$$

Random Variables and Probability Distributions

2.1 RANDOM VARIABLES

Suppose that to each point of a sample space we assign a number. We then have a *function* defined on the sample space. This function is called a *random variable* (or *stochastic variable*) or more precisely a *random function* (*stochastic function*). It is usually denoted by a capital letter such as X or Y. In general, a random variable has some specified physical, geometrical, or other significance.

Example 2.1 Suppose that a coin is tossed twice so that the sample space is $S = \{HH, HT, TH, TT\}$. Let X represent the number of heads that can come up. With each sample point we can associate a number for X as shown in Table 2.1. Thus, for example, in the case of HH (i.e., 2 heads), $X = 2$ while for TH (1 head), $X = 1$. It follows that X is a random variable.

It should be noted that many other random variables could also be defined on this sample space, for example, the square of the number of heads or the number of heads minus the number of tails.

Table 2.1

Sample Point	*HH*	*HT*	*TH*	*TT*
X	2	1	1	0

A random variable that takes on a finite or countably infinite number of values (see page 4) is called a *discrete random variable* while one which takes on a noncountably infinite number of values is called a *nondiscrete random variable*.

2.2 DISCRETE PROBABILITY DISTRIBUTIONS

Let X be a discrete random variable, and suppose that the possible values that it can assume are given by $x_1, x_2, x_3, \ldots$, arranged in some order. Suppose also that these values are assumed with probabilities given by

$$P(X = x_k) = f(x_k) \qquad k = 1, 2, \ldots \tag{1}$$

It is convenient to introduce the *probability function*, also referred to as *probability distribution*, given by

$$P(X = x) = f(x) \tag{2}$$

For $x = x_k$, this reduces to (1) while for other values of x, $f(x) = 0$.

In general, $f(x)$ is a probability function if

1. $f(x) \geq 0$
2. $\sum f(x) = 1$

where the sum in 2 is taken over all possible values of x.

Example 2.2 Find the probability function corresponding to the random variable X of Example 2.1. Assuming that the coin is fair, we have

$$P(HH) = \frac{1}{4} \qquad P(HT) = \frac{1}{4} \qquad P(TH) = \frac{1}{4} \qquad P(TT) = \frac{1}{4}$$

Then

$$P(X = 0) = P(TT) = \frac{1}{4}$$

$$P(X = 1) = P(HT \cup TH) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(HH) = \frac{1}{4}$$

The probability function is thus given by Table 2.2.

Table 2.2

x	0	1	2
$f(x)$	1/4	1/2	1/4

2.3 DISTRIBUTION FUNCTIONS FOR RANDOM VARIABLES

The *cumulative distribution function*, or briefly the *distribution function*, for a random variable X is defined by

$$F(x) = P(X \le x) \tag{3}$$

where x is any real number, i.e., $-\infty < x < \infty$.

The distribution function $F(x)$ has the following properties:

1. $F(x)$ is nondecreasing [i.e., $F(x) \le F(y)$ if $x \le y$].
2. $\lim_{x \to -\infty} F(x) = 0$; $\lim_{x \to \infty} F(x) = 1$.
3. $F(x)$ is continuous from the right [i.e., $\lim_{h \to 0^+} F(x+h) = F(x)$ for all x].

2.4 DISTRIBUTION FUNCTIONS FOR DISCRETE RANDOM VARIABLES

The distribution function for a discrete random variable X can be obtained from its probability function by noting that, for all x in $(-\infty, \infty)$,

$$F(x) = P(X \le x) = \sum_{u \le x} f(u) \tag{4}$$

where the sum is taken over all values u taken on by X for which $u \le x$.

If X takes on only a finite number of values $x_1, x_2, \ldots, x_n$, then the distribution function is given by

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \le x < x_2 \\ f(x_1) + f(x_2) & x_2 \le x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \le x < \infty \end{cases} \tag{5}$$

Example 2.3 (a) Find the distribution function for the random variable X of Example 2.2. (b) Obtain its graph.

(a) The distribution function is

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{4} & 0 \le x < 1 \\ \frac{3}{4} & 1 \le x < 2 \\ 1 & 2 \le x < \infty \end{cases}$$

(b) The graph of $F(x)$ is shown in Fig. 2.1.

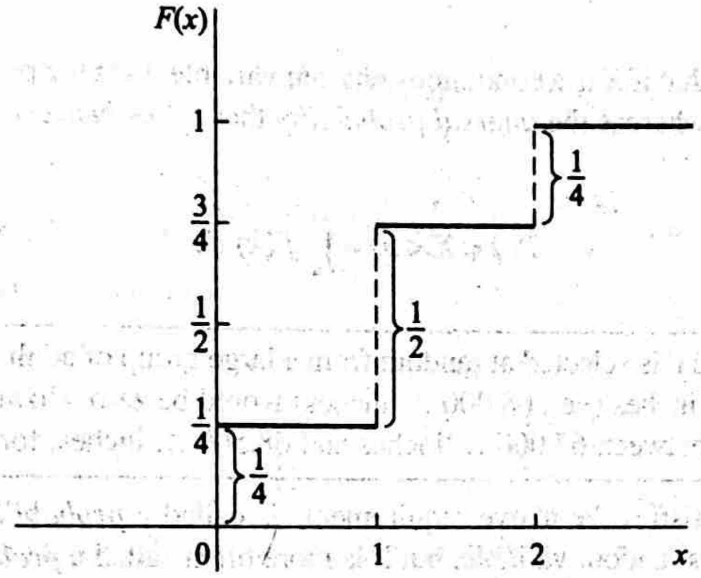


Fig. 2.1

The following things about the above distribution function, which are true in general, should be noted.

1. The magnitudes of the jumps at 0, 1, 2 are $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ which are precisely the probabilities in Table 2.2. This fact enables one to obtain the probability function from the distribution function.
2. Because of the appearance of the graph of Fig. 2.1, it is often called a *staircase function* or *step function*. The value of the function at an integer is obtained from the higher step; thus the value at 1 is $\frac{3}{4}$ and not $\frac{1}{4}$. This is expressed mathematically by stating that the distribution function is *continuous from the right* at 0, 1, 2.
3. As we proceed from left to right (i.e. going *upstairs*), the distribution function either remains the same or increases, taking on values from 0 to 1. Because of this, it is said to be a *monotonically increasing function*.

It is clear from the above remarks and the properties of distribution functions that the probability function of a discrete random variable can be obtained from the distribution function by noting that

$$f(x) = F(x) - \lim_{u \to x^-} F(u) \tag{6}$$

c) Distribution function for Con. r. v.

$$F_X(x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f(u)\, du$$

where the p.d.f. $f(x)$ is such that $(-\infty < x < \infty)$,

i) $f(x) \geq 0$

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

Again

i) As $P(X = c) = F(c) - F(c-0)$,

hence $P(X=c) = 0$ ($\because F(x)$ is continuous) for X is a con r. v.

ii) $P(a < X < b) = \int_a^b f(x)\, dx$

Hence if an individual is selected at random from a large group of adult males, the prob. that his height X is precisely 68" would be zero. However, there is a +ve prob that X lies between 67.5" & 68.5".

$P(X=x) = f(x) = F(x) - \lim_{u \to x^-} F(u)$ (jump)

$\Rightarrow P(X=c) = F(c) - \lim_{x \to c^-} F(x)$

$= F(c) - F(c-0)$

$= F(c) - F(c) = 0$.

Example 2.5 (a) Find the constant c such that the function

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

is a density function, and (b) compute $P(1 < X < 2)$.

(a) Since $f(x)$ satisfies Property 1 if $c \geq 0$, it must satisfy Property 2 in order to be a density function. Now

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^3 cx^2 dx = \left.\frac{cx^3}{3}\right|_0^3 = 9c$$

and since this must equal 1, we have $c = 1/9$.

(b) $$P(1 < X < 2) = \int_1^2 \frac{1}{9} x^2 dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{1}{27}$$

In case $f(x)$ is continuous, which we shall assume unless otherwise stated, the probability that X is equal to any particular value is zero. In such case we can replace either or both of the signs $<$ in (8) by $\leq$. Thus, in Example 2.5,

$$P(1 \le X \le 2) = P(1 \le X < 2) = P(1 < X \le 2) = P(1 < X < 2) = \frac{7}{27}$$

Example 2.6 (a) Find the distribution function for the random variable of Example 2.5. (b) Use the result of (a) to find $P(1 < x \le 2)$.

(a) We have

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\, du$$

If $x < 0$, then $F(x) = 0$. If $0 \le x < 3$, then

$$F(x) = \int_{0}^{x} f(u)\, du = \int_{0}^{x} \frac{1}{9} u^2 du = \frac{x^3}{27}$$

If $x \ge 3$, then

$$F(x) = \int_{0}^{3} f(u)\, du + \int_{3}^{x} f(u)\, du = \int_{0}^{3} \frac{1}{9} u^2 du + \int_{3}^{x} 0\, du = 1$$

Thus the required distribution function is

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3/27 & 0 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

Note that $F(x)$ increases monotonically from 0 to 1 as is required for a distribution function. It should also be noted that $F(x)$ in this case is continuous.

(b) We have

$$\begin{aligned} P(1 < X \le 2) &= P(X \le 2) - P(X \le 1) \\ &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27} \end{aligned}$$

as in Example 2.5.

The probability that X is between x and $x + \Delta x$ is given by

$$P(x \le X \le x + \Delta x) = \int_{x}^{x+\Delta x} f(u)\, du \qquad (9)$$

so that if Δx is small, we have approximately

$$P(x \le X \le x + \Delta x) = f(x)\, \Delta x \qquad (10)$$

We also see from (7) on differentiating both sides that

$$\frac{dF(x)}{dx} = f(x) \qquad (11)$$

at all points where $f(x)$ is continuous; i.e., the derivative of the distribution function is the density function.