INTRODUCTION TO PROBABILITY AND PROBABILITY DISTRIBUTIONS

2 authors:

John Benjamin Ofosu
Methodist University College Ghana

Christian Akrong Hesse
Methodist University College Ghana

INTRODUCTION TO

PROBABILITY

AND

PROBABILITY DISTRIBUTIONS

J. B. OFOSU, BSc, PhD, FSS
Professor of Statistics
Methodist University College Ghana

C. A. HESSE, BSc, MPhil
Lecturer in Statistics
Methodist University College Ghana

PREFACE

This book has been written primarily to answer the growing need for a one-semester course in probability and probability distributions for University and Polytechnic students in engineering and physical sciences. No previous course in probability is needed in order to study the book. The mathematical preparation required is the conventional full-year calculus course which includes series expansion, multiple integration and partial differentiation.

The book has nine chapters. Chapter 1 covers the basic tools of probability theory. In Chapter 2, we discuss concepts of random variables and probability distributions. Chapter 3 covers numerical characteristics of random variables and in Chapter 4 we discuss moment generating functions and characteristic functions. Chapters 5 and 6 treat important probability distributions, their applications, and relationships between probability distributions. Chapter 7 extends the concept of univariate random variables to bivariate random variables. In Chapter 8, we discuss distributions of functions of random variables and in Chapter 9, we discuss order statistics, probability inequalities and modes of convergence.

The book is written with the realization that concepts of probability and probability distributions – even though they often appear deceptively simple – are in fact difficult to comprehend. Every basic concept and method is therefore explained in full, in a language that is easily understood.

A prominent feature of the book is the inclusion of many examples. Each example is carefully selected to illustrate the application of a particular statistical technique and or interpretation of results. Another feature is that each chapter has an extensive collection of exercises. Many of these exercises are from published sources, including past examination questions from King Saud University in Saudi Arabia, and Methodist University College Ghana. Answers to all the exercises are given at the end of the book.

We are grateful to Professor O. A. Y. Jackson of Methodist University College Ghana, for reading a draft of the book and offering helpful comments and suggestions. We are also indebted to Professor Abdullah Al-Shiha of King Saud University for his permission to publish the statistical tables he used the Minitab software package to prepare. These tables are given in the Appendix. Last, but not least, we thank King Saud University and Methodist University College Ghana, for permission to use their past examination questions in probability and probability distributions.

We have gone to great lengths to make this text both pedagogically sound and error-free. If you have any suggestions, or find potential errors, please contact us at **jonofosu@hotmail.com or akrongh@yahoo.com**.

J. B. Ofosu

C. A. Hesse

July, 2009

CONTENTS

FOREWORD

By Prof. A. H. O. Mensah
Methodist University College Ghana

I am privileged to provide a foreword for this book on "Probability and Probability Distributions" – a piece of work produced by two of our most competent and highly respected professors here in our Methodist University College Ghana for university and polytechnic students in engineering and the physical sciences.

This highly valuable product supplies a need for a one-semester course in probability and probability distributions in a most illuminating and facilitating form.

Even though, as stated by the authors, the concepts of probability and probability distributions are often difficult to comprehend, their approach, style and pains-taking treatment of the subject make it easy for the student to understand and the teacher to deliver.

The book covers several aspects of the subject including the basic tools of probability theory, concepts of random variables and probability distributions, the numerical characteristics of random variables, moment generating functions and characteristic functions, extension of univariate random variables to bivariate random variables, distributions of functions of random variables, order statistics, probability inequalities and modes of convergence.

This book is certainly an invaluable asset for those who learn and those who teach.

Get copies now!

CHAPTER ONE

Probability

1.1 Random experiments, sample space and events

1.1.1 Random experiments

If we record the number of road traffic accidents in Ghana every day, we will find a large variation in the numbers we get each day. On some days, many accidents will be recorded while on other days, very few accidents will be recorded. That is, there is a large variation in the number of road traffic accidents which occur in Ghana every day. We cannot therefore tell in advance, the number of road traffic accidents in Ghana on a given day. The same sort of remark can be made about the number of babies who are born in Accra every week, the number of customers who enter a bank in a given time interval when the bank is open. Such experiments, because their outcomes are uncertain, are called ***random experiments***.

Example 1.1

Picking a ball from a box containing 20 numbered balls, is a random experiment, since the process can lead to one of the 20 possible outcomes.

Example 1.2

Consider the experiment of rolling a six-sided die and observing the number which appears on the uppermost face of the die. The result can be any of the numbers 1, 2, 3, …, 6. This is a random experiment since the outcome is uncertain.

Example 1.3

If we measure the distance between two points A and B, many times, under the same conditions, we expect to have the same result. This is therefore not a random experiment. It is a ***deterministic experiment***.

If a deterministic experiment is repeated many times under exactly the same conditions, we expect to have the same result.

Probability allows us to quantify the variability in the outcome of a random experiment. However, before we can introduce probability, it is necessary to specify the space of outcomes and the events on which it will be defined.

1.1.2 Sample space

In statistics, the set of all possible outcomes of an experiment is called the ***sample space*** of the experiment, because it usually consists of the things that can happen when one takes a sample. Sample spaces are usually denoted by the letter S.

Each outcome in a sample space is called an ***element*** or a ***member*** of the sample space, or simply a ***sample point***.

Example 1.4

Consider the experiment of rolling a red die and a green die and observing the number which appears on the uppermost face of each die. The sample space of the experiment consists of the following array of 36 outcomes.

The first coordinate of each point is the number which appears on the red die, while the second coordinate is the number which appears on the green die.

	Green die					
Red die	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
	⋮	⋮	⋮	⋮	⋮	⋮
	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Using the standard notation for sets, we can express this sample space as follows:

$$S = \{(i, j): i = 1, 2, 3, 4, 5, 6;\ j = 1, 2, 3, 4, 5, 6\}.$$

Example 1.5

Define a sample space for each of the following experiments.

(a) The heights, in centimetres, of five children are 60, 65, 70, 45, 48. Select a child from this group of children, then measure and record the child's height.

(b) Select a number at random from the interval [0, 2] of real numbers. Record the value of the number selected.

Solution

(a) $S = \{60, 65, 70, 45, 48\}$. (b) $S = \{x: 0 \le x \le 2$, where x is a real number$\}$.

In some experiments, it is helpful to list the elements of the sample space systematically by means of a ***tree diagram***. The following example illustrates the idea.

Example 1.6

Suppose that three items are selected at random from a manufacturing process. Each item is inspected and classified as defective, D, or non-defective, N. To list the elements of the sample space, we first construct the tree diagram as shown in Fig. 1.1. Now, the various paths along the branches of the tree give the distinct sample points. Starting with the first path, we get the sample point DDD.

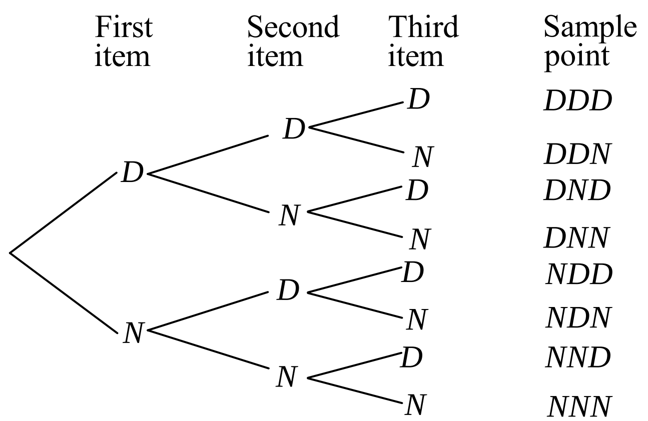

***Fig. 1.1**: A tree diagram for Example 1.6*

As we proceed along the other paths, we see that the sample space of the experiment is

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

Sample spaces are classified according to the number of elements they contain. A sample space is said to be a ***finite sample space*** if it consists of a finite number of elements. It is a ***discrete sample space*** if the elements can be placed in a one-to-one correspondence with the positive integers, or if it is finite. If the elements (points) of a sample space is an interval of the number line, then the sample space is said to be a ***continuous sample space*** (see Example 1.5(b)).

1.1.3 Event

A subset of a sample space is called an ***event***. The empty set, ϕ, is a subset of S and S is also a subset of S. ϕ and S are therefore events. We call ϕ the ***impossible event*** and S the ***certain event***. A subset of S containing one element of S is called a **simple event**.

For any given experiment, we may be interested in the occurrence of certain events rather than in the outcome of a specific element in the sample space. For instance, in Example 1.4, we may be interested in the event, A, that the sum of the numbers which appear is greater than 10. This event can be expressed as $A = \{(5, 6), (6, 5), (6, 6)\}$. As a further illustration, we may be interested in the event, B, that the number of defective items is greater than 1 in Example 1.6. This event will occur if the outcome is an element of the subset $B = \{DDN, DND, NDD, DDD\}$ of the sample space.

It is important to understand that an event occurs if and only if any of its elements is the outcome of an experiment. For instance, in Example 1.4, $P = \{(1, 2), (2, 3), (6, 4)\}$, $Q = \{(1, 4), (4, 4), (3, 6)\}$ and $R = \{(6, 4), (2, 5), (1, 3)\}$ are events. If, when the dice are rolled, the number 6 appears on the red die and the number 4 appears on the green die, then the events P and R have occurred, since each of these events has (6, 4) as an element. Event Q did not occur since $(6, 4) \notin Q$.

Example 1.7

Given the sample space $S = \{t: t \geq 0\}$, where t is the life in years of an electronic component, the event, A, that the component fails before the end of the third year, is the subset $A = \{t: 0 \leq t < 3\}$.

Equally likely events

Consider the experiment of rolling a six-sided die. If the die is symmetrical, then all the six numbers have the same chance of appearing. The six events $\{1\}, \{2\}, \{3\}, \ldots, \{6\}$, are then said to be ***equally likely***. Frequently, the manner in which an experiment is performed, determines whether or not all the possible outcomes of the experiment are equally likely. For example, if identical balls of different colours are placed in a box and a ball is selected without looking, then this gives rise to equally likely outcomes. A sample space in which all the simple events are equally likely is called an ***equiprobable sample space***. The expression "***at random***" will be used only with respect to an equiprobable sample space. Thus, the statement ***"a number is chosen at random from a set S"***, will mean that S is an equiprobable sample space.

1.2 Operations on events

Since an event is a subset of a sample space, we can combine events to form new events, using the various set operations. The sample space is considered as the universal set. If *A* and *B* are two events defined on the same sample space, then:

(1) $A \cup B$ denotes the event "*A* or *B* or both". Thus the event $A \cup B$ occurs if either *A* occurs or *B* occurs or both *A* and *B* occur.

(2) $A \cap B$ denotes the event "both *A* and *B*". Thus the event $A \cap B$ occurs if both *A* and *B* occur.

(3) A' (or $\overline{A}$) denotes the event which occurs if and only if *A* does not occur.

Example 1.8

Let *P* be the event that an employee selected at random from an oil drilling company smokes cigarettes. Let *Q* be the event that an employee selected drinks alcoholic beverages. Then $P \cup Q$ is the event that an employee selected either drinks or smokes, or drinks and smokes. $P \cap Q$ is the event that an employee selected drinks and smokes.

Example 1.9

If $A = \{x: 3 < x < 9\}$ and $B = \{y: 5 \leq y < 12\}$, then $A \cup B = \{z: 3 < z < 12\}$, and $A \cap B = \{y: 5 \leq y < 9\}$.

Sample spaces and events, particularly relationships among events, are often depicted by means of Venn diagrams like those of Fig. 1.2. In each case, the sample space is represented by a rectangle, whereas events are represented by regions within the rectangle, usually by circles or parts of circles. The shaded regions of the four diagrams of Fig. 1.2 represent event *A*, the complement of *A*, the union of events *A* and *B*, and the intersection of events *A* and *B*.

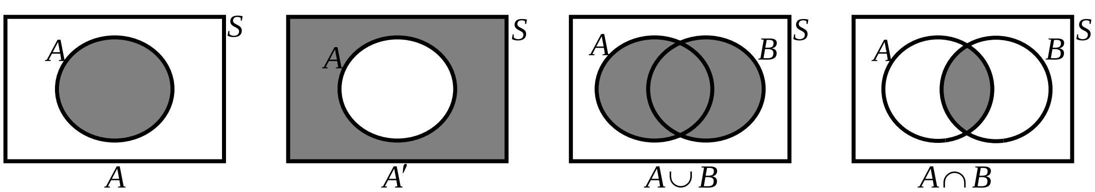

***Fig. 1.2**: Venn diagrams showing the complement, union and intersection*

When we deal with three events, we draw the circles as in Fig. 1.3. In this diagram, the circles divide the sample space into eight regions, numbered 1 through 8, and it is easy to determine whether the corresponding events are parts of *A* or A', *B* or B', and *C* or C'.

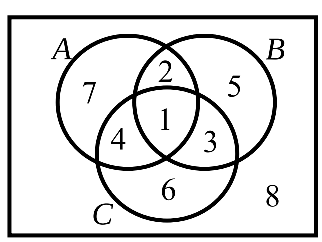

***Fig. 1.3**: A Venn diagram*

Example 1.10

A manufacturer of small motors is concerned with three major types of defects. If A is the event that the shaft size is too large, B is the event that the windings are improper, and C is the event that the electrical connections are unsatisfactory, express in words the events represented by the following regions of the Venn diagram in Fig. 1.3.

(a) region 2, (b) regions 1 and 3 together,
(c) region 7, (d) regions 3, 5, 6, and 8 together.

Solution

(a) Since this region is contained in A and B but not in C, it represents the event that the shaft size is too large and the windings are improper, but the electrical connections are satisfactory.
(b) Since this region is common to B and C, it represents the event that the windings are improper and the electrical connections are unsatisfactory.
(c) This region is in A but not in B and not in C. The region therefore represents the event that the shaft size is too large, the windings are proper and the electrical connections are satisfactory.
(d) Since this is the entire region outside A, it represents the event that the shaft size is not too large.

Mutually exclusive (or disjoint) events

Any two events that cannot occur simultaneously, so that their intersection is the impossible event, are said to be ***mutually exclusive*** (or disjoint). Thus two events A and B are mutually exclusive if and only if $A \cap B = \phi$. In general, a collection of events $A_1, A_2, \ldots, A_n$, is said to be mutually exclusive if there is no overlap among any of them. That is, if $A_i \cap A_j = \phi$ $(i \neq j,\ i, j = 1, 2, \ldots n)$. Fig. 1.4 shows several mutually exclusive events

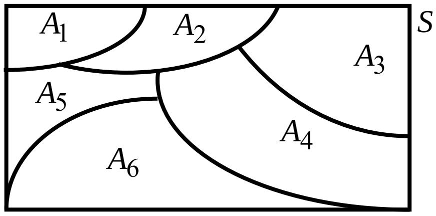

***Fig. 1.4**: Mutually exclusive events*

De Morgan's Laws

Venn diagrams are often used to verify relationships among sets, thus making it unnecessary to give formal proofs based on the algebra of sets. To illustrate, let us show that $(A \cup B)' = A' \cap B'$, which expresses the fact that the complement of the union of two sets equals the intersection of their complements. To begin, note that in Figures 1.5 and 1.6, A and B are events defined on the same sample space S. In Fig. 1.5, the shaded area represents the event $(A \cup B)'$.

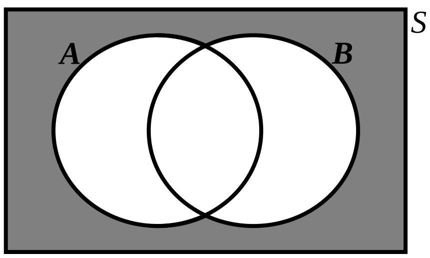

***Fig. 1.5**: $(A \cup B)'$ is shaded*

In Fig. 1.6, the area shaded vertically represents the event A' while the area shaded horizontally represents the event B'. It follows that the cross-shaded area represents the event $A' \cap B'$. But the total shaded area in Fig. 1.5 is identical to the cross-shaded area in Fig. 1.6. We can therefore state the following theorem.

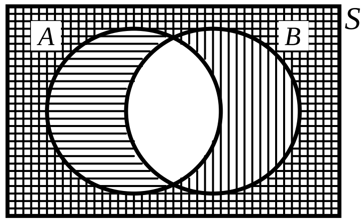

Fig. 1.6*: A' and B' are shaded*

Theorem 1.1

$$(A \cup B)' = A' \cap B'.$$

Similarly, we can use Venn diagrams to verify the following two theorems.

Theorem 1.2

$$(A \cap B)' = A' \cup B'.$$

Theorem 1.3

If A, B and C are events defined on the same sample space, then
$(A \cup B \cup C)' = A' \cap B' \cap C'$, $(A \cap B \cap C)' = A' \cup B' \cup C'$, and $(A' \cup B' \cup C)' = A \cap B \cap C'$.

The results given in Theorems 1.1, 1.2 and 1.3 are called ***de Morgan's laws***.

Other useful facts concerning operations on events

The following results can be verified by means of Venn diagrams.

1. **Commutative law**

   $A \cup B = B \cup A$; $A \cap B = B \cap A$.

2. **Associative law**

   $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$.

   Because of these two statements, we can use the simpler notations $A \cup B \cup C$ and $A \cap B \cap C$ without fear of ambiguity when referring to $A \cup (B \cup C)$ and $A \cap (B \cap C)$, respectively.

3. **Distributive law**

   $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

4. **Other useful results**

   (a) $S' = \phi$, $\phi' = S$, $(A')' = A$.

   (b) $A \cup S = S$, $A \cap \phi = \phi$.

   (c) $A \cup A' = S$, $A \cap A' = \phi$.

Exercise 1(a)

1. Use Venn diagrams to verify that:
   (a) $A \cup (A \cap B) = A$, (b) $(A \cap B) \cup (A \cap B') = A$, (c) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

2. A experiment involves tossing a red and a green die, and recording the numbers that come up (see Example 1.4 on page 2).
   (a) List the elements corresponding to the event, A, that the sum is greater than 8.
   (b) List the elements corresponding to the event, B, that a 2 occurs on either die.
   (c) List the elements corresponding to the event, C, that a number greater than 4 comes up on the green die.
   (d) List the elements corresponding to the following events
   (i) $A \cap C$, (ii) $A \cap B$, (iii) $B \cap C$.

3. Consider the experiment of rolling two dice (see Example 1.4 on page 2).
   (a) Let F = "the sum of the two numbers which appear on the dice is 8". List the sample points in F.
   (b) Express the following events in words:
   $G = \{(1, 1), (1, 2), (2, 1)\}$,
   $H = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\}$.
   (c) Let I = "the sum of the two numbers which appear on the dice is even". List the sample points in I.
   (d) Let J = "the number on the red die is 3". List the sample points in J.
   (e) Let $K = \{(3, 3)\}$. If K occurs, does J occur?
   (f) If J occurs, does K occur?
   (g) Are the events G and H mutually exclusive?
   (h) List the sample points in $G \cap H$.
   (i) Let $L = \{(2, 2)\}$. Are the events G and L mutually exclusive?

4. Let A, B and C be 3 events defined on the same sample space. Find the simplest expression for each of the following events:
   (a) None of the events A, B or C occurred.
   (b) At least one of the events A, B, C occurred.
   (c) Exactly one of the events A, B, C occurred.

   Use set notation to make the following statements:
   (d) If A occurred, then B occurred.
   (e) Events B and C cannot occur simultaneously.

5. Let A, B and C be events defined on the same sample space of an experiment. Which of the following statements are true?
   (a) A and A' are mutually exclusive events.
   (b) If A and B are mutually exclusive, then A' and B' are mutually exclusive.
   (c) If the occurrence of A implies the occurrence of B, then the occurrence of A' implies the occurrence of B'.
   (d) If the occurrence of A implies the occurrence of B, then A and B' are mutually exclusive.

(e) If A and B are mutually exclusive and B and C are mutually exclusive, then A and C are mutually exclusive.

(f) If A and B are mutually exclusive, then the occurrence of A implies the occurrence of B'.

1.3 Counting sample points

In this section, we discuss some techniques for determining, without direct enumeration, the number of possible outcomes of an experiment. Such techniques are useful in finding probabilities of some complex events.

The multiplication principle

We begin with the following basic principle:

Theorem 1.4 (The multiplication theorem)

If an operation can be performed in n_1 ways and after it is performed in any one of these ways, a second operation can be performed in n_2 ways and, after it is performed in any one of these ways, a third operation can be performed in n_3 ways, and so on for k operations, then the k operations can be performed together in $n_1 n_2 \ldots n_k$ ways.

Example 1.11

How many lunches consisting of a soup, sandwich, dessert, and a drink are possible if we can select 4 soups, 3 kinds of sandwiches, 5 desserts and 4 drinks?

Solution

Here, $n_1 = 4,\ n_2 = 3,\ n_3 = 5$ and $n_4 = 4$. Hence there are

$$n_1 \times n_2 \times n_3 \times n_4 = 4 \times 3 \times 5 \times 4 = 240$$

different ways to choose a lunch.

Example 1.12

How many even three-digit numbers can be formed from the digits 3, 2, 5, 6, and 9 if each digit can be used only once?

Solution

Since the number must be even, we have $n_1 = 2$ choices for the units position. For each of these, we have $n_2 = 4$ choices for the tens position and then $n_3 = 3$ choices for the hundreds position. Therefore we can form a total of

$$n_1 \times n_2 \times n_3 = 2 \times 4 \times 3 = 24$$

even three-digit numbers.

Permutations

Frequently, we are interested in a sample space that contains, as elements, all possible arrangements of a group of objects. For example, we may want to know how many different arrangements are possible for sitting 6 people around a table, or we may ask how many different ways are possible for drawing 2 lottery tickets from a total of 20. The different arrangements are called **permutations**.

Consider the following example.

Example 1.13

In how many ways can three different books, *A*, *B* and *C*, be arranged on a shelf?

Solution

The three books can be arranged on a shelf in 6 different ways, as follows:

ABC, *ACB*, *BCA*, *BAC*, *CAB*, *CBA*.

It can be seen that the order of the books is important. Each of the 6 arrangements is therefore a different arrangement of the books.

Notice that the first book can be chosen in 3 different ways. Following this, the second book can be chosen in 2 different ways and following this, the third book can be chosen in one way. The three books can therefore be arranged on a shelf in $3 \times 2 \times 1 = 6$ ways.

In general, n distinct objects can be arranged in

$n(n-1)(n-2)...(3)(2)(1)$ ways.

We represent this product by the symbol $n!$, which is read "n factorial". Three objects can be arranged in 3! = (3)(2)(1) = 6 ways. By definition, 1! = 1 and 0! = 1.

Permutations of *n* different things taken r at a time

The number of permutations of n different objects, taken r at a time, is denoted by ${}^nP_r$. To obtain a formula for ${}^nP_r$, we note that the first object can be chosen in n different ways. Following this, the second object can be chosen in $(n - 1)$ different ways and following this, the third object can be chosen in $(n - 2)$ different ways. Continuing in this manner, the r^{th} (i.e. the last) object can be chosen in $n - (r-1) = (n - r + 1)$ ways. Hence, by the multiplication principle (see Theorem 1.4 on page 8),

$$\begin{aligned} {}^nP_r &= n(n-1)(n-2)\ldots(n-r+1) \\ &= \frac{n(n-1)(n-2)\ldots(n-r+1)(n-r)\ldots(3)(2)(1)}{(n-r)\ldots(3)(2)(1)} \\ &= \frac{n!}{(n-r)!} \quad (r < n). \end{aligned}$$

Theorem 1.5

There are $n!/(n-r)!$ permutations of r objects chosen from n different objects.

Example 1.14

Two lottery tickets are drawn from 20 for a first and a second prize. In how many ways can this be done?

Solution

The total number of ways is

$$^{20}P_2 = \frac{20!}{(20-2)!} = \frac{20!}{18!} = \frac{20 \times 19 \times 18!}{18!} = 20(19) = 380.$$

Example 1.15

(a) Three different Mathematics books, four different French books, and two different Physics books, are to be arranged on a shelf. How many different arrangements are possible if
   (i) the books in each particular subject must all stand together,
   (ii) only the mathematics books must stand together?

(b) In how many ways can 9 people be seated on a bench if only 3 seats are available?

Solution

(a) (i) The mathematics books can be arranged among themselves in 3! = 6 ways, the French books in 4! = 24 ways, the Physics books in 2! = 2 ways, and the three groups in 3! = 6 ways. Hence, the required number of arrangements is $6 \times 24 \times 2 \times 6 = 1\,728$.

(ii) Consider the mathematics books as one big book. Then we have 7 books which can be arranged in 7! ways. In all of these arrangements, the mathematics books are together. But the mathematics books can be arranged among themselves in 3! ways. Hence, the required number of arrangements is 7!3! = 30 240.

(b) The first seat can be selected in 9 ways and when this has been done, there are 8 ways of selecting the second seat and 7 ways of selecting the third seat. Hence the required number of arrangements is 9(8)(7) =504.

Example 1.16

In how many ways can 6 boys and 2 girls be arranged in a row if

(a) the two girls are together,   (b) the two girls are not together.

Solution

(a) Since the girls are together, they can be considered as one unit. We then have 7 objects (6 boys and the unit of 2 girls) to be arranged. This can be done in 7! ways. In all these arrangements, the two girls are together. But the two girls can be arranged among themselves in 2! ways. Hence the required number of arrangments is 2! × 7! = 10 080.

(b) Number of arrangements without restriction = 8!.

Number of arrangements with the girls together = 2! × 7!.

Therefore the number of arrangements with the girls not together is

$(8! - 2!7!) = 8 \times 7! - 2!7! = 7!(8 - 2) = 7! \times 6 = 30\,240.$

Permutations with repetitions

So far we have considered permutations of distinct objects. That is, all the objects were completely different. The following theorem gives permutations with repetitions.

Theorem 1.6

Given a set of n objects having n_1 elements alike of one kind, and n_2 elements alike of another kind, and n_3 elements alike of a third kind, and so on for k kinds of objects, then the number of different arrangements of the n objects, taken all together is $\binom{n}{n_1, n_2, \ldots, n_k} = \dfrac{n!}{n_1!n_2!\ldots n_k!}$, where $n_1 + n_2 + \ldots + n_k = n$.

Example 1.17

How many different ways can 3 red, 4 yellow, and 2 blue bulbs be arranged in a string of Christmas tree lights with 9 sockets?

Solution

The total number of distinct arrangements is

$$\binom{9}{3, 4, 2} = \frac{9!}{3!4!2!} = 1\,260.$$

Circular permutations

Permutations that occur by arranging objects in a circle are called **circular permutations**. It can be shown that the number of permutations of n distinct objects arranged in a circle is $(n-1)!$.

Combinations

In many problems, we are interested in the number of ways of selecting r objects from n without regard to order. These selections are called ***combinations***. A combination is actually a partition with two cells, one cell containing the r objects selected and the other cell containing the $(n-r)$ objects that are left.

The number of such combinations, denoted by $\binom{n}{r,\ n-r}$, is usually shortened $\binom{n}{r}$ or ${}^{n}C_{r}$.

Theorem 1.7

The number of combinations of n distinct objects taken r at a time is

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

Example 1.18

From 4 chemists and 3 physicists, find the number of committees that can be formed consisting of 2 chemists and 1 physicist.

Solution

The number of ways of selecting 2 chemists from 4 is $\binom{4}{2} = \frac{4!}{2!\,2!} = 6.$

The number of ways of selecting 1 physicist from 3 is $\binom{3}{1} = \frac{3!}{1!\,2!} = 3.$

Using the multiplication theorem (see Theorem 1.4 on page 8) with $n_1 = 6$ and $n_2 = 3$, it can be seen that we can form $n_1 \times n_2 = 6 \times 3 = 18$ committees with 2 chemists and 1 physicist.

Exercise 1(b)

1. In how many ways can 3 of 20 laboratory assistants be chosen to assist with an experiment?
2. If an experiment consists of throwing a die and then drawing a letter at random from the English alphabet, how many points are in the sample space?
3. A certain shoe comes in 5 different styles with each style available in 4 distinct colours. If the store wishes to display pairs of these shoes showing all of its various style and colours, how many different pairs would the store have on display?
4. Find the number of ways 6 teachers can be assigned to 4 sections of an introductory psychology course if no teacher is assigned to more than one section.
5. How many distinct permutations can be made from the letters of the word INFINITY?
6. In how many ways can 5 different trees be planted in a circle?
7. How many ways are there to select 3 candidates from 8 equally qualified recent graduates for openings in an accounting firm?
8. A football team plays 12 games during a season. In how many ways can the team end the season with 7 wins, 3 losses, and 2 ties?
9. (a) In how many ways can 10 students line up at the bursar's window?
   (b) How many 3-letter words can be formed from the letters of the English alphabet if no letter can be used more than once?

10. Out of 6 mathematicians and 8 physicists, a committee consisting of 3 mathematicians and 4 physicists is to be formed. In how many ways can this be done if
    (a) any mathematician and any physicist can be selected,
    (b) two particular mathematicians cannot be on the committee,
    (c) 2 particular physicists must be selected?

11. Out of seven women and nine men, a committee consisting of three women and four men is to be formed. In how many ways can this be done if one particular man must be on the committee?

1.4 The probability of an event

1.4.1 Introduction

It is frequently useful to quantify the likelihood, or chance, that an outcome of a random experiment will occur. For example, we may hear a physician say that a patient has 50-50 chance of surviving a certain operation. Another physician may say that she is 95% certain that a patient has a particular disease. A public health nurse may say that 80% of certain clients will break an appointment. As these examples suggest, most people express probabilities in terms of percentages. However, in dealing with probabilities mathematically, it is convenient to express probabilities as fractions. Thus, we measure the probability of the occurrence of some event by a number between zero and one. The more likely the event, the closer the number is to one; and the more unlikely the event, the closer the number is to zero.

1.4.2 Classical probability

The classical treatment of probability dates back to the 17th century and the work of two mathematicians, Pascal and Fermat [see Todhunter (1931) and David (1962)]. Much of this theory developed out of attempts to solve problems related to games of chance, such as those involving the rolling of dice [see Jeffreys (1939), Ore (1960), and Keynes (1921)]. We can calculate the probability of an event in the classical sense as follows.

Definition 1.1

If a trial of an experiment can result in m mutually exclusive and equally likely outcomes, and if exactly h of these outcomes correspond to an event A, then the probability of event A is given by

$$P(A) = \frac{h}{m} = \frac{\text{number of ways that } A \text{ can occur}}{\text{number of ways the sample space } S \text{ can occur}}.$$

Thus, if all the simple events in S are equally likely, then

$$P(A) = \frac{n(A)}{n(S)} \text{ for all } A \subset S,$$

where $n(A)$ denotes the number of elements in A. We emphasize that the above expression for $P(A)$ is applicable only when all the simple events in S are equally likely.

It is important to realize that here, we are using the same symbol A to represent two different things. In the expression $n(A)$, A represents a set (for example, the set of even integers less than 7) whereas when we write $P(A)$, A represents an event (for example, the score on a die is even).

Example 1.19

A mixture of candies contains 6 mints, 4 toffees, and 3 chocolates. If a person makes a random selection of one of these candies, find the probability of getting
(a) a mint, (b) a toffee or a chocolate.

Solution

Let M, T, and C represent the events that the person selects, respectively, a mint, toffee, or chocolate candy. The total number of candies is 13, all of which are equally likely to be selected.

(a) Since 6 of the 13 candies are mints,

$$P(M) = \tfrac{6}{13}.$$

(b) $$P(T \cup C) = \frac{n(T \cup C)}{n(S)} = \frac{7}{13}.$$

Example 1.20

The following table shows 100 patients classified according to blood group and sex.

	Blood group		
	A	***B***	***O***
Male	30	20	17
Female	15	10	8

If a patient is selected at random from this group, find the probability that the patient selected:
(a) is a male and has blood group B, (b) is a female and has blood group A.

Solution

There are 100 ways in which we can select a patient from the 100 patients. Since the patient is selected at random, all the 100 ways of selecting a patient are equally likely.

(a) There are 20 males with blood group B. Therefore the probability that the patient selected is a male and has blood group B is $\frac{20}{100} = 0.2$.

(b) There are 15 females with blood group A. Therefore the probability that the patient selected is a female and has blood group A is $\frac{15}{100} = 0.15$.

One advantage of the classical definition of probability is that it does not require experimentation. Furthermore, if the outcomes are truely equally likely, then the probability assigned to an event is not an approximation. It is an accurate description of the frequency with which the event will occur.

1.4.3 Relative frequency probability

The relative frequency approach was developed by Fisher (1921) and Von Mises (1941), and depends on the repeatability of some process and the ability to count the number of repetitions, as well as the number of times that some event of interest occurs. In this context, we may define the probability of observing some characteristic, A, of an event as follows:

Definition 1.2

If some process is repeated a large number of times n, and if some resulting event with the characteristic A occurs m times, the relative frequency of occurrence of A, m/n, will be approximately equal to the probability of A. Thus,

$$P(A) = \lim_{n \to \infty} \frac{\mathrm{n}(A)}{n}.$$

The disadvantage in this approach is that the experiment must be repeatable. Remember that any probability obtained this way is an approximation. It is a value based on n trials. Further testing might result in a different approximate value.

Example 1.21

The following table gives the frequency distribution of the heights of 150 students. If a student is selected at random from this group, find the probability that the student selected is taller than the modal height of the students.

Height(cm)	130	140	150	160	170	180	190
Frequency	8	16	28	44	33	17	4

Solution

The modal height of the students is 160 cm. This is the height with the highest frequency. The number of students who are taller than 160 cm is (33 + 17 + 4) = 54. An estimate of the required probability is the relative frequency

$$\frac{54}{150} = 0.36.$$

1.4.4 Subjective probability

In the early 1950s, Savage (1972) gave considerable impetus to what is called **subjective** concept of probability. This view holds that probability measures the confidence that a particular individual has in the truth of a particular proposition. This concept does not depend on the repeatability of any process. By applying this concept of probability, one can calculate the probability of an event that can only happen once, for example, the probability that a cure for HIV/AIDS will be discovered within the next 8 years.

Although the subjective view of probability has enjoyed increased attention over the years, it has not been fully accepted by statisticians who have traditional orientations.

1.5 Some probability laws

In the last section, we considered how to interpret probabilities. In this section, we consider some laws that govern their behaviour.

1.5.1 Axioms of probability

In 1933, the axiomatic approach to probability was formalized by the Russian mathematician A. N. Kolmogorov (1964). The basis of this approach is embodied in three axioms from which a whole system of probability theory is constructed through the use of mathematical logic. The three axioms are as follows.

Axioms of probability

Let S be the sample space of an experiment and P be a set function which assigns a number $P(A)$ to every $A \subset S$. Then the function $P(A)$ is said to be a probability function if it satisfies the following three axioms:

Axiom 1: $P(S) = 1.$

Axiom 2: $P(A) \geq 0$ for every event A.

Axiom 3: If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$.

1.5.2 Elementary Theorems

The following theorems can be derived from the axioms of probability.

Theorem 1.8

$P(\phi) = 0.$

Proof

$S \cup \phi = S$ and S and ϕ are mutually exclusive. Hence by Axiom 3,

$$P(S) + P(\phi) = P(S)$$
$$\Rightarrow \quad P(\phi) = 0.$$

Theorem 1.9:

$P(A') = 1 - P(A).$

Proof

$A \cup A' = S$ and A and A' are mutually exclusive. Hence by Axiom 3,

$$P(A) + P(A') = P(S)$$
$$= 1 \text{ (by Axiom 1)}$$
$$\Rightarrow \quad P(A') = 1 - P(A).$$

Sometimes it is more difficult to calculate the probability that an event occurs than it is to calculate the probability that the event does not occur. Should this be the case for some event A, we simply find $P(A')$ first and then, using Theorem 1.9, find $P(A)$ by subtraction.

Theorem 1.10

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Proof

$A \cup B$ can be decomposed into two mutually exclusive events, $A \cap B'$ and B (see Fig. 1.7). Thus,

$$A \cup B = (A \cap B') \cup B$$

Applying Axiom 3, we obtain

$$P(A \cup B) = P(A \cap B') + P(B)$$

$$\Rightarrow P(A \cap B') = P(A \cup B) - P(B) \quad \text{......................(1.1)}$$

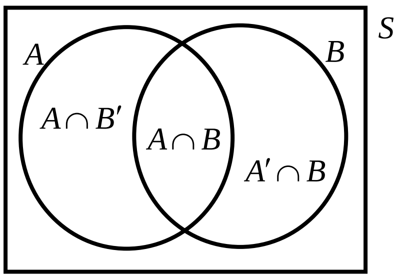

Fig. 1.7: *A Venn diagram for Theorem 1.10*

We can also decompose A into two mutually exclusive events $A \cap B'$ and $A \cap B$. Thus,

$$A = (A \cap B') \cup (A \cap B). \quad \text{......................(1.2)}$$

Applying Axiom 3, we obtain

$$P(A) = P(A \cap B') + P(A \cap B). \quad \text{......................(1.3)}$$

Using Equations (1.1) and (1.3), we obtain

$$P(A) = \{P(A \cup B) - P(B)\} + P(A \cap B),$$

giving $P(A \cup B) = P(A) + P(B) - P(A \cap B).$

Theorem 1.10 is often called the **addition rule of probability**.

Corollary 1.1

If the events A and B are mutually exclusive, then $A \cap B = \phi$ and so by Theorem 1.8, $P(A \cap B) = 0$. Theorem 1.10 then becomes

$$P(A \cup B) = P(A) + P(B).$$

Corollary 1.1 can be extended by mathematical induction to the following corollary.

Corollary 1.2

If the events $A_1, A_2, ..., A_n$ are mutually exclusive, then;

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n).$$

The following corollary gives an extension of Theorem 1.10 to 3 events.

Corollary 1.3

If A, B, and C are three events defined on the same sample space, then;

$$\begin{aligned} P(A \cup B \cup C) = {} & P(A) + P(B) + P(C) - P(A \cap B) \\ & - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

Proof

Let $B \cup C = D$. Then,

$$\begin{aligned} P(A \cup B \cup C) &= P(A \cup D) = P(A) + P(D) - P(A \cap D) \\ &= P(A) + \{P(B) + P(C) - P(B \cap C)\} - P\{(A \cap B) \cup (A \cap C)\} \\ &= P(A) + P(B) + P(C) - P(B \cap C) - \{P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)\} \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

Example 1.22

The probability that Akosua passes Mathematics is $\frac{2}{3}$, and the probability that she passes English is $\frac{4}{9}$. If the probability that she passes both courses is $\frac{1}{4}$, what is the probability that she passes at least one of the two courses?

Solution

Let M denote the event "Akosua passes Mathematics" and E the event "Akosua passes English". We wish to find $P(M \cup E)$. By the addition rule of probability, (see Theorem 1.10 on page 17),

$$P(M \cup E) = P(M) + P(E) - P(M \cap E) = \frac{2}{3} + \frac{4}{9} - \frac{1}{4} = \frac{31}{36}.$$

Example 1.23

Refer to Example 1.20 on page 14. If a patient is selected at random from the 100 patients, find the probability that the patient selected:

(a) is a male or has blood group A,

(b) does not have blood group A,

(c) is a female or does not have blood group B.

Solution

There are 100 ways in which we can select a patient from the 100 patients. Since the patient is selected at random, all the 100 ways of selecting a patient are equally likely.

(a) Let M denote the event "a patient is a male" and A the event "a patient has blood group A". We wish to find $P(M \cup A)$. By the addition rule of probability,

$$\begin{aligned} P(M \cup A) &= P(M) + P(A) - P(M \cap A) \\ &= \frac{n(M)}{100} + \frac{n(A)}{100} - \frac{n(M \cap A)}{100} \\ &= \frac{67}{100} + \frac{45}{100} - \frac{30}{100} = \frac{82}{100} = 0.82. \end{aligned}$$

(b) We wish to find $P(A')$. By Theorem 1.9,

$$P(A') = 1 - P(A) = 1 - \frac{n(A)}{100} = 1 - \frac{45}{100} = 0.55.$$

(c) Let F denote the event "a patient selected is a female" and B the event "a patient selected has blood group B". We wish to find $P(F \cup B')$. By the addition rule of probability,

$$\begin{aligned} P(F \cup B') &= P(F) + P(B') - P(F \cap B') \\ &= P(F) + \{1 - P(B)\} - P(F \cap B') \\ &= \frac{33}{100} + \left(1 - \frac{30}{100}\right) - \frac{23}{100} = 1 - \frac{20}{100} = 0.8. \end{aligned}$$

Example 1.24

Of 200 students in a certain Senior High School, 60 study Mathematics, 40 study Biology, 30 study Chemistry, 10 study Mathematics and Biology, 5 study Mathematics and Chemistry, 3 study Biology and Chemistry and 1 studies the three subjects. If a student is selected at random from this group, find the probability that the student selected studies at least one of the three subjects.

Solution

Let S = {the 200 students}, M = {those who study Mathematics}, B = {those who study Biology} and C = {those who study Chemistry}. Then, $n(M) = 60$, $n(B) = 40$, $n(C) = 30$, $n(M \cap B) = 10$, $n(M \cap C) = 5$, $n(B \cap C) = 3$ and $n(M \cap B \cap C) = 1$. We are required to calculate $P(M \cup B \cup C)$. By Corollary 1.3,

$$\begin{aligned} P(M \cup B \cup C) &= P(M) + P(B) + P(C) - P(M \cap B) \\ &\qquad - P(M \cap C) - P(B \cap C) + P(M \cap B \cap C) \\ &= \frac{60}{200} + \frac{40}{200} + \frac{30}{200} - \frac{10}{200} - \frac{5}{200} - \frac{3}{200} + \frac{1}{200} = \frac{113}{200}. \end{aligned}$$

Example 1.25

If the probabilities are, respectively, 0.08, 0.14, 0.22, and 0.24 that a person buying a new car will choose the colour green, white, red, or black, calculate the probability that a given buyer will purchase a new car that comes in one of these colours.

Solution

Let G, W, R, and B be the events that a buyer selects, respectively, a green, white, red or black car. Since the four colours are mutually exclusive, the required probability is

$$P(G \cup W \cup R \cup B) = P(G) + P(W) + P(R) + P(B)$$
$$= 0.08 + 0.14 + 0.22 + 0.24 = 0.68.$$

1.5.3 Two-set problems

If A and B are any two events defined on a sample space S, then we can draw Fig. 1.8. It can be seen that S can be split into the following four mutually exclusive events:
$A \cap B$, $A' \cap B$, $A \cap B'$ and $A' \cap B'$.

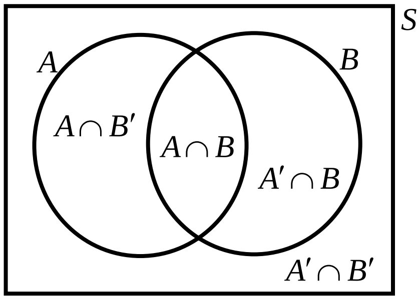

***Fig. 1.8**: Two-set problems*

Notice that:

$$A = (A \cap B') \cup (A \cap B)$$

Since $A \cap B'$ and $A \cap B$ are mutually exclusive,

$$P(A) = P(A \cap B') + P(A \cap B) \quad \text{.....(1.4)}$$

Similarly,

$$P(B) = P(A' \cap B) + P(A \cap B) \quad \text{.....(1.5)}$$

Moreover,

$$(A \cap B') \cup (A \cap B) \cup (A' \cap B) \cup (A' \cap B') = S$$

and since the four events are mutually exclusive,

$$P(A \cap B') + P(A \cap B) + P(A' \cap B) + P(A' \cap B') = P(S) = 1.$$

Example 1.26

The probability that a new airport will get an award for its design is 0.04, the probability that it will get an award for the efficient use of materials is 0.2 and the probability that it will get both awards is 0.03. Find the probability that it will get:

(a) at least one of the two awards,
(b) only one of the two awards,
(c) none of the two awards.

Solution

Let D denote the event "the airport will get an award for its design", and E the event "the airport will get an award for the efficient use of materials".

We are given that $P(D) = 0.04$, $P(E) = 0.2$ and $P(D \cap E) = 0.03$. We can therefore draw Fig. 1.9. Notice that, since $P(D) = 0.04$, and $P(D \cap E) = 0.03$, $P(D \cap E') = 0.04 - 0.03 = 0.01$.

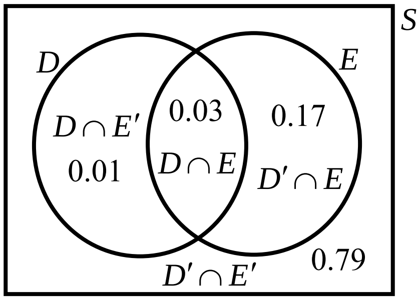

***Fig. 1.9**: A Venn diagram for Example 1.26*

(a) We wish to find $P(D \cup E)$. From Fig. 1.9,

$$P(D \cup E) = 0.17 + 0.03 + 0.01 = 0.21$$

Alternatively,

$$P(D \cup E) = P(D) + P(E) - P(D \cap E)$$
$$= 0.04 + 0.2 - 0.03 = 0.21.$$

(b) The probability that it will get only one of the awards is

$$P(D \cap E') + P(D' \cap E) = 0.01 + 0.17 = 0.18.$$

(c) We wish to find $P(D' \cap E')$. From Fig. 1.9,

$$P(D' \cap E') = 0.79.$$

Alternatively, using Theorem 1.9, we obtain

$$\begin{aligned} P(D' \cap E') &= 1 - P\left[(D' \cap E')'\right] \\ &= 1 - P(D \cup E) = 1 - 0.21, \quad \text{(from part (a))} \\ &= 0.79. \end{aligned}$$

Example 1.27

The events A, B and C have probabilities $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ and $P(C) = \frac{1}{4}$. Furthermore, $A \cap C = \phi$, $B \cap C = \phi$ and $P(A \cap B) = \frac{1}{6}$. Find:

(a) $P[(A \cap B)']$, (b) $P(A \cap B')$, (c) $P[(A \cup B)']$,
(d) $P(A' \cap B')$, (e) $P(A \cup B \cup C)$.

Solution

(a) Applying Theorem 1.9, we obtain

$$P[(A \cap B)'] = 1 - P(A \cap B) = 1 - \frac{1}{6} = \frac{5}{6}.$$

(b) $$\begin{aligned} P(A \cap B') &= P(A) - P(A \cap B), \quad \text{(see Equation (1.4) on page 20)} \\ &= \frac{1}{2} - \frac{1}{6} = \frac{1}{3}. \end{aligned}$$

(c) Applying Theorem 1.9, we obtain

$$\begin{aligned} P[(A \cup B)'] &= 1 - P(A \cup B) = 1 - \{P(A) + P(B) - P(A \cap B)\} \\ &= 1 - \left(\frac{1}{2} + \frac{1}{3} - \frac{1}{6}\right) = 1 - \frac{2}{3} = \frac{1}{3}. \end{aligned}$$

(d) $$\begin{aligned} P(A' \cap B') &= P[(A \cup B)'], \quad \text{(see Theorem 1.1 on page 6)} \\ &= \frac{1}{3}, \quad \text{(from part (c))}. \end{aligned}$$

(e) Applying Corollary 1.3, we obtain

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \\ &= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} - \frac{1}{6} - 0 - 0 + 0 = \frac{11}{12}. \end{aligned}$$

Exercise 1(c)

1. A die is loaded in such a way that an even number is twice as likely to occur as an odd number. Find the probability that a number less than 4 occurs on a single toss of the die.

2. In Question 1, let A be the event that an even number turns up and let B be the event that a number divisible by 3 occurs. Find: (a) $P(A \cup B)$, (b) $P(A \cap B)$.

3. If $A \subset B$, prove that $P(A) \leq P(B)$. Hint: $B = A \cup (A' \cap B)$.

4. Prove that the probability of any event A is at most 1. Hint: $A \subseteq S$.

5. Let A and B be events with $P(A) = 0.25$, $P(B) = 0.40$ and $P(A \cap B) = 0.15$. Find
   (a) $P(A' \cap B')$, (b) $P(A \cap B')$, (c) $P(A' \cap B)$.

6. A certain carton of eggs has 3 bad and 9 good eggs.
   (a) If an omelette is made of 3 eggs randomly chosen from the carton, what is the probability that there are no bad eggs in the omelette?
   (b) What is the probability of having exactly 2 bad eggs in the omelette?

7. Samples of a cast aluminum part are classified on the basis of surface finish (in microinches) and length measurements. The results of 100 parts are summarized below.

		length	
		excellent	good
surface finish	excellent	75	7
	good	10	8

   Let A denote the event that a sample has excellent surface finish, and let B denote the event that a sample has excellent length. Find
   (a) $P(A)$, (b) $P(B)$, (c) $P(A')$, (d) $P(A \cap B)$, (e) $P(A \cup B)$.

8. Samples of foam from two suppliers are classified for conformance to specifications. The results from 40 samples are summarized below

		Conforms	
		yes	no
Supplier	1	18	2
	2	17	3

   Let A denote the event that a sample is from supplier 1, and let B denote the event that a sample conforms to specifications. Find
   (a) $P(A)$, (b) $P(B)$, (c) $P(A')$, (d) $P(A \cap B)$, (e) $P(A \cup B)$, (f) $P(A' \cap B')$.

9. In a certain population of women, 4% have breast cancer, 20% are smokers and 3% are both smokers and have breast cancer. If a woman is selected at random from this population, find the probability that the person selected is:
   (a) a smoker or has breast cancer,
   (b) a smoker and does not have breast cancer,
   (c) not a smoker and does not have breast cancer.

1.6 Conditional probability

A box contains n white balls and m red balls. All the balls are of the same size. Without looking, a person takes a ball from the box. Without replacing the ball in the box, he then takes a second ball. Let B denote the event "the first ball drawn is white" and A the event "the second ball drawn is red". For the first draw, any of the $(m + n)$ balls in the box is equally likely to be drawn, while n of these are white. Therefore,

$$P(B) = \frac{n}{m+n}.$$

If the first ball drawn was white, then the probability that the second ball drawn is red is $\frac{m}{(m+n-1)}$, since only m of the remaining $(m+n-1)$ balls are red.

We introduce a new notation to describe this probability. We call it the ***conditional probability*** of event A given B. We denote it by $P(A|B)$. This is usually read as "the probability of A, given B". Thus, in the above example,

$$P(A \mid B) = \frac{m}{m+n-1}.$$

We now give a formal definition of conditional probability.

Definition 1.3

If A and B are any two events defined on the same sample space S, the conditional probability of A given B, is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0 \quad \ldots\ldots(1.6)$$

and if $P(B) = 0$, then $P(A \mid B)$ is undefined.

In particular, if S is a finite, equiprobable sample space (see page 3), then:

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)}, \quad P(B) = \frac{n(B)}{n(S)}, \text{ and so}$$

$$P(A|B) = \left\{\frac{n(A \cap B)}{n(S)}\right\} \bigg/ \left\{\frac{n(B)}{n(S)}\right\} = \frac{n(A \cap B)}{n(B)} \quad \ldots\ldots(1.7)$$

Example 1.28

Consider the data given in Example 1.20 on page 14. If a patient is chosen at random from the 100 patients, find the probability that the patient chosen has blood group A given that he is a male.

Solution

Let M denote the event "the patient chosen is a male" and A the event "the patient chosen has blood group A". We wish to find $P(A|M)$.

$$P(A|M) = \frac{P(A \cap M)}{P(M)} = \frac{30/100}{67/100} = \frac{30}{67}.$$

Alternatively, since the sample space is finite and equiprobable,

$$P(A|M) = \frac{n(A \cap M)}{n(M)} = \frac{30}{67}.$$

Example 1.29

The probability that a regularly scheduled flight departs on time is $P(D) = 0.83;$ the probability that it arrives on time is $P(A) = 0.82;$ and the probability that it departs and arrives on time is $P(D \cap A) = 0.78.$ Find the probability that a plane:

(a) arrives on time given that it departed on time,
(b) departed on time given that it has arrived on time,
(c) arrives on time, given that it did not depart on time.

Solution

(a) The probability that a plane arrives on time given that it departed on time is

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

(b) The probability that a plane departed on time given that it has arrived on time is

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

(c) We wish to find $P(A|D')$. Now,

$$P(A|D') = \frac{P(A \cap D')}{P(D')} = \frac{P(A) - P(A \cap D)}{1 - P(D)} = \frac{0.82 - 0.78}{1 - 0.83} = 0.24.$$

The notion of conditional probability provides the capability of re-evaluating the idea of probability of an event in the light of additional information, that is, when it is known that another event has occurred. The probability $P(A|B)$ is an updating of $P(A)$ based on the knowledge that event B has occurred.

1.7 The multiplication rule

If we multiply each side of Equation (1.6) by $P(B),$ we obtain the following multiplication rule, which enables us to calculate the probability that two events will both occur.

If in an experiment, the events A and B can both occur, then

$$P(A \cap B) = P(B)P(A|B) \quad \text{(1.8)}$$

Since the events $A \cap B$ and $B \cap A$ are equivalent, it follows from Equation (1.8) that we can also write

$$P(A \cap B) = P(B \cap A) = P(A)P(B|A) \quad \text{(1.9)}$$

In other words, it does not matter which event is referred to as A and which event is referred to as B.

Example 1.30

Suppose that we have a fuse box containing 20 fuses, of which 5 are defective. If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

Solution

Let A denote the event that the first fuse is defective and B the event that the second fuse is defective. We wish to find $P(A \cap B)$. The probability of first removing a defective fuse is $\frac{5}{20}$. If the first fuse is defective, then the probability of removing a second defective fuse from the remaining 4 is $\frac{4}{19}$. By the multiplication rule,

$$P(A \cap B) = P(A)P(B|A) = \left(\frac{5}{20}\right)\left(\frac{4}{19}\right) = \frac{1}{19}.$$

Example 1.31

Bag 1 contains 4 white balls and 3 green balls, and bag 2 contains 3 white balls and 5 green balls. A ball is drawn from bag 1 and placed unseen in bag 2. Find the probability that a ball now drawn from bag 2 is (a) green, (b) white.

Solution

Let G_1, G_2, W_1 and W_2 represent, respectively, the events of drawing a green ball from bag 1, a green ball from bag 2, a white ball from bag 1 and a white ball from bag 2.

(a) We wish to find $P(G_2)$. We can express G_2 in the form $G_2 = (G_1 \cap G_2) \cup (W_1 \cap G_2)$. The events $(G_1 \cap G_2)$ and $(W_1 \cap G_2)$ are mutually exclusive and so

$$\begin{aligned} P(G_2) &= P(G_1 \cap G_2) + P(W_1 \cap G_2) \\ &= P(G_1)P(G_2|G_1) + P(W_1)P(G_2|W_1). \end{aligned}$$

Using the tree diagram in Fig. 1.10, we obtain

$$P(G_2) = \left(\frac{3}{7}\right)\left(\frac{6}{9}\right) + \left(\frac{4}{7}\right)\left(\frac{5}{9}\right) = \frac{38}{63}.$$

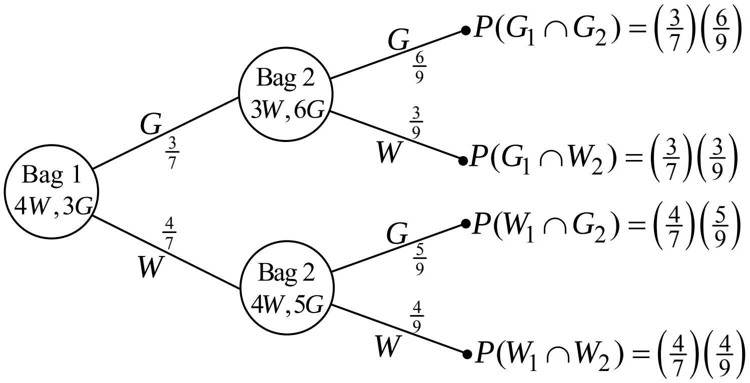

***Fig. 1.10**: Tree diagram for Example 1.31*

(b) We wish to find $P(W_2)$. We can express W_2 in the form $W_2 = (G_1 \cap W_2) \cup (W_1 \cap W_2)$. The events $G_1 \cap W_2$ and $W_1 \cap W_2$ are mutually exclusive and so

$$\begin{aligned} P(W_2) &= P(G_1 \cap W_2) + P(W_1 \cap W_2) \\ &= P(G_1)P(W_2 | G_1) + P(W_1)P(W_2 | W_1) \\ &= \tfrac{3}{7} \times \tfrac{3}{9} + \tfrac{4}{7} \times \tfrac{4}{9} = \tfrac{25}{63}. \end{aligned}$$

The multiplication rule can be applied to two or more events. For three events *A*, *B* and *C*, the multiplication rule takes the following form.

Theorem 1.11

$P(A \cap B \cap C) = P(A)\, P(B | A)\, P(C | A \cap B)$, where $P(A) \neq 0$ and $P(A \cap B) \neq 0$.

Proof

By the associative law,

$$\begin{aligned} A \cap B \cap C &= (A \cap B) \cap C. \text{ Therefore,} \\ P(A \cap B \cap C) &= P[(A \cap B) \cap C] \\ &= P(A \cap B)\, P(C | A \cap B) \\ &= P(A)\, P(B | A)\, P(C | A \cap B). \end{aligned}$$

Theorem 1.11 can be extended by mathematical induction to the following theorem.

Theorem 1.12

For any events $A_1, A_2, ..., A_n \ (n > 2)$

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2)...P(A_n | A_1 \cap A_2 ... \cap A_{n-1}).$$

Example 1.32

A box contains 5 red, 4 white and 3 blue balls. If three balls are drawn successively from the box, find the probability that they are drawn in the order red, white and blue if each ball is not replaced.

Solution

Let *R* be the event "red on first draw", *W* the event "white on second draw" and *B* the event "blue on third draw". We wish to find $P(R \cap W \cap B)$. Since there are 5 balls out of 12 balls, $P(R) = \frac{5}{12}$ If the first ball drawn is red, then there are 4 white balls out of the 11 balls remaining in the box. Hence $P(W | R) = \frac{4}{11}$. If the first ball is red and the second ball is white, then there are 3 blue balls out of the

10 balls remaining in the box. It follows that $P(B|R\cap W)=\frac{3}{10}$. Hence by Theorem 1.11,

$$P(R\cap W\cap B)=\left(\frac{5}{12}\right)\left(\frac{4}{11}\right)\left(\frac{3}{10}\right)=\frac{1}{22}.$$

1.8 Independent events

In Example 1.30, if the first fuse is replaced and the fuses are thoroughly rearranged before the second fuse is removed, then

$$P(B|A) = \frac{5}{20} = P(B).$$

This means that the occurrence or non-occurrence of event A does not affect the probability that the event B occurs. The event A is said to be **independent** of the event B. It can be proved that, if an event A is independent of an event B, then the event B is also independent of the event A. We therefore simply say that A and B are independent events.

If the events A and B are independent, then the multiplication theorem becomes $P(A\cap B)=P(A)P(B)$. This result illustrates the following general result.

Definition 1.4

Two events with nonzero probabilities are independent if and only if, any one of the following equivalent statements is true.

(a) $P(A|B)=P(A)$, (b) $P(B|A)=P(B)$, (c) $P(A\cap B)=P(A)P(B)$.

Two events that are not independent are said to be dependent. Usually, physical conditions under which an experiment is performed will enable us to decide whether or not two or more events are independent. In particular, the outcomes of unrelated parts of an experiment can be treated as independent.

Example 1.33

A small town has one fire engine and one ambulance available for emergencies. The probability that a fire engine is available when needed is 0.96, and the probability that the ambulance is available when called is 0.90. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available.

Solution

Let A and B represent the respective events that the fire engine and the ambulance are available. The two events are independent and so

$$P(A\cap B) = P(A)P(B) = (0.96)(0.90) = 0.864.$$

Example 1.34

A pair of fair dice is thrown twice. Find the probability of getting totals of 7 and 11.

Solution

Let A_1, A_2, B_1 and B_2 be the respective events that a total of 7 occurs on the first throw, a total of 7 occurs on the second throw, a total of 11 occurs on the first throw and a total of 11 occurs on the second throw. We are interested in the probability of the event $(A_1 \cap B_2) \cup (A_2 \cap B_1)$. It is clear that the events A_1, A_2, B_1 and B_2 are independent. Moreover, $A_1 \cap B_2$ and $A_2 \cap B_1$ are mutually exclusive events. Hence,

$$\begin{aligned} P[(A_1 \cap B_2) \cup (A_2 \cap B_1)] &= P(A_1 \cap B_2) + P(A_2 \cap B_1) \\ &= P(A_1)P(B_2) + P(A_2)P(B_1) \\ &= \left(\tfrac{6}{36}\right)\left(\tfrac{2}{36}\right) + \left(\tfrac{6}{36}\right)\left(\tfrac{2}{36}\right) = \tfrac{1}{54}. \end{aligned}$$

Occasionally, we must deal with more than two events. Again, the question arises: When are these events considered independent? The following definition answers this question by extending our previous definition to include more than two events.

Definition 1.5

Let $A_1, A_2, \ldots, A_n$ be a finite collection of events $(n \geq 3)$. These events are mutually independent if and only if the probability of the intersection of any 2, 3, …, n of these events is equal to the product of their respective probabilities.

For $n = 3$ events, Definition 1.5 reduces to the following definition.

Definition 1.6

The three events A_1, A_2, and A_3 are independent if and only if:

$$P(A_1 \cap A_2) = P(A_1)P(A_2), \qquad P(A_1 \cap A_3) = P(A_1)P(A_3),$$
$$P(A_2 \cap A_3) = P(A_2)P(A_3), \qquad P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

That is, three events A_1, A_2, and A_3 are independent if and only if they are pairwise independent, and $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$. It can be proved that if the above equations are satisfied, so is any equation we obtain by replacing an event by its complement on both sides of one of the original equations. For instance,

$$P(A_1 \cap A_2' \cap A_3) = P(A_1)P(A_2')P(A_3).$$

We may also replace any two or three events by their complements on both sides of the equations and obtain a true result. For instance,

$$P(A_1' \cap A_2') = P(A_1')P(A_2'), \qquad P(A_1' \cap A_2 \cap A_3') = P(A_1')P(A_2)P(A_3').$$

Example 1.35

The probability that Kofi hits a target is $\frac{1}{4}$ and the corresponding probabilities for Kojo and Kwame are $\frac{1}{3}$ and $\frac{2}{5}$, respectively. If they all fire together, find the probability that

(a) they all miss,
(b) exactly one shot hits the target,
(c) at least one shot hits the target,
(d) Kofi hits the target given that exactly one hit is registered.

Solution

Let A, B and C denote the events that the target is hit by Kofi, Kojo, and Kwame, respectively. Then $P(A)=\frac{1}{4}$, $P(B)=\frac{1}{3}$ and $P(C)=\frac{2}{5}$. Since they fire together, the events A, B and C are independent .

(a) If E is the event "they all miss", then $E = A'\cap B'\cap C'$,

and so
$$\begin{aligned} P(E) &= P(A'\cap B'\cap C') \\ &= P(A')P(B')P(C'), \text{ since } A, B \text{ and } C \text{ are independent} \\ &= \left(1-\tfrac{1}{4}\right)\left(1-\tfrac{1}{3}\right)\left(1-\tfrac{2}{5}\right) = \tfrac{3}{10}. \end{aligned}$$

(b) Let F denote the event "exactly one shot hits the target". Then,

$$F = (A\cap B'\cap C')\cup(A'\cap B\cap C')\cup(A'\cap B'\cap C).$$

In other words, if exactly one shot hits the target, then it was either only Kofi, $A\cap B'\cap C'$, or only Kojo, $A'\cap B\cap C'$, or only Kwame, $A'\cap B'\cap C$. Since the three events are mutually exclusive, we obtain

$$\begin{aligned} P(F) &= P(A\cap B'\cap C')+P(A'\cap B\cap C')+P(A'\cap B'\cap C) \\ &= P(A)P(B')P(C')+P(A')P(B)P(C')+P(A')P(B')P(C), \text{ by independence} \\ &= \tfrac{1}{4}\left(1-\tfrac{1}{3}\right)\left(1-\tfrac{2}{5}\right)+\left(1-\tfrac{1}{4}\right)\left(\tfrac{1}{3}\right)\left(1-\tfrac{2}{5}\right)+\left(1-\tfrac{1}{4}\right)\left(1-\tfrac{1}{3}\right)\left(\tfrac{2}{5}\right) \\ &= \frac{1}{10}+\frac{3}{20}+\frac{1}{5} = \frac{9}{20}. \end{aligned}$$

(c) Let G denote the event "at least one shot hits the target". Then, $G = A\cup B\cup C$. By de Morgan's law (see Theorem 1.3), $(A\cup B\cup C)' = A'\cap B'\cap C'$.

Hence, by Theorem 1.9,

$$\begin{aligned} P(G) &= 1-P(A'\cap B'\cap C') \\ &= 1-\tfrac{3}{10} = \tfrac{7}{10}, \text{ (from part (a)).} \end{aligned}$$

(d) We wish to find $P(A\,|\,F)$, the probability that Kofi hits the target given that exactly one hit is registered.

Now, $P(A|F) = \dfrac{P(A \cap F)}{P(F)}$ and $A \cap F = A \cap B' \cap C'$.

Therefore, $P(A|F) = \dfrac{P(A \cap B' \cap C')}{P(F)} = \dfrac{1/10}{9/20} = \frac{2}{9}$, (from part (b)).

Exercise 1(d)

1. In a space lot, the primary computer system is backed up by two secondary systems. They operate independently of one another and each is 90% reliable. Find the probability that all three systems will be operable at the time of the launch.
2. A study of major flash floods that occurred over the last 15 years indicates that the probability that a flash flood warning will be issued is 0.5 and the probability of a dam failure during the flood is 0.33. The probability of a dam failure given that a warning is issued is 0.17. Find the probability that a flash flood warning will be issued and a dam failure will occur.
3. Show that if A_1 and A_2 are independent, then A_1 and A_2' are also independent. Hint: $A_1 = (A_1 \cap A_2) \cup (A_1 \cap A_2')$.
4. Kofi feels that the probability that he will get an A in the first Physics test is $\frac{1}{2}$ and the probability that he will get $A's$ in the first and second Physics tests is $\frac{1}{3}$. If Kofi is correct, what is the conditional probability that he will get an A in the second test, given that he gets an A in the first test?
5. In rolling 2 balanced dice, if the sum of the two values is 7, what is the probability that one of the values is 1?
6. A random sample of 200 adults are classified below by sex and their level of education attained.

Education	**Male**	**Female**
Elementary	38	45
High School	28	50
University	22	17

If a person is chosen at random from this group, find the probability that:
(a) the person is a male, given that the person has High School education,
(b) the person does not have a university degree, given that the person is a female.

7. In an experiment to study the relationship of hypertension and smoking habits, the following data were collected for 180 individuals.

	Non-smokers	Moderate smokers	Heavy smokers
Hypertension	21	36	30
No hypertension	48	26	19

If one of these individuals is selected at random, find the probability that the person is

(a) experiencing hypertension, given that he/she is a heavy smoker;

(b) a non-smoker, given that he/she is experiencing no hypertension.

8. The probability that a married man watches a certain television show is 0.4 and the probability that a married woman watches the show is 0.5. The probability that a man watches the show, given that his wife does is 0.7. Find the probability that

   (a) a married couple watches the show;

   (b) a wife watches the show given that her husband does;

   (c) at least 1 person of a married couple will watch the show.

9. A town has 2 fire engines operating independently. The probability that a specific engine is available when needed is 0.96.

   (a) What is the probability that neither is available when needed?

   (b) What is the probability that exactly one fire engine is available when needed?

1.9 Bayes' theorem

1.9.1 The total probability rule

In Fig. 1.11, the events A_1, A_2, A_3, A_4, A_5 are mutually exclusive and $S = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$. These events are said to form a **partition of the sample space** S. By a partition of S, we mean a collection of mutually exclusive events whose union is S. In general, the events A_1, A_2, …, A_n form a partition of the sample space S if

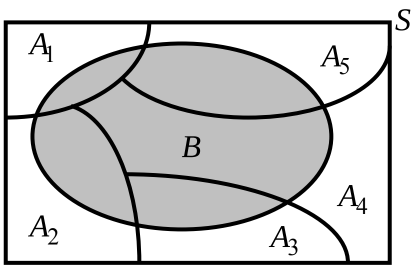

***Fig. 1.11**: A partition of a sample space*

(a) $A_i \neq \phi$ $(i = 1, 2, \ldots, n)$, (b) $A_i \cap A_j = \phi$ $(i \neq j,\ i, j = 1, 2, \ldots, n)$, (c) $S = \bigcup_{i=1}^{n} A_i$.

In Fig. 1.11, it can be seen that if B is an event defined on the sample space S such that $P(B) > 0$, then

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_5 \cap B).$$

Since $(A_1 \cap B), (A_2 \cap B), \ldots, (A_5 \cap B)$ are mutually exclusive events,

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_5 \cap B) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \ldots + P(A_5)P(B|A_5). \end{aligned}$$

The following theorem gives the general result.

Theorem 1.13

If A_1, A_2, …, A_n form a partition of a sample space S, then for any event B defined on S such that $P(B) > 0$,

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$

This result is called the **total probability rule.**

Example 1.36

In a certain assembly plant, three machines A, B, and C make 30%, 45% and 25%, respectively, of the products. It is known from past experience that 2% of the products made by machine A, 3% of the products made by machine B and 2% of the products made by machine C are defective. If a finished product is selected at random, what is the probability that it is defective?

Solution

Let A_1 denote the event "the finished product was made by machine A",

A_2 denote the event "the finished product was made by machine B",

A_3 denote the event "the finished product was made by machine C",

and let D denote the event "the finished product is defective". We wish to find $P(D)$. We are given that $P(A_1) = 0.3$, $P(A_2) = 0.45$, $P(A_3) = 0.25$, $P(D|A_1) = 0.02$, $P(D|A_2) = 0.03$ and $P(D|A_3) = 0.02$. A_1, A_2 and A_3 form a partition of the sample space. Hence,

$$\begin{aligned} P(D) &= P(A_1)P(D|A_1) + P(A_2)P(D|A_2) + P(A_3)P(D|A_3) \\ &= 0.3 \times 0.02 + 0.45 \times 0.03 + 0.25 \times 0.02 \\ &= 0.0245. \end{aligned}$$

The probability that a finished product selected at random is defective is 0.0245.

1.9.2 Bayes' theorem

Consider the following example.

Example 1.37

In Example 1.36, if a finished product is found to be defective, what is the probability that it was made by machine A_1?

Solution

We wish to find $P(A_1|D)$. By the multiplication rule,

$$P\left(A_1|D\right) = \frac{P(A_1 \cap D)}{P(D)} = \frac{P(A_1)P(D|A_1)}{P(D)} = \frac{(0.3)(0.02)}{0.0245} = 0.245.$$

The probability that a defective finished product was made by machine A_1 is 0.245.

Example 1.37 was solved by using **Bayes' theorem**. We now state the theorem.

Theorem 1.14 (**Bayes' theorem**)

Let $A_1, A_2, \ldots, A_n$ be a collection of events which partition a sample space S. Let B be an event defined on S such that $P(B) \neq 0$. Then for any of the events A_j, $(j = 1, 2, \ldots, n)$

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}, \quad \text{for } j = 1, 2, \ldots, n.$$

Proof

By the definition of conditional probability,

$$P\left(A_j|B\right) = \frac{P(A_j \cap B)}{P(B)}.$$

Using Theorem 1.13 in the denominator, we obtain

$$P\left(A_j|B\right) = \frac{P(A_j \cap B)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)},$$

which completes the proof.

Bayes' theorem was named after the English philosopher and theologian, Reverend Thomas Bayes (1702 – 1761). The theorem is applicable in situations where quantities of the form $P(B|A_i)$ and $P(A_i)$ are known and we wish to determine $P(A_i|B)$. The following example illustrates an application of Bayes' theorem.

Example 1.38

A consulting firm rents cars from three agencies: 30% from agency A, 20% from agency B and 50% from agency C. 15% of the cars from A, 10% of the cars from B and 6% of the cars from C have bad tyres. If a car rented by the firm has bad tyres, find the probability that it came from agency C.

Solution

Let A_1 denote the event "the car came from agency *A*",

A_2 denote the event "the car came from agency *B*",

A_3 denote the event "the car came from agency *C*",

and let T denote the event "a car rented by the firm has bad tyres". We wish to find $P(A_3|T)$. We are given $P(A_1)=0.3$, $P(A_2)=0.2$, $P(A_3)=0.5$, $P(T|A_1)=0.15$, $P(T|A_2)=0.1$ and $P(T|A_3)=0.06$. A_1, A_2 and A_3 are mutually exclusive and $P(A_1)+P(A_2)+P(A_3)=1$, and so A_1, A_2 and A_3 form a partition of the sample space. Hence, by Bayes' theorem,

$$\begin{aligned} P(A_3|T) &= \frac{P(A_3)P(T|A_3)}{P(A_1)P(T|A_1)+P(A_2)P(T|A_2)+P(A_3)P(T|A_3)} \\ &= \frac{0.5\times 0.06}{0.3\times 0.15\ +\ 0.2\times 0.1\ +\ 0.5\times 0.06} \\ &= 0.3158. \end{aligned}$$

Exercise 1(e)

1. A factory employs three machine operators, George, Andrew and Eric, to produce its brand of goods. George works 45% of the time, Andrew works 30% of the time and Eric works 25% of the time. Each operator is prone to produce defective items. George produces defective items 2% of the time, Andrew produces defective items 4% of the time while Eric produces defective items 6% of the time. If a defective item is produced, what is the probability that it was produced by Andrew?

2. In a certain assembly plant, three machines B_1, B_2 and B_3, make 30%, 45% and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by B_1, B_2 and B_3, respectively, are defective.
   (a) If a finished product is selected at random, what is the probability that it is defective?
   (b) If a finished product is found to be defective, what is the probability that it was produced by B_3?

3. A large industrial firm uses local hotels *A*, *B* and *C* to provide overnight accommodation for its clients. From past experience, it is known that 20% of the clients are assigned rooms at hotel *A*, 50% at hotel *B*, and 30% at hotel *C*. If the plumbing is faulty in 5% of the rooms at hotel *A*, in 4% of the rooms at hotel *B*, and in 8% of the rooms at hotel *C*, what is the probability that:
   (a) a client will be assigned a room with faulty plumbing?
   (b) a person with a room having faulty plumbing was assigned accommodation at hotel *B*?

4. Suppose that at a certain accounting office, 30%, 25% and 45% of the statements are prepared by Mr. George, Mr. Charles and Mrs. Joyce, respectively. These employees are very reliable.

Nevertheless, they are in error some of the time. Suppose that 0.01%, 0.005% and 0.003% of the statements prepared by Mr. George, Mr. Charles and Mrs. Joyce, respectively, are in error. If a statement from the accounting office is in error, what is the probability that it was prepared (caused) by Mr. George?

5. A certain construction company buys 20%, 30%, and 50% of their nails from hardware suppliers *A*, *B*, and *C*, respectively. Suppose it is known that 0.05%, 0.02% and 0.01% of the nails from *A*, *B*, and *C*, respectively, are defective.
   (a) What percentage of the nails purchased by the construction company are defective?
   (b) If a nail purchased by the construction company is defective, what is the probability that it came from the supplier *C*?

References

David, F. N. (1962). Games, Gods and Gambling. *Hafner*, *New York*.

Fisher, R. A. (1921). On the mathematical foundations of theoretical statistics. PTRS 222 : 309.

Jeffreys, H. (1939). Theory of probability. *Oxford*: *Clarendon*.

Keynes, J. M. (1921). A treatrise on probability. *Macmillan Co. Ltd., London*.

Kolmogorov, A. N. (1964). Foundation of the theory of Probability. *Chelsea, New York* (original German edition, published in 1933).

Mises, R. von (1941). On the foundation of probability and statistics. AMS 12 : 191.

Ore, Oyestein (1960). Pascal and the invention of probability theory. *American Mathematical Monthly*, **67**, 409 – 419.

Savage, L. J. (1972). Foundation of Statistics, Second Revised Edition. *Dover, New York*.

Todhunter, I. (1931). A History of Mathematical Theory of Probability. *G. E. Stechert*, *New York*.

CHAPTER TWO

Random Variables and Probability Distributions

2.1 The concept of a random variable

Results of random experiments are often summarized in terms of numerical values. Consider, for example, the experiment of testing two electronic components. When an electronic component is tested, it is either defective or non-defective. The sample space of the experiment may therefore be written as $S = \{NN, DN, ND, DD\}$, where N denotes non-defective and D denotes defective.

Let X denote the number of electronic components which are defective. One is naturally interested in the possible values of X. Can X take the value 3? What about the value 1.5? The values X can take are 0, 1 and 2. Notice that X takes the value 0 at the sample point NN and the value 1 at the sample points DN and ND. What value does X take at the sample point DD?

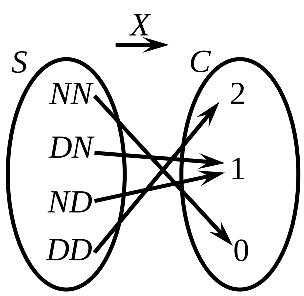

Fig. 2.1

It can be seen that X assigns a unique real number $X(s)$ to each sample point s of S (see Fig. 2.1). X is therefore a function with domain S and co-domain $C = \{0, 1, 2\}$. Such a function is called a **random variable**.

Definition 2.1:

A random variable is a function that assigns a real number to each element in the sample space of a random experiment.

A random variable is denoted by an uppercase letter, such as X, and a corresponding lowercase letter, such as x, is used to denote a possible value of X. We refer to the set of possible values of a random variable X as ***the range of X***.

Example 2.1

Three balls are drawn in succession without replacement from a box containing 5 white and 4 green balls. Let Y denote the number of white balls selected. The possible outcomes and the values y of Y are:

outcome	GGG	GGW	GWG	WGG	GWW	WGW	WWG	WWW
y	0	1	1	1	2	2	2	3

where G denotes "green" and W denotes "white" and the i^{th} letter in a triple, denotes the colour of the i^{th} ball drawn (i = 1, 2, 3). For example, GWG means the first ball drawn is green, the second ball drawn is white and the third ball drawn is green.

Example 2.2

Let W denote the number of times a die is thrown until a 3 occurs. The possible outcomes and the values w, of the random variable W are:

outcome	F,	NF,	NNF,	$NNNF$, …
w	1,	2,	3,	4,…

where F and N represent, respectively, the occurrence, and non-occurrence of a 3. The random variable W takes the values 1, 2, 3, 4, … . The range of W is said to be **countably infinite**.

Example 2.3

Let H denote the height, in metres, of a patient selected from a hospital. Values of H depend on the outcomes of the experiment. H is therefore a random variable.

In Example 2.1, the range of Y is finite while in Example 2.2, the range of W is countably infinite. The random variables Y and W are examples of discrete random variables.

It is easy to distinguish a discrete random variable from one that is not discrete. Just ask the question: "What are the possible values for the variable?" If the answer is a finite set or a countably infinite set, then the random variable is discrete; otherwise, it is not. This idea leads to the following definition.

Definition 2.2 (Discrete random variable)

A random variable is discrete if it can assume a finite or a countably infinite set of values.

The random variable H in Example 2.3 is different from the random variables Y and W in Examples 2.1 and 2.2. The range of H is neither finite nor countably infinite. H can assume any value in some interval of real numbers. H is an example of a continuous random variable. We therefore have the following definition.

Definition 2.3 (Continuous random variable)

If the range of a random variable X contains an interval (either finite or infinite) of real numbers, then X is a ***continuous random variable***.

In most practical problems, continuous random variables represent measured data, such as heights, weights, temperatures, distances, or life periods, whereas discrete random variables represent count data, such as the number of defectives in a sample of n items or the number of road traffic accidents in Accra in a week.

2.2 Discrete probability distributions

When dealing with a random variable, it is not enough just to determine what values are possible. We also need to determine what is probable. Consider the following example.

Example 2.4

Fifty taxi drivers were asked of the number of road traffic accidents they have had in a year. The results are given in Table 2.1.

***Table 2.1*: Number of accidents per year of 50 taxi drivers**

Number of accidents	0	1	2	3	4	5
Frequency	15	12	9	7	5	2

Suppose we select a taxi driver from this group and X is the number of road traffic accidents the person selected had in a year. What values can X take? X can take the values 0, 1, 2, 4 and 5. What is the probability that $X = 0$? Out of the 50 equally likely ways of selecting a taxi driver, there are 15 ways in which the taxi driver selected had no accident in a year. Hence the probability that $X = 0$ is $\frac{15}{50}$ or 0.30. This is written as $P(X = 0) = 0.30.$ Similarly, $P(X = 1) = \frac{12}{50} = 0.24.$ Table 2.2 gives the possible values x, of X and their probabilities. Table 2.2 is called the **probability distribution** of X. Note that the values of X exhaust all possible cases and hence the probabilities add up to 1.

Table 2.2:* The probability distribution of *X

x	0	1	2	3	4	5
$P(X = x)$	0.30	0.24	0.18	0.14	0.10	0.04

Example 2.5

A shipment of 8 similar microcomputers 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 of the number of defectives.

Solution

Let X be the number of defective computers purchased by the school. X can take the values 0, 1, and 2. Now,

$$P(X=0) = \frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} = \frac{10}{28}, \qquad P(X=1) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} = \frac{15}{28}, \qquad P(X=2) = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} = \frac{3}{28}.$$

The probability distribution of X is given in the following table.

x	0	1	2
$P(X = x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Representation of the probability distribution of a discrete random variable

The probability distribution of a discrete random variable X can be represented by a table, a formula or a graph.

Tabular form

For a random variable that can assume a small number of values, it is simplest to present its probability distribution in the form of a table having two rows: the upper row contains the possible values the random variable assumes and the lower row contains the corresponding probabilities of the values (see Table 2.3).

Table 2.3: **The probability distribution of *X***

x	x_1	x_2	$\cdots$	x_n
$P(X=x)$	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$

Formula

Frequently, it is convenient to represent the probability distribution of a discrete random variable by a formula. For example,

$$f(x)=\tfrac{1}{7}, \qquad x = 1, 2, \ldots, 7$$

defines a probability distribution of a discrete random variable.

The values of a discrete random variable are often called **mass points**; and $f(x_j)$ denotes the **mass** associated with the mass point x_j. The function $f(x)=P(X=x)$ is therefore called the **probability mass function** of the random variable X. Other terms used are **frequency function** and **probability function**. Also, the notation $p(x)$ is sometimes used instead of $f(x)$ for probability mass functions.

Definition 2.4 Probability mass function

A function $f(x)$ is the probability mass function of a discrete random variable X if it has the following two properties:

(1) $f(x)\geq 0$ for all x, (2) $\sum_{\text{all } x} f(x)=1.$

Graphical form

The probability distribution of a discrete random variable can also be represented graphically, as shown in Fig. 2.2. Such a graph is called a **probability graph** (see Fig. 2.2).

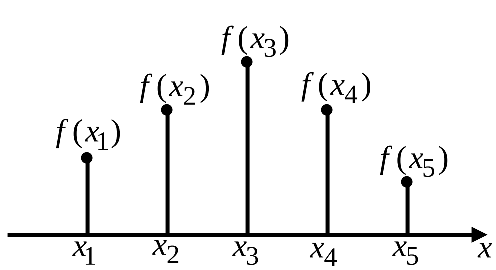

Fig. 2.2: *Probability graph*

To obtain the probability graph of a discrete random variable X, vertical lines are drawn above the possible values x_i of X, on the horizontal axis. The height of each line is equal to the probability of the corresponding value of X.

The probability distribution of a discrete random variable can also be represented by a **probability histogram** (see Fig. 2.3). Similar to the probability graph, the height of each rectangle of a probability histogram is equal to the probability that the random variable takes on the value which corresponds to the mid-point of the base.

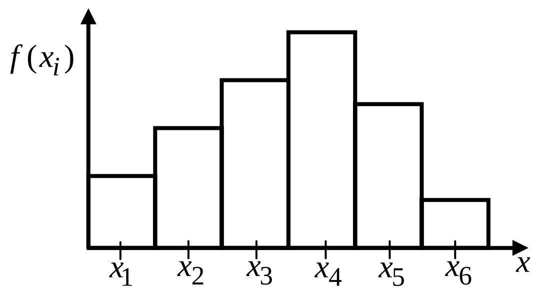

Fig. 2.3: *Probability histogram*

Random variables are so important in random experiments that sometimes we essentially ignore the original sample space of the experiment and focus on the probability distribution of the random variable. In this manner, a random variable can simplify the description and analysis of a random experiment.

Example 2.6

Determine whether each of the following can serve as a probability mass function of a discrete random variable:

(a) $f(x) = \frac{1}{2}(x-2)$, $x = 1, 2, 3, 4$. (b) $g(x) = \frac{1}{10}(x+1)$, $x = 0, 1, 2, 3$.

(c) $h(x) = \frac{1}{20}x^2$, $x = 0, 1, 2, 3, 4$.

Solution

(a) $f(1) = \frac{1}{2}(1-2) = -\frac{1}{2}$. $f(1)$ is negative and so $f(x)$ cannot serve as a probability mass function of a discrete random variable.

(b) $g(x) \geq 0$ for all values of x, and $\sum_{x=0}^{3} g(x) = \frac{1}{10}(1+2+3+4) = 1$. $g(x)$ is therefore the probability mass function of a discrete random variable.

(c) $h(x) \geq 0$ for all values of x and $\sum_{x=0}^{4} h(x) = \frac{1}{20}(0+1+4+9+16) = 1.5$. $\sum_{x=0}^{4} h(x) \neq 1$ and so $h(x)$ cannot serve as a probability mass function of a discrete random variable.

Example 2.7

The following table gives the probability distribution of a random variable X.

x	1	2	3	4
$p(x)$	$\frac{1}{4}$	c	$\frac{1}{2}$	$\frac{1}{8}$

(a) Find the value of the constant c.

(b) Represent the probability distribution by

(i) a probability graph, (ii) a probability histogram.

(c) Find (i) $P(X>1)$, (ii) $P(0<X<2)$, (iii) $P(X \geq 2)$.

Solution

(a) $\sum_{x=0}^{3} p(x) = 1 \Rightarrow \frac{1}{4} + c + \frac{1}{2} + \frac{1}{8} = 1 \Rightarrow c = \frac{1}{8}$.

(b) (i) Fig. 2.4 is the required probability graph.

(ii) Fig. 2.5 is the required probability histogram.

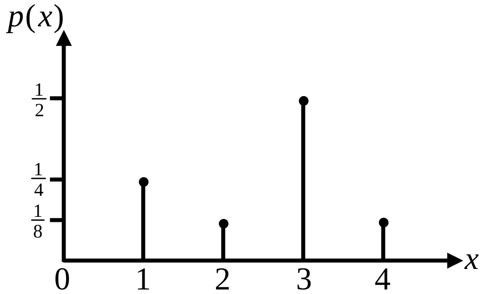

***Fig. 2.4:** Probability graph for Example 2.7*

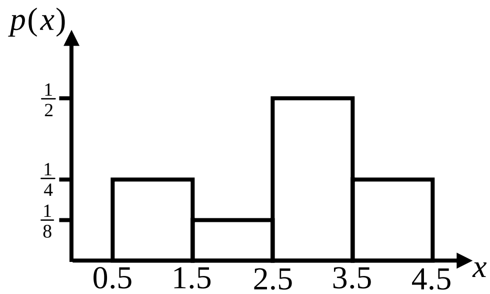

***Fig. 2.5:** Probability histogram for Example 2.7*

(c) (i) $P(X>1) = 1 - P(X \leq 1) = 1 - P(X=1) = 1 - \frac{1}{4} = \frac{3}{4}$.

Alternatively,

$P(X>1) = P(X=2) + P(X=3) + P(X=4) = \frac{1}{8} + \frac{1}{2} + \frac{1}{8} = \frac{3}{4}$.

(ii) $P(0<X<2) = P(X=1) = \frac{1}{4}$.

(iii) $P(X \geq 2) = P(X=2) + P(X=3) + P(X=4) = \frac{1}{8} + \frac{1}{2} + \frac{1}{8} = \frac{3}{4}$.

Alternatively,

$P(X \geq 2) = 1 - P(X<2) = 1 - P(X=1) = 1 - \frac{1}{4} = \frac{3}{4}$.

Exercise 2(a)

1. Classify the following random variables as discrete or continuous.
   (a) the weight of grain produced per acre,
   (b) the number of road traffic accidents per year in Accra,
   (c) the distance between two towns,
   (d) the length of time to write an examination.

2. A discrete random variable X has a probability mass function given by
   $f(x) = c(x+1), \quad x = 0, 1, 2, 3.$
   (a) Find the value of the constant c.
   (b) Draw (i) a probability graph,
   (ii) a probability histogram, to represent $f(x)$.
   (c) Find: (i) $P(0 \le X < 2)$, (ii) $P(X > 1)$.

3. Determine whether each of the following functions can serve as a probability mass function of a discrete random variable:
   (a) $f(x) = \frac{1}{2}(x-1), \quad x = 0, 1, 2, 3.$
   (b) $g(x) = \frac{1}{10}x, \quad x = 1, 2, 3, 4.$
   (c) $h(x) = \frac{1}{6}x^2, \quad x = -1, 0, 1, 2.$

4. Let X be a random variable whose probability mass function is defined by the values
   $f(-2) = \frac{1}{10}, \ f(0) = \frac{2}{10}, \ f(4) = \frac{4}{10}, \ f(11) = \frac{3}{10}.$
   Find: (a) $P(-2 \le X < 4)$, (b) $P(X > 0)$, (c) $P(X \le 4)$.

5. Check whether the following functions satisfy the conditions of a probability mass function.
   (a) $f(x) = \frac{1}{4}, \quad x = -3, 0, 1, 4.$
   (b) $f(x) = \frac{1}{x}, \quad x = 1, 2, 3, 4.$
   (c) $f(x) = 1 - x, \quad x = 0, \frac{1}{2}, \frac{3}{2}.$
   (d) $f(x) = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, 4, \ldots .$

6. Consider a throw of two fair dice. Let X denote the sum of the numbers on the two dice.
   (a) Find the probability mass function of X.
   (b) Find: (i) $P(X = 7)$, (ii) $P(X > 8)$, (iii) $P(3 < X < 11)$.

7. The sample space of an experiment is $\{a, b, c, d, e, f\}$, and each outcome is equally likely. A random variable X, is defined as follows:

outcome	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

   (a) Determine the probability distribution of X.
   (b) Find: (i) $P(X = 1.5)$, (ii) $P(0.5 < X < 2.7)$, (iii) $P(X > 3)$, (iv) $P(0 \le X < 2)$.

8. Determine the value of c so that each of the following can serve as a probability mass function of a discrete random variable.

   (a) $f(x) = c(x^2 + 4)$, $\quad x = 0, 1, 2, 3.$

   (b) $f(x) = c\binom{2}{x}\binom{3}{3-x}$, $\quad x = 0, 1, 2.$

9. A discrete random variable X has the probability mass function given by

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

   (a) Find the value of a. (b) Find $P(X = 3)$.

2.3 Continuous probability distributions

In Section 2.2, we learnt that for a complete characterization of a discrete random variable, it is necessary and sufficient to know the probability mass function of the random variable. Corresponding to every continuous random variable X, there is a function f, called the **probability density function** (p.d.f.) of X such that

(a) $f(x) \geq 0$, (b) $\int_{-\infty}^{\infty} f(x)dx = 1$, (c) $P(a \leq X \leq b) = \int_a^b f(x)dx.$

For a complete characterization of a continuous random variable, it is necessary and sufficient to know the p.d.f. of the random variable.

Geometrically, relation (c) means the following: the probability that a continuous random variable X takes values in the interval (a, b) is equal to the area of the region defined by the p.d.f. of X, the straight lines $x = a$ and $x = b$, and the x-axis (see Fig. 2.6).

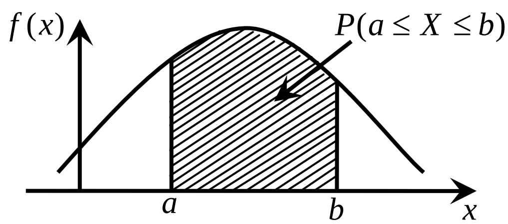

Fig. 2.6: *Probability as an area under a p.d.f.*

A consequence of X being a continuous random variable is that for any value in the range of X, say x,

$$P(X = x) = \int_x^x f(t)\,dt = 0 \quad \text{.....(2.1)}$$

Thus a continuous random variable has a probability of zero of assuming **exactly** any of its values. At first this may seem startling, but it becomes more plausible when we consider an example. Consider a random variable whose values are the heights of all people over twenty years of age. Between any two values, say 162.99 and 163.01 centimetres, there are infinite number of heights, one of which is 163 centimetres. The probability of selecting a person at random who is **exactly** 163 centimetres tall and not one of the infinitely large set of heights so close to 163 centimetres that you cannot humanly measure the difference, is remote, and thus we assign a probability of zero to the event.

As an immediate consequence of Equation (2.1), if X is a continuous random variable, then for any numbers a and b, with $a \leq b$,

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b). \quad \text{......(2.2)}$$

That is, it does not matter whether we include an endpoint of the interval or not. This is not true, though, when X is discrete.

Example 2.8

(a) Show that

$$f(x) = \begin{cases} \frac{1}{9}x^2, & 0 < x < 3, \\ 0, & \text{elsewhere.} \end{cases}$$

is the p.d.f. of a continuous random variable X.

(b) Sketch the graph of $f(x)$.

Solution

(a) We have to show that $f(x) \geq 0$ for all x and $\int_{-\infty}^{\infty} f(x)dx = 1$. It is clear that $f(x) \geq 0$ for all x.

Moreover,
$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{3} f(x)dx + \int_{3}^{\infty} f(x)dx.$$
$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{3} \frac{1}{9}x^2 dx + \int_{3}^{\infty} 0\,dx$$
$$= 0 + \left[\frac{1}{27}x^3\right]_0^3 + 0 = 1.$$

Hence, $f(x)$ is the p.d.f. of a continuous random variable.

(b) Fig. 2.7 shows a sketch of the graph of $f(x)$.

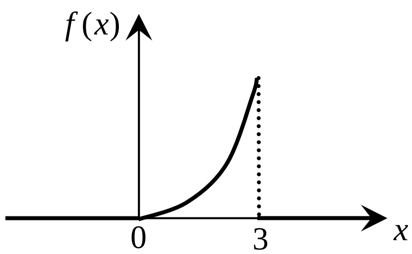

Fig. 2.7: *The p.d.f. of X*

Example 2.9

Refer to Example 2.8. Find:

(a) $P(X \leq 2)$, (b) $P(X > 1)$, (c) $P(2 \leq X \leq 3)$.

Solution

(a)
$$P(X \leq 2) = \int_{-\infty}^{2} f(x)\,dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{2} f(x)dx$$
$$= \int_{-\infty}^{0} 0dx + \int_{0}^{2} \frac{1}{9}x^2 dx$$
$$= 0 + \left[\frac{1}{27}x^3\right]_0^2 = \frac{8}{27}.$$

(b) $P(X > 1) \;=\; \int_1^{\infty} f(x)\,dx \;=\; \int_1^3 f(x)dx \;+\; \int_3^{\infty} f(x)dx$

$$= \int_1^3 \tfrac{1}{9}x^2 dx \;+\; \int_3^{\infty} 0\,dx$$

$$= \left[\tfrac{1}{27}x^3\right]_1^3 \;+\; 0 \;=\; \tfrac{1}{27}(27-1) \;=\; \tfrac{26}{27}.$$

(c) $P(2 \le X \le 3) \;=\; \int_2^3 \tfrac{1}{9}x^2 dx \;=\; \left[\tfrac{1}{27}x^3\right]_2^3 \;=\; \tfrac{1}{27}(27-8) \;=\; \tfrac{19}{27}.$

Example 2.10

A random variable X has the p.d.f. given by

$$f(x) = \begin{cases} c\sqrt{x}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the value of the constant c.

(b) Calculate $P\left(X < \tfrac{1}{4}\right)$.

Solution

(a) The value of c is given by the equation

$$1 \;=\; \int_{-\infty}^{\infty} f(x)\,dx \;=\; \int_{-\infty}^{0} f(x)dx \;+\; \int_0^1 f(x)dx \;+\; \int_1^{\infty} f(x)dx$$

$$= \int_{-\infty}^{0} 0\,dx \;+\; \int_0^1 c\sqrt{x}\,dx \;+\; \int_1^{\infty} 0\,dx \;=\; 0 \;+\; \left[\frac{cx^{\frac{3}{2}}}{3/2}\right]_0^1 \;+\; 0 \;=\; \tfrac{2}{3}c$$

$$\Rightarrow \quad c \;=\; 1\tfrac{1}{2}.$$

Thus, $f(x) = \begin{cases} \tfrac{3}{2}\sqrt{x}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$

(b) $P\left(X < \tfrac{1}{4}\right) \;=\; \int_{-\infty}^{\frac{1}{4}} f(x)\,dx \;=\; \int_{-\infty}^{0} f(x)\,dx \;+\; \int_0^{\frac{1}{4}} f(x)\,dx$

$$= \int_{-\infty}^{0} 0\,dx \;+\; \int_0^{\frac{1}{4}} \tfrac{3}{2}\sqrt{x}\,dx \;=\; 0 \;+\; \left[x^{\frac{3}{2}}\right]_0^{\frac{1}{4}} \;=\; \tfrac{1}{8}.$$

The following example illustrates how conditional probability can be applied to random variables.

Example 2.11

A random variable X has p.d.f. given by

$$f(x) = \begin{cases} \frac{1}{10}, & 0 < x < 10, \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X > 8 | X > 5)$, (b) $P(X > 7 | X < 9)$.

Solution

(a) $$P(X > 8 | X > 5) = \frac{P(X > 8,\ X > 5)}{P(X > 5)} = \frac{P(X > 8)}{P(X > 5)}$$
$$= \left(\int_8^{10} \tfrac{1}{10}\,dx\right) \Big/ \left(\int_5^{10} \tfrac{1}{10}\,dx\right)$$
$$= \left[\tfrac{1}{10}x\right]_8^{10} \Big/ \left[\tfrac{1}{10}x\right]_5^{10} = (10-8)/(10-5) = \tfrac{2}{5}.$$

(b) $$P(X > 7 | X < 9) = \frac{P(X > 7,\ X < 9)}{P(X < 9)} = \frac{P(7 < X < 9)}{P(X < 9)}$$
$$= \left(\int_7^{9} \tfrac{1}{10}\,dx\right) \Big/ \left(\int_0^{9} \tfrac{1}{10}\,dx\right)$$
$$= \left[\tfrac{1}{10}x\right]_7^{9} \Big/ \left[\tfrac{1}{10}x\right]_0^{9} = (9-7)/(9-0) = \tfrac{2}{9}.$$

Exercise 2(b)

1. Show that the following functions are probability density functions for some value of c and determine c.

   (a) $f(x) = \begin{cases} ce^{-4x}, & x \geq 0, \\ 0, & \text{elsewhere.} \end{cases}$ (b) $f(x) = \begin{cases} cx^2, & -1 < x < 10, \\ 0, & \text{elsewhere.} \end{cases}$

   (c) $f(x) = \begin{cases} c(1+2x), & 0 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$ (d) $f(x) = \begin{cases} \frac{1}{2}e^{-cx}, & x \geq 0, \\ 0, & \text{elsewhere.} \end{cases}$

2. Suppose that in a certain region, the daily rainfall (in inches) is a continuous random variable X with p.d.f. $f(x)$ given by
$$f(x) = \begin{cases} \frac{3}{4}(2x - x^2), & 0 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$
   Find the probability that on a given day in this region the rainfall is
   (a) not more than 1 inch, (b) greater than 1.5 inches,
   (c) equal to 1 inch, (d) less than 1 inch.

3. Let X be a continuous random variable with p.d.f.
$$f(x) = \begin{cases} -4x, & -0.5 < x < 0, \\ 4x, & 0 < x < 0.5, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Sketch the graph of $f(x)$.

(b) Find: (i) $P(X \le -0.3)$, (ii) $P(X \le 0.3)$, (iii) $P(-0.2 \le X \le 0.2)$.

4. The pressure (measured in kg/cm^2) at a certain valve is a random variable X whose p.d.f. is

$$f(x) = \begin{cases} \frac{6}{27}(3x - x^2), & 0 < x < 3, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that the pressure at this valve is

(a) less than 2 kg/cm^2, (b) greater than 2 kg/cm^2, (c) between 1.5 and 2.5 kg/cm^2.

5. Let X denote the length in minutes of a long-distance telephone conversation. Assume that the p.d.f. of X is given by

$$f(x) = \frac{1}{10} e^{-x/10}, \qquad x > 0.$$

(a) Verify that f is a p.d.f. of a continuous random variable.

(b) Find the probability that a randomly selected call will last:

(i) at most 7 minutes, (ii) at least 7 minutes, (iii) exactly 7 minutes.

6. A continuous random variable X has the p.d.f.

$$f(x) = \begin{cases} \frac{2}{27}(1 + x), & 2 < x < 5, \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X < 4)$, (b) $P(3 < X < 4)$.

7. The proportion of people who respond to a certain mail-order solicitation is a continuous random variable X with p.d.f.

$$f(x) = \begin{cases} \frac{2}{5}(x + 2), & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find $P(0 < X < 0.8)$.

(b) Find the probability that more than $\frac{1}{4}$ but fewer than $\frac{1}{2}$ of the people contacted will respond to this type of solicitation.

8. A continuous random variable X that can assume values between $x = 1$ and $x = 3$ has a p.d.f. given by $f(x) = \frac{1}{2}$.

(a) Show that the area under the curve is equal to 1.

(b) Find: (i) $P(2 < X < 2.5)$, (ii) $P(X \le 1.6)$.

9. The p.d.f. of the length X millimetres of a hinge for fastening a door is

$$f(x) = 1.25, \qquad 74.6 < x < 75.4.$$

(a) Find $P(X < 74.8)$.

(b) If the specifications for this process are from 74.7 to 75.3 millimetres, what proportion of hinges meet specifications?

10. The p.d.f. of the length Y metres, of a metal rod is

$$f(y) = 2, \qquad 2.3 < y < 2.8.$$

If the specifications for this process are from 2.25 to 2.75 metres, what proportion of the bars fail to meet the specifications?

11. The p.d.f. of the time X seconds, required to complete an assembly operation is

$$f(x) = 0.1, \quad 30 < x < 40.$$

(a) Determine the proportion of assemblies that require more than 35 seconds to complete.
(b) What time is exceeded by 90% of the assemblies?

12. Which of the following functions are probability density functions?

(a) $f(x) = \begin{cases} x, & -0.5 < x < 0.5, \\ 0, & \text{elsewhere.} \end{cases}$ (b) $g(x) = \begin{cases} \frac{1}{2}, & -1 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$

(c) $f(x) = \begin{cases} 1/x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$ (d) $h(x) = \begin{cases} \frac{1}{3}, & 0 < x < 1, \\ \frac{2}{3}, & 2 < x < 3. \end{cases}$

13. The random variable X has the p.d.f.

$$f(x) = \begin{cases} \frac{1}{30}, & 0 < x < 30, \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X > 25 | X > 15)$, (b) $P(X < 20 | X > 15)$,
(c) $P(X > 15 | X < 22)$, (d) $P(X < 13 | X < 18)$.

2.4 The cumulative distribution function

There are many problems where we may wish to compute the probability that the observed value of a random variable X will be less than or equal to some real number x. For example, what are the chances that a certain candidate will get no more than 30% of the votes? What are the chances that the prices of gold will remain at or below \$800 per ounce? Writing $F(x) = P(X \le x)$ for every real number x, we define $F(x)$ to be the cumulative distribution function of X, or more simply, the distribution function of the random variable X.

2.4.1 The cumulative distribution function of a discrete random variable

Definition 2.5 (Cumulative distribution function)

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

$$F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i).$$

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

$$F(x) = \begin{cases} 0, & -\infty < x < x_1, \\ f(x_1), & x_1 \le x < x_2, \\ f(x_1) + f(x_2), & x_2 \le x < x_3, \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ f(x_1) + f(x_2) + \ldots + f(x_n) = 1, & x_n \le x < \infty. \end{cases}$$

Fig. 2.8 depicts the graph of $F(x)$. It can be seen that $F(x)$ is discontinuous at the points $x_1, x_2, \ldots, x_n$. At these points, $F(x)$ is continuous from the right but discontinuous from the left. Because of the shape of its graph, the cumulative distribution function of a discrete random variable is called a **staircase function** or a **step function**. Notice that $F(x)$ has a jump of height $f(x_i)$ at the point x_i and is constant in the interval (x_i, x_{i+1}).

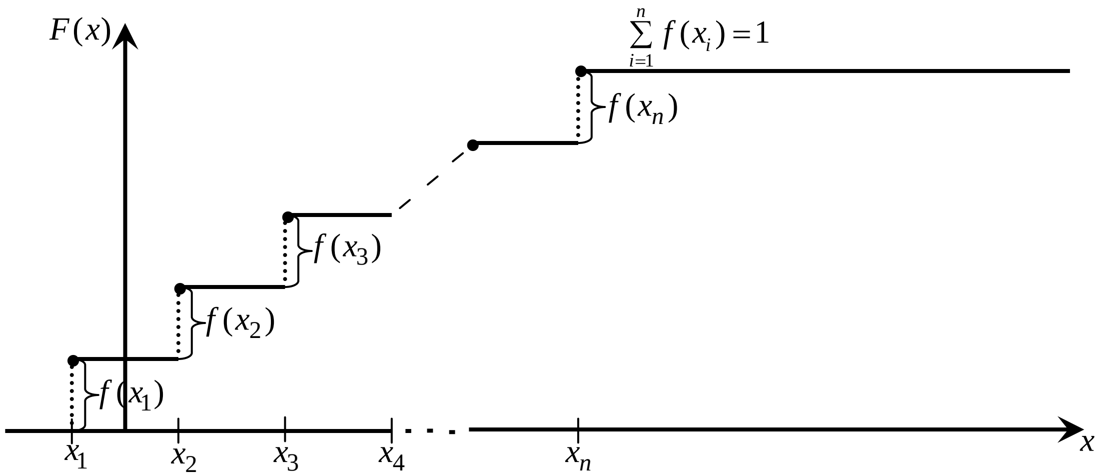

***Fig. 2.8**: A typical cumulative distribution function of a discrete random variable*

Properties of $F(x)$

The cumulative distribution function of a discrete random variable has the following properties:

(1) $0 \le F(x) \le 1$ for all x.

(2) If $x \le y$, then $F(x) \le F(y)$. This means that F is a non-decreasing function.

(3) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

(4) F is right-continuous. That is, $\lim_{x \to a+} F(x) = F(a)$ for all points a.

Notice that $F(x)$ is not continuous from the left, that is, $\lim_{x \to a-} F(x) \ne F(a)$ for all points a. This is because we have defined $F(x)$ by $F(x) = P(X \le x)$. If $F(x)$ is defined by $F(x) = P(X < x)$, it will be continuous from the left, but not from the right.

Example 2.12

The following table gives the probability mass function of X. Find the cumulative distribution function of X and sketch its graph.

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Solution

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

If $0 \le x < 1$, $F(x) = f(0) = \frac{1}{16}.$

If $1 \le x < 2$, $F(x) = f(0) + f(1) = \frac{1}{16} + \frac{1}{4} = \frac{5}{16}.$

If $2 \le x < 3$, $F(x) = f(0) + f(1) + f(2) = \frac{1}{16} + \frac{1}{4} + \frac{3}{8} = \frac{11}{16}.$

If $3 \le x < 4$, $F(x) = f(0) + f(1) + f(2) + f(3) = \frac{1}{16} + \frac{1}{4} + \frac{3}{8} + \frac{1}{4} = \frac{15}{16}.$

If $x \ge 4$, $F(x) = f(0)+f(1)+f(2)+f(3)+f(4) = \frac{1}{16} + \frac{1}{4} + \frac{3}{8} + \frac{1}{4} + \frac{1}{16} = 1.$

The cumulative distribution function of X is therefore given by

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

Fig. 2.9 shows the graph of $F(x)$.

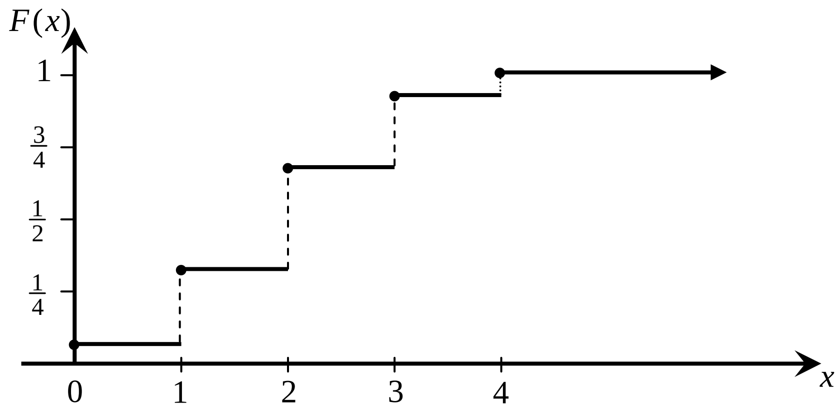

***Fig. 2.9**: The cumulative distribution function for Example 2.12*

Notice that even if the random variable X can assume only integers, the cumulative distribution function of X can be defined for non-integers. For example, in Example 2.12, $F(1.5) = \frac{5}{16}$, $F(2.5) = \frac{11}{16}$.

To find the probability mass function, $f(x)$, corresponding to a given cumulative distribution function, $F(x)$, we first find the points where $F(x)$ is discontinuous. We then find the magnitudes of the jumps at these points. If $F(x)$ has a jump at the point x_i, then $f(x_i)$ is equal to the magnitude of this jump. Example 2.13 illustrates the procedure.

Example 2.13

Suppose the cumulative distribution function of X is

$$F(x) = \begin{cases} 0, & x < -2. \\ 0.2, & -2 \le x < 0. \\ 0.7, & 0 \le x < 2. \\ 1, & 2 \le x. \end{cases}$$

Determine the probability mass function of X.

Solution

$F(x)$ is discontinuous at the points –2, 0 and 2. The jump at the point –2 is 0.2 – 0 = 0.2 and so $f(-2) = 0.2.$ The jump at the point 0 is 0.7 – 0.2 = 0.5, and so $f(0) = 0.5.$ Similarly $f(2) = 1 - 0.7 = 0.3.$ The following table gives the probability mass function of X.

x	–2	0	2
$f(x)$	0.2	0.5	0.3

2.4.2 The cumulative distribution function of a continuous random variable

As in the case of discrete random variables, we are often interested in the probability that the value of a particular continuous random variable will be less than or equal to a given number. Again, as in the case of discrete random variables, the mathematical function used to designate a probability of this type is called a **cumulative distribution function**. The formal definition follows.

Definition 2.6 (Cumulative distribution function)

Let X be a continuous random variable with probability density function f. The function $F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$ for $-\infty < x < \infty$, is called the cumulative distribution function or the distribution function of the random variable X.

As an immediate consequence of Definition 2.6, we can write the following two results

$$P(a < X < b) = \int_{-\infty}^{b} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt = F(b) - F(a) \quad \text{.......................(2.3)}$$

and $$f(x) = \frac{d}{dx}F(x) \quad \text{.......................(2.4)}$$

Example 2.14

A continuous random variable X has the p.d.f.

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

(a) Find the distribution function of X and sketch its graph.

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

Solution

(a) If $x \leq -1$, then

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0.$$

If $-1 \leq x \leq 2$, then

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{x} \tfrac{1}{3}t^2\,dt = \left[\tfrac{1}{9}t^3\right]_{-1}^{x} = \tfrac{1}{9}(x^3+1)$$

If $x \geq 2$, then

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-1} f(t)\,dt + \int_{-1}^{2} f(t)\,dt + \int_{2}^{x} f(t)\,dt \\ &= \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{2} \tfrac{1}{3}t^2\,dt + \int_{2}^{x} 0\,dt \\ &= 0 + \left[\tfrac{1}{9}t^3\right]_{-1}^{2} + 0 = \tfrac{1}{9}(8+1) = 1. \end{aligned}$$

Therefore,

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

Fig. 2.10 shows the graph of $F(x)$. Notice that $F(x)$ is a continuous, non-decreasing function.

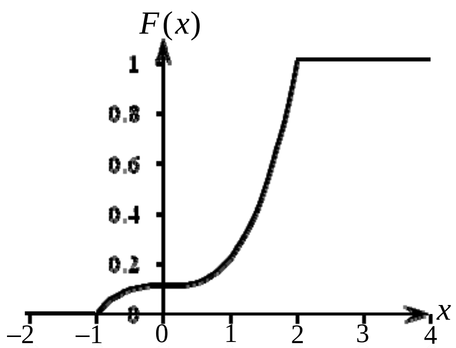

***Fig. 2.10**: The distribution function for Example 2.14*

(b) $P(0 < X \leq 1) = F(1) - F(0) = \frac{1}{9}(1^2 + 1) - \frac{1}{9}(0 + 1) = \frac{1}{9}.$

Example 2.15

Let X be a continuous random variable with cumulative distribution function given by

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{4}x^2, & 0 \leq x \leq 2, \\ 1, & x \geq 2. \end{cases}$$

Find the p.d.f. of X.

Solution

By applying Equation (2.4) on page 51, we find that the p.d.f. of X is given by

$$f(x) = \frac{d\,F(x)}{dx} = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Properties of $F(x)$

Let $F(x)$ be the distribution function of a continuous random variable. $F(x)$ has the following properties:

1. $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0, \qquad F(+\infty) = \lim\limits_{x \to \infty} F(x) = 1.$
2. The function $F(x)$ is the probability of an event, and so $0 \le F(x) \le 1, \quad -\infty < x < \infty.$
3. $F(x)$ is continuous everywhere.
4. $\dfrac{d}{dx}F(x)$ exists at all points x.
5. $F(x)$ is a non-decreasing function of x, that is, if $x_1 \le x_2$, then $F(x_1) \le F(x_2)$.

Exercise 2(c)

1. Let X be a discrete random variable whose only possible values are 1, 2 and 5. Find the cumulative distribution function of X if the probability mass function of X is defined by the following values: $f(1) = \frac{1}{4},\ f(2) = \frac{1}{2},\ f(5) = \frac{1}{4}$.

2. Let X be a discrete random variable whose cumulative distribution function is

$$F(x) = \begin{cases} 0, & x < -3, \\ \frac{1}{6}, & -3 \le x < 6, \\ \frac{1}{2}, & 6 \le x < 10, \\ 1, & x \ge 10. \end{cases}$$

   (a) Find (i) $P(X \le 4)$, (ii) $P(-5 < X \le 4)$, (iii) $P(X = 4)$.
   (b) Find the probability mass function of X.

3. Let X be a discrete random variable with cumulative distribution function

$$F(x) = \begin{cases} 0, & x < 1, \\ 0.1, & 1 \le x < 3, \\ 0.4, & 3 \le x < 5, \\ 0.9, & 5 \le x < 5.5, \\ 1.0, & x \ge 5.5. \end{cases}$$

   (a) Find (i) $P(X \le 3)$, (ii) $P(X \le 4)$, (iii) $P(1.5 < X \le 5.2)$.
   (b) Find the probability mass function of X.

4. Let X be a discrete random variable whose only possible values are -5, -1, 0, and 7. Find the cumulative distribution function of X if the probability mass function of X is defined by the values $f(-5) = 0.3$, $f(-1) = 0.1$, $f(0) = 0.2$, and $f(7) = 0.4$.

5. Let

$$F(x) = \begin{cases} 0, & x \le 0, \\ x^2, & 0 \le x \le 1, \\ 1, & x \ge 1. \end{cases}$$

be the cumulative distribution function of the random variable X. Find

(a) $P(X \le -1)$, (b) $P(X \le 0.5)$, (c) $P(X > 0.4)$,

(d) $P(0.2 < X \le 0.5)$, (e) $P(X > 0.4 \mid X > 0.2)$.

6. If a certain type of motor is in operating condition after 4 years of service, the total number of years it will be in operating condition is a random variable X with cumulative distribution function

$$F(x) = \begin{cases} 0, & x \le 4, \\ 1 - 16/x^2, & x \ge 4. \end{cases}$$

(a) Find the p.d.f. of X.
(b) Find the probability that a 4-year-old motor of this type will be in operating condition for:
(i) at least 5 years, (ii) less than 10 years, (iii) between 5 and 7 years.

7. The weekly profit (in thousands of Ghana cedis) from a certain concession is a random variable X whose distribution function is given by

$$F(x) = \begin{cases} 0, & x \le 0, \\ 3x - 3x^2 + x^3, & 0 \le x \le 1, \\ 1, & x \ge 1. \end{cases}$$

(a) Find the probability of a weekly profit of less than GH¢2 000.00.
(b) Find the probability of a weekly profit of at least GH¢500.00.

8. The p.d.f. of a continuous random variable is

$$f(x) = \begin{cases} kx, & 1 \le x \le 5, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the value of the constant k.
(b) Find (i) $P(X < 4)$, (ii) $P(2 \le X \le 3)$, (iii) $P(X > 7)$, (iv) $P(X > 4 \mid X > 2)$.
(c) Find the cumulative distribution function of X.
(d) Use the cumulative distribution function to find the probabilities in part (b).
(e) Sketch the graph of the cumulative distribution function.
(f) Is $F(x)$ continuous everywhere?

9. The distribution function of the continuous random variable X is

$$F(x)=\begin{cases}0, & x\le 3,\\ 1-k/x^2, & x\ge 3.\end{cases}$$

(a) Find the value of k.

(b) Sketch the graph of $F(x)$.

(c) Find (i) $P(X<4)$, (ii) $P(4<X<5)$, (iii) $P(X>6)$.

(d) Find the p.d.f. of X.

10. Which of the following are cumulative distribution functions of continuous random variables?

(a) $F(x)=\begin{cases}0, & x\le -1,\\ x^2, & -1\le x\le 1,\\ 1, & x\ge 1.\end{cases}$ (b) $F(x)=\begin{cases}0, & x\le 0,\\ x^2, & 0\le x\le 1,\\ 1, & x\ge 1.\end{cases}$

(c) $F(x)=\begin{cases}0, & x<0,\\ \frac{1}{2}, & 0\le x<1,\\ 1, & x\ge 1.\end{cases}$ (d) $F(x)=\begin{cases}-1, & x\le -1,\\ x, & -1\le x\le 1,\\ 1, & x\ge 1.\end{cases}$

11. The time until a chemical reaction is complete (in milliseconds) is a random variable X with the following cumulative distribution function

$$F(x)=\begin{cases}0, & x<0,\\ 1-e^{-0.01x}, & x\ge 0.\end{cases}$$

Determine the p.d.f. of X. What proportion of reactions is complete within 200 milliseconds?

12. Suppose the cumulative distribution function of the random variable X is

$$F(x)=\begin{cases}0, & x\le -2,\\ 0.25x+0.5, & -2\le x\le 2,\\ 1, & 2\le x.\end{cases}$$

(a) Find the p.d.f. of X.

(b) Calculate:

(i) $P(X<1.8)$, (ii) $P(X>-1.5)$, (iii) $P(X<-2)$, (iv) $P(X>1 \mid X>0.5)$.

13. A continuous random variable X has the p.d.f.

$$f(x)=\begin{cases}e^{-cx}, & x\ge 0,\\ 0, & x<0,\end{cases}$$

where c is a constant.

(a) Find the value of c.

(b) Find the cumulative distribution function of X and sketch its graph.

(c) Find (i) $P(X>8 \mid X>3)$, (ii) $P(X>1 \mid X<4)$.

14. The distribution function of X is given by

$$F(x)=\begin{cases} 0, & x\le 0,\\ x/8, & 0\le x\le 2,\\ x^2/16, & 2\le x\le 4,\\ 1, & x\ge 4.\end{cases}$$

(a) Find the p.d.f. of X.

(b) Find (i) $P(1\le X\le 3)$, (ii) $P(X<3)$, (iii) $P(X>1 \mid X>2)$.

15. The shelf life, in days, for bottles of a certain prescribed medicine is a random variable with p.d.f.

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

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

(a) at least 200 days, (b) between 80 and 120 days.

16. A random variable X has distribution function given by

$$F(x) = \begin{cases} 0, & x\le -1,\\ \frac{1}{2}(x+1)^2, & -1\le x\le 0,\\ 1-\frac{1}{2}(1-x)^2, & 0\le x\le 1,\\ 1, & x\ge 1.\end{cases}$$

Find the p.d.f. of X.

17. A random variable X has the p.d.f. given by

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

(a) Find the cumulative distribution function of X and sketch its graph.

(b) Compute (i) $P(X\le 2)$, (ii) $P(X\le 2 \mid 1\le X<3)$.