UNIT 2 PROBABILITY CONCEPTS

Structure **Page No.**

2.1 INTRODUCTION

Sunil is an enterprising class VIII student who wants to use his free hours more fruitfully. A news paper agent agrees to employ him for one hour in the morning between 5.30 AM to 6.30 AM for distributing news papers in a residential colony where there are 85 regular subscribers. In addition, Sunil finds that there are about 10 irregular customers who may buy the paper from him on a day to day basis. On every additional news paper Sunil sells, he makes an extra income of 30 paise. But on every unsold news paper that he takes back to the agent he looses 10 paise. Sunil has to decide how many newspapers he should collect from the agent each morning so that he makes the maximum possible gain. His dilemma is about the 10 irregular customers as he has to decide how many of these 10 will actually buy from him on any given day. If you ask him whether he knows probability theory, he may be surprised and say he knows nothing. Actually he may already be using some of the ideas of probability theory while not being aware of or articulate about them. This is a very commonly occurring situation.

When asked, "Do you know anything about probability?" most people are quick to answer, "No!" Usually that is not the case at all. The words probable and probably are used commonly in everyday language. We say, "It will probably rain tomorrow" or "there is 0% chance of rain today." Such statements often have a vague, subjective quality and based sometimes on certain information and at other times on intuition only.

We live in an uncertain world. When we get up in the morning we cannot say exactly who we are going to meet, what the weather will be like or what event will be on the television news during the day.

In our everyday lives, we cope with this uncertainty by making hundreds of guesses, calculated risks and some gambles. We don't take a coat with us for the weekend because it is unlikely to be cold. We allow a particular length of time to travel to an important interview because it will probably be enough.

All these decisions are made by assessing the relative probability (chance) of all the possible outcomes -- even if we do this unconsciously and intuitively. Business decisions are made in a similar climate of uncertainty. A publisher must decide how large the print run of a new book should be to avoid unusual storage of unsold copies and yet ensure availability. A stock market dealer decides to sell a particular share because a financial model tells her that the price is likely to fall.

The penalties of estimating chances inaccurately and hence making a wrong decision vary from minor inconvenience, to loss of income to bankruptcy. So, in business (and other fields) we endeavour to measure uncertainty using some scientific method. Rather than make vague statements containing 'likely', 'may be' or 'probably'; we need to be more precise.

Historically, the oldest way of measuring uncertainties is the probability concept. Probability theory had its beginnings over 300 years ago, when gamblers of that period asked mathematicians to develop a system for predicting outcomes of a turn of the roulette wheel or a roll of a pair of dice. The word probability is associated with a quantitative approach to predicting the outcome of an event (the outcome of a presidential election, the side effects of a new medication, etc.).

In this unit we shall see how uncertainties can actually be measured, how they can be assigned numbers (called probabilities) and how these numbers are to be interpreted. After starting with some preliminaries of the probabilities we shall concentrate on the rules which probabilities must obey. This includes the basic postulates, the relationship between probabilities and odds, the addition rules, the definition of conditional probability, the multiplication rules etc.

Objectives

After reading this unit, you should be able to

- describe trials, events, sample spaces associated with an experiment;
- express the union, intersection, complement of two or more events in terms of a new event;
-] define the probability of the occurrence of an event and obtain it;
- obtain the conditional probability of an event.

2.2 PRELIMINARIES

What is an experiment? Many of you will relate experiments with all that you were expected to do in physics, chemistry or biology laboratories in your schools and colleges. For example, you may perhaps recollect that in the chemistry laboratory, one of the experiments you performed was to explore as to what would happen if sulfuric acid is poured in a jar containing zinc. Yet in another experiment in the physics laboratory, you might have performed the act of inserting a battery of a certain specification in a given circuit with a view to finding out the quantum of the flow of electricity through the said circuit. Thus, generally speaking, an **experiment** is merely the performance of an act for generating an observation on the phenomenon under study. In fact, when you pick up an item from a lot consisting of a number of items coming out of a manufacturing process, say, to decide whether the picked up item is defective or not, then also you are performing an experiment. Whenever you are investing a certain sum of money in the share market to see to what extent your money grows over a certain specified time interval, you are performing an experiment too. When you are stocking a number of units of a particular brand of a consumer good in your store in anticipation of sale, that is also an experiment. Tossing a coin, rolling a die, observing the number of road accidents on a given day in a city etc. are all experiments. In each such case, a certain act is performed and its outcome is observed. Is there then any difference between the former type of laboratory experiments and the latter type of experiments that we presumably perform in some form or the other in our daily lives? The answer to this question is 'Yes'. The difference is in terms of the outcomes that you associate with the experiments. **Notice** that in the former type of experiments performed under the controlled conditions of a laboratory, the outcomes are known a priori from the conditions under which these are carried out. It is known apriori that sulfuric acid and zinc together will yield hydrogen gas and zinc sulfate. The quantum of electricity flow through the circuit is known from the specifications of the inserted battery through the celebrated Ohm's Law. Such experiments are **deterministic** in the sense that the conditions under which these are carried out would inform us what the result is going to be even before the experiment is performed. However, when you are blindly picking up an item from a manufactured lot, you have no way of knowing a priori whether the item being picked up would be defective or not. When in your role as a shopkeeper, you stock certain units of a particular commodity, you will have no prior knowledge what your sales will be like over a specified period. When you toss a coin (it is equivalent to picking up an item blindly from a lot containing both defective and non-defective items - how?), you do not know the result of the experiment beforehand. Experiments whose outcomes cannot be precisely predicted primarily because the totality of the factors influencing the outcomes are either not precisely identifiable or not controllable at the time of experimentation, even if known, are called **random or stochastic** experiments. If you look at the dilemma of Sunil, the news paper boy, do you think what Sunil observes regarding the buying behaviour of the ten irregular customers on a given day can be modelled as the outcome of a random experiment? As Sunil does not know whether any of the specified irregular customer will actually buy from him, it is possible to think of this as a random experiment.

For the purposes of this block, by an experiment, we shall always mean a random experiment only. We now introduce some of the commonly occurring terms of the probability theory.

2.2.1 Trials, Sample Space, Events

You must have often observed that a random experiment may comprise of a series of smaller sub-experiments. These are called **trials**. Consider for instance the following situations.

Example 1: Suppose the experiment consists of observing the results of three successive tosses of a coin. Each toss is a trial and the experiment consists of three trials so that it is completed only after the third toss (trial) is over.

* * *

Example 2: Suppose from a lot of manufactured items, ten items are chosen successively following a certain mechanism for checking. The underlying experiment is completed only after the selection of the tenth item is completed; the experiment obviously comprises of 10 **trials**.

* * *

Example 3: If you consider Example 1 once again you would notice that each toss (trial) results into either a head (H) or a tail (T). In all there are 8 possible outcomes of the experiment viz., s_1 = (H,H,H), s_2 = (H,H,T), s_3 = (H,T,H), s_4 = (T,H,H), s_5 =(T,T,H), s_6 = (T,H,T), s_7 =(H,T,T) and s_8 = (T,T,T).

Each of the above outcome $s_1, s_2, \ldots, s_8$ is called a **sample point. Notice** that each sample point has three entries separated by commas. For example, the sample point s_2 indicates that the first toss results in a head, the second also produces head while the third one leads to a tail. Thus the sequence in which H's and T's appear is important. The set $\zeta = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\}$ of all possible outcomes is called the **sample space** of the experiment. Sample spaces are classified as **finite** or **infinite** according to the number of sample points (finite/infinite) they contain. For instance, an infinite sample space arises when we throw a dart at a target and there is a continuum of points we may hit. Sample space that arises in the case of Sunil, the news paper boy is an example of a finite sample space. What do you think is the sample space in this case? In this case the random experiment is to observe the 10 irregular customers on any day and note down whether each one "buys" or "does not buy" the newspaper. Therefore the sample space contains 2^{10} simple points which are sequences of the from (B, NB, NB, B, B, B, NB ....) etc. The sample space is, of course, finite. A specific collection or subset of sample points, say $E_1 = \{s_2, s_3, s_4\}$ is called an **event**. Event is a **simple event** if it consists of a single sample point. In this case $\{s_i\}$ is a simple event for $i = 1, 2, \ldots, 8$. A **word of caution** here that not every subset of a sample space is an event. We shall not explain it here as it is beyond the scope of this course.

* * *

The following examples will further strengthen your understanding of various terms defined.

Example 4: Suppose our experiment consists in observing the number of road accidents in a given city on a given day. Obviously, $\zeta = \{0, 1, 2, \ldots, b\}$, where b is the maximum possible number of accidents in a day and this can very well be infinity, in which case the sample space is infinite. The event E that there are five or less number of accidents on that day can be described by $E = \{0, 1, 2, 3, 4, 5\}$.

* * *

Example 5: Suppose, we are interested in noting down the time (in hours) to failure of an equipment. Here, the sample space $\zeta = \{x | 0 < x \leq b\} =]0, b]$, the half open interval between 0 and b, where b is the maximum possible life of the equipment. Here also b may be infinite. In any case, ζ is not only infinite, but also uncountable i.e. the elements of ζ cannot be put into one-to-one correspondence with the set of natural numbers. The event E that the equipment survives for **at least** 500 hours of operation can be described as $E = \{x | 500 \leq x \leq b\} = [500, b]$.

* * *

We shall now consider an example which shows the importance of order of selection of items/objects under consideration.

Example 6: Suppose an urn contains three marbles which are identical in all respect excepting that two of them are red in colour while the third one is white. The experiment is to pick up two marbles, one after the other, blindly and observing their colour. This experiment comprises of two trials. How does the experimenter report the result of the experiment? For this purpose, we may identify the two red marbles as r_1, r_2 and the white marble may be identified as w. Thus, a typical outcome, for instance, can be (w, r_2); this means that the white marble was picked up at the first draw while the second red marble appeared in the next draw. **Note** that the order in which w and r_2 appeared in (w, r_2) is important; the first letter corresponds to the first selection while the second letter corresponds to the second. Thus the sample space ζ of this random experiment is the following set:

$$\zeta = \{(r_1, r_2), (r_1, w), (r_2, r_1), (r_2, w), (w, r_1), (w, r_2)\}$$

How do we describe the 'event E that' of the two marbles picked up, one is red and the other is white'? Well, this event materialises if one of the two red marbles is picked up in one of the two trials while the white marble is picked up in the other trial. Thus, E materialises if the outcome of the experiment is either (r_1, w) or (r_2, w) or (w, r_1) or (w, r_2); thus, $E = \{(r_1, w), (r_2, w), (w, r_1), (w, r_2)\}$.

Let us now modify the conditions of the experiment a little. Suppose, we return the first marble selected back to the urn before the second marble is selected. Then, there will be some new possibilities. Specifically, now it becomes possible for the same marble to be selected twice. Hence, the sample space gets modified to

$$\zeta = \{(r_1, r_1), (r_1, r_2), (r_1, w), (r_2, r_1), (r_2, r_2), (r_2, w), (w, r_1), (w, r_2), (w, w)\}$$

but, the above event E remains unaltered.

Incidentally, in our initial experiment, marbles were being drawn **without replacement** while in the modified experiment, the marbles were being drawn **with replacement**.

Selection is said to be done **with/without replacement** if each object chosen is **returned/not returned** to the population before the next object is drawn

The former experiment is equivalent to drawing two marbles in succession without the previously drawn marble being returned to the urn before the next draw. In the latter experiment the previously drawn marble was being replaced before the next draw, so that the number of marbles in the urn was always the same and it is possible that same marble is chosen more than once.

***

You may now try the following exercises:

E1) In each of the following exercises, an experiment is described. Specify the relevant sample spaces:

a)A machine manufactures a certain item. An item produced by the machine is tested to determine whether or not it is defective.

b)An urn contains six balls, which are coloured differently. A ball is drawn from the urn and its colour is noted.

c)A person is asked to which of the following categories he belongs: unemployed, self-employed, employed by the Central Government, employed by a state or local Government, employed by a private organisation, employed by a employer who does not fulfil any of the above descriptions. The person's answer is recorded.

d)An urn contains ten cards numbered 1 through 10. A card is drawn, its number noted and the card is replaced. Another card is drawn and its number is noted.

E2) Suppose a six-faced die is thrown twice. Describe each of the following events:

i)The maximum score is 6.

ii)The total score is 9.

iii)Each throw results in an even score.

iv)Each throw results in an even score larger than 2.

v)The scores on the two throws differ by at least 2.

It must have been quite clear to you by now that events associated with a random experiment can always be represented by the collection of sample points each of which leads to the occurrence of the said event. The phrase 'event E occurs' is an alternative way of saying that the outcome of the experiment has been one of those that lead to the materialisation of said event. Let us now see how an event can be expressed in terms of two or more events by forming unions, intersections and complements.

2.2.2 Algebra of Events

Let ζ be a fixed sample space. We have already defined an event as a collection of sample points from ζ. Imagine that the (conceptual) experiment underlying ζ is being performed. The phrase "the event E occurs" would mean that the experiment results in an outcome that is included in the event E. Similarly, non-occurrence of the event E would mean that the experiment results into an outcome that is not an element of the event E. Thus, the collection of all sample points that are not included in the event E is also an event which is complementary to E and is denoted as E^c. The event E^c is therefore the event which contains all those sample points of ζ which are not in E. As such, it is easy to see that the event E occurs if and only if the event E^c does not take place. The events E and E^c are **complementary events** and taken together they comprise the entire sample space, i.e., $E \cup E^c = \zeta$.

You may recall that ζ is an event which consists of all the sample points. Hence, its complement is an empty set in the sense that it does not contain any sample point and is called the **null event**, usually denoted as ϕ so that $\zeta^c = \phi$.

Let us once again consider Example 3. Consider the event E that the three tosses produce at least one head. Thus, $E = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$ so that the complementary event $E^c = \{s_8\}$, which is the event of not scoring a head at all.

Again in Example 6 in the case of selection without replacement, event that the white marble is picked up at least once is defined as $E = \{(r_1, w), (r_2, w), (w, r_2)(w, r_1)\}$. Hence, $E^c = \{(r_1, r_2), (r_2, r_1)\}$ i.e. the event of not picking the white marble at all.

Let us now consider two events E and F. We write $E \cup F$, read as **E "union" F**, to denote the collection of sample points, which are responsible for occurrence of either E or F or both. Thus, $E \cup F$ is a new event and it occurs if and only if either E or F or both occur i.e. if and only if at least one of the events E or F occurs. Generalising this idea, we can define a new event $\cup_{j=1}^{k} E_j$, read as **"union" of the k events** $E_1, E_2, \ldots, E_k$, as the event which consists of all sample points that are in at least one of the events $E_1, E_2, \ldots, E_k$ and it occurs if and only if at least one of the events $E_1, E_2, \ldots, E_k$ occurs.

Example 7: Consider Example 3 and let E_j be the event that the three tosses produce j heads, $j = 0, 1, 2, 3$. Hence, $\cup_{j=0}^{2} E_j$ is the event that the three tosses produce at most 2 heads i.e. not all the three tosses produce heads.

* * *

Again, let E and F be two given events. We write $E \cap F$, read as E "**intersection**" F, to denote the collection of sample points any of whose occurrence implies the occurrence of both E and F. Thus, $E \cap F$ is a new event and it occurs if and only if both the events E and F occur. Generalising this idea, we can define a new event $\cap_{j=1}^{k} E_j$ read as "intersection" of the k events $E_1, E_2, \ldots, E_k$, as the event which consists of sample points that are common to each of the events $E_1, E_2, \ldots, E_k$, and it occurs only if all the k events $E_1, E_2, \ldots, E_k$ occur simultaneously.

Further, two events E and F are said to be **mutually exclusive** or **disjoint** if they do not have a common sample point i.e. $E \cap F = \phi$. Two mutually exclusive events then cannot occur simultaneously. In the coin-tossing experiment for instance, the two events, heads and tails, are mutually exclusive: if one occurs, the other cannot occur. To have a better understanding of these events let us once again look at Example 3.

Let E be the event of scoring an odd number of heads and F be the event that tail appears in the first two tosses, so that $E = \{s_1, s_5, s_6, s_7\}$ and $F = \{s_5, s_8\}$. Now $E \cap F = \{s_5\}$, the event that only the third toss yields a head. Thus events E and F are not mutually exclusive.

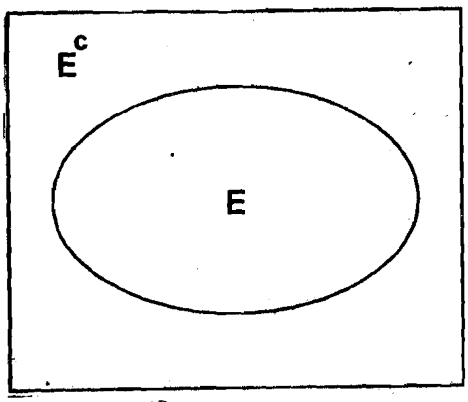

Fig. 1(a)

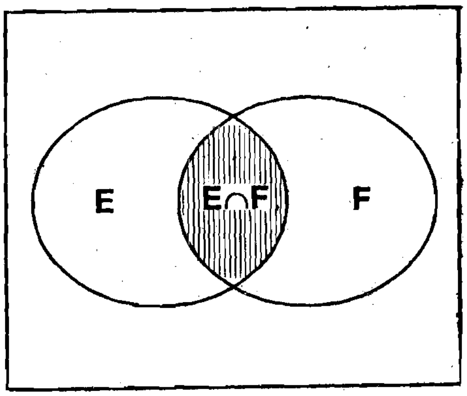

Fig. 1(b)

The above relations between events can be best viewed through a Venn diagram. A rectangle is drawn to represent the sample space ζ. All the sample points are represented within the rectangle by means of points. An event is represented by the region enclosed by a closed curve containing all the sample points leading to that event. The space inside the rectangle but outside the closed curve representing E represents the complementary event E^c (See Fig.1(a) in the previous page.) Similarly, in Fig.1(b), the space inside the curve represented by the broken line represent the event $E \cup F$ and the shaded portion represents $E \cap F$.

Why don't you try the following exercises now.

E3) Suppose A, B and C are events in a sample space ζ. Specify whether the following relations are true or false.:

a) $(A \cap B) \cap (A \cup B) = \phi$

b) $(A^c \cap B^c)^c = A \cup B$

c) $A^c \cap (A \cap B) = A \cup B^c$

d) $A \cap (B \cup B^c) = A \cap (A \cup B) = A$

After the above preliminaries, we are now ready to bring in the concept of probability of events.

2.3 PROBABILITY CONCEPTS

As is clear by now, the outcome of a random experiment being uncertain, none of the various events associated with a sample space can be predicted with certainty before the underlying experiment is performed and the outcome of it is noted. However, some events may intuitively seem to be more likely than the rest. For example, talking about human beings, the event that a person will live 20 years seems to be more likely compared to the event that the person will live 200 years. Such thoughts motivate us to explore if one can construct a scale of measurement to distinguish between likelihoods of various events. Towards this, a small but extremely significant fact comes to our help. Before we elaborate on this, we need a couple of definitions.

Consider an event E associated with a random experiment; suppose the experiment is repeated n times under identical conditions and suppose the event E (which is not likely to occur with every performance of the experiment) occurs $f_n(E)$ times in these n repetitions. Then, $f_n(E)$ is called the **frequency** of the event E in n repetitions of the experiment and $r_n(E) = f_n(E)/n$ is called the **relative frequency** of the event E in n repetitions of the experiment. Let us consider the following example.

Example 8: Consider the experiment of throwing a coin. Suppose we repeat the process of throwing a coin 5 times and suppose the following are the frequencies of a head:

Table 1

No. of repetitions (n)	**Frequency of head** ($f_n(H)$)	**Relative frequency of head** $r_n(H)$
1	0	0
2	1	1/2
3	2	2/3
4	3	3/4
5	3	3/5

Notice that the third column in Table-1 gives the relative frequencies $r_n(H)$ of heads. We can keep on increasing the number of repetitions n and continue calculating the values of $r_n(H)$ in Table 1.

* * *

Merely to fix ideas regarding the concept of probability of an event, we present below a very naive approach which in no way is rigorous, but it helps to see things better at this stage.

2.3.1 Probability of an Event

While the outcome of a random experiment and hence the occurrence of any event E cannot be predicted beforehand, it is interesting to know that the relative frequency $r_n(E)$ of E though may initially fluctuate significantly, would settle down around a constant eventually i.e.

$$r_n(E) \approx r_{n+1}(E) \approx r_{n+2}(E) \approx \dots$$

for all large values of n. We shall not give a formal mathematically rigorous proof of this phenomenon at this stage but we shall try to explain it through a simple argument (we caution you at this stage that this argument is not strictly mathematically rigorous. It merely helps you to visualise that such a result is a possibility).

Observe that the frequency $f_{n+1}(E)$ of E in n+1 repetitions must be equal to either $f_n(E)$ or $f_n(E)+1$ depending on whether E did not or did occur at the $(n+1)$-th repetition. Thus,

$$\begin{aligned}
r_{n+1}(E) - r_n(E) &= f_{n+1}(E)/(n+1) - f_n(E)/n \\
&= \{nf_{n+1}(E) - (n+1)f_n(E)\} / \{(n+1)n\} \\
&= \{nf_n(E) - (n+1)f_n(E)\} / \{(n+1)n\}\text{ , or} \\
&\quad \{n(f_n(E)+1) - (n+1)f_n(E)\} / \{(n+1)n\} \\
&= -f_n(E)/n(n+1), \text{or} \\
&\quad (n - f_n(E)/(n+1)n,
\end{aligned}$$

so that

$$\begin{aligned}
|r_{n+1}(E) - r_n(E)| &= \left|-\frac{f_n(E)/n}{n+1}\right| \le \frac{1}{n+1}, \text{ or} \\
&= \left|\frac{1 - f_n(E)/n}{n+1}\right| \le \frac{1}{n+1}
\end{aligned}$$

as $f_n(E)/n$ being the relative frequency of E in n repetitions can never exceed 1. Thus, the difference between $r_{n+1}(E)$ and $r_n(E)$ can be made as small as we please by increasing the number of repetitions n.

In any case, the constant around which the relative frequency of an event E settles down as the number of repetitions becomes large (i.e. $\lim_{n\to\infty} r_n(E)$) is called the **probability of E** and is denoted as P(E). Thus, P(E) can be interpreted to be the proportion of times one would expect the event E to take place when the underlying random experiment is repeated a large number of times under identical conditions. How large should the number of repetitions be in order for the relative frequency to settle down is another matter. Let us now illustrate through examples what we have discussed above.

Example 9: When we say that the probability of scoring a head when it is tossed is 0.5, we merely mean that 50% of a **large number of tosses** should result in heads. We do not mean that in 10 tosses, 5 will turn heads and 5 tails. However if the coin is tossed N times and N is sufficiently large, then we may expect nearly N/2 heads and N/2 tails.

* * *

Example 10: When we say that the probability of a certain machine failing before 500 hours of operation is 2%, we merely mean that out of a large number of such machines, about 2% of them will fail before 500 hours of operation.

* * *

Evidently, then such a numerical assessment of likelihood of the event in a given situation helps make an appropriate decision. For example, suppose we have two brands A and B of an equipment available in the market such that the probability of failure before 500 hours of operation is 0.02 for an equipment of brand A and 0.10 for an equipment of brand B. We will obviously choose the former variety as the long term proportion of these failing before 500 hours of operation is less.

Notice that the relative frequency of any event E in n repetitions, $r_n(E)$, satisfies the inequalities $0 \le r_n(E) \le 1$ and therefore P(E) the limit of $r_n(E)$ as $n \to \infty$, must satisfy

$$0 \le (P(E) \le 1 \qquad (1)$$

Further, since ζ denotes the event that the experiment will result into one of the outcomes listed in ζ and since by definition, it consists of all possible outcomes, $r_n(\zeta) = 1$, for every n, so that

$$P(\zeta) = 1 \qquad (2)$$

Consider two **mutually exclusive** events E and F. Then, to calculate $r_n(E \cup F)$, i.e., the proportion of times either E or F is observed, we must add the proportion of times E is observed and the proportion of times F is observed so that $r_n(E \cup F) = r_n(E) + r_n(F)$, and hence in the limit,

$$P(E \cup F) = P(E) + P(F), \qquad (3)$$

Remember that E and F cannot occur together as they are mutually exclusive. Arguing in a similar manner, one observes that if $E_1, E_2, \ldots, E_k$ are k mutually exclusive events, then

$$P(\cup_{j=1}^{k} E_j) = P(E_1) + P(E_2) + \ldots + P(E_k) \qquad (4)$$

i.e., the probability of observing at least one of the k mutually exclusive events $E_1, E_2, \ldots, E_k$ (and hence exactly one) is equal to the sum of their individual probabilities.

From the above properties (1)-(4) of probability, the following useful results follow immediately:

PROPOSITION 1: In view of relations (1) and (2), for any event E,

$$1 = P(\zeta) = P(E \cup E^c)$$

so that for any event E,

$$P(E^c) = 1 - P(E). \qquad (5)$$

You will realise later that this result is particularly useful in many problems where it is easier to compute $P(A^c)$ than P(A). So whenever we wish to evaluate P(A), we compute $P(A^c)$ and get the desired result by subtraction from unity

PROPOSITION 2: In Eqn.(5) above, substituting ζ for E, so that $\zeta^c = \phi$, we obtain

$$P(\phi) = 1 - P(\zeta) = 1 - 1 = 0 \qquad (6)$$

Notice that you may come across situations where the converse of result (6) is not true. That is, if P(A) = 0, we cannot in general conclude that $A = \phi$, for, there are situations in which we assign probability zero to an event that can occur.

PROPOSITION 3: Consider a finite sample space $\zeta = \{s_1, s_2, \ldots, s_m\}$. Suppose, from the nature of the experiment, it is reasonable to assume that all the possible outcomes $s_1, s_2, \ldots, s_m$ are equally likely i.e. after taking into consideration all the relevant information pertaining to the experiment, none of the m outcomes seems to be more likely than the rest. For $i = 1, 2, \ldots, m$, let E_i be the simple event defined as

$$E_i = \{s_i\},$$

Since, all these simple events are equally likely,

$$P(E_1) = P(E_2) = \ldots = P(E_m) = p, \text{ say}$$

Now,

$$1 = P(\zeta) = P(\cup_{j=1}^{m} E_j) = P(E_1) + P(E_2) + P(E_3) + \ldots + P(E_m)$$

so that for $i = 1, 2, \ldots, m$,

$$P(E_i) = p = 1/m$$

Also, any event E that consists of k of the m sample points ($k \leq m$) can be represented as the union of the k simple events (which are trivially mutually exclusive). Arguing as above, we conclude that

k events are mutually exclusive if no two of them have any element in common

$$P(E) = kp \tag{7}$$

$$\left.\begin{aligned} &= k/m \\ &= \frac{(\text{Total no. of sample points in E})}{(\text{Total no. of sample points in } \zeta)} \\ &= \frac{(\text{Total no. of outcomes favouring E})}{(\text{Total no. of the outcomes of the experiment}))} \\ &= \frac{(\text{Total no. of ways E can occur})}{(\text{Total no. of the outcomes of the experiment.})} \end{aligned}\right\} \tag{8}$$

Example 11: Let us once again consider Example 3. Whenever the coin is unbiased, H and T are equally likely and then all the eight simple events are equally likely. Then for an event E that the three tosses produce at least one head, Formula (8) gives $P(E) = 7/8$, since E consists of 7 sample points and the sample space S consists of 8 sample points.

* * *

Example 12: Suppose a fair die is rolled once and the score on the top face is recorded. Obviously, the sample space $\zeta = \{1, 2, 3, 4, 5, 6\}$. Since the die is fair, all the six outcomes are equally probable so that the probability of scoring i is $1/6$, for $i = 1, 2, \ldots, 6$. Now, consider the event E that the top face shows an odd score; clearly, $E = \{1, 3, 5\}$. Thus, $P(E) = 3/6 = 1/2$.

* * *

Example 13: Consider two fair dice which are rolled simultaneously and the scores are recorded. Hence, the sample space consisting of $6 \times 6 = 36$ points is

$$\begin{aligned} \zeta &= \{(1,1), (1,2), \ldots, (6,6)\} \\ &= \{(i,j) : i = 1, 2, \ldots, 6; j = 1, 2, \ldots, 6\} \end{aligned}$$

Here, a typical sample point is denoted by (i,j) where i is the score on the first die and j is the score on the second die. Now, consider the event E that the sum of the two scores is 7 or more. Then,
$E = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1), (2,6), (3,5), (4,4), (5,3), (6,2), (3,6), (4,5), (5,4), (6,3), (4,6), (5,5), (6,4), (5,6), (6,5), (6,6)\}$. Since the dice are fair, all the 36 outcomes listed in ζ are equally likely and since E comprises of 21 sample points, $P(E) = 21/36 = 7/12$.

Now consider the event F that none of the two scores is even. To calculate the probability of F, we can surely list out all the outcomes, which favour this event. However, to economise on computations, we may like to use a simple argument as follows. Since none of the scores is even, then, both must be odd. Now, on each die, the possible odd scores are 1,3 or 5. Also with each odd score on the first die, any of the three possible odd scores on the second die may be associated. Thus with each odd score on the first die, there will then be three possible results of the complete experiment and since there are 3 possible odd scores on the first die, the total number of sample points in F is $3 \times 3 = 9$. Thus, $P(F) = 9/36 = 1/4$.

Let us consider another event F' described by the condition that the only possible scores on the two dice is either 1 or 3. Then the number of possible sample points in F' are $2 \times 2 = 4$ and $P(F') = 4/36 = 1/9$.

You may **notice** here that $E \cap F' = \phi$, since the sum of the two scores corresponding to any sample point in F' can be at most 6 so that $P(E \cap F') = 0$.

* * *

Example 14: Consider the case of drawing marbles without replacement as in

Example 6. Suppose, we are interested in working out the probability of the event E that the two marbles drawn are of different colour. When we are reaching out for marbles blindly and the marbles are of identical size, we are not favouring any particular marble so that all the possible pairs are equally likely to be picked up. Now, the total number of sample points in S is 6 and the number of sample points favouring the occurrence of the event E is 4, so that P(E) = 4/6 = 2/3.

* * *

And now some exercises for you

E4) In Example 14 above find the probability of the event E when drawing is done, with replacement.

E5) The manager of a shoe store sells from 0 to 4 pairs of shoes of a recently introduced design every day. Based on experience, the following probabilities are assigned to daily sales of 0, 1,2,3, or 4 pairs:

P(0)	=	0 .08
P(1)	=	0.18
P(2)	=	0.32
P(3)	=	0.30
P(4)	=	0.12
		1.00

a)Are these valid probability assignments? Why or why not?

b)Let A be the event that 2 or fewer are sold in a day. Find P(A).

c)Let B be the event that 3 or more are sold in a day. Find P(B)

Before proceeding further we have the following remarks

Remark 1: For a population of N objects $a_1, a_2, \ldots, a_N$, choosing one object at random from these N objects would mean that the selection mechanism is such that each of the N objects has probability 1/N of being chosen.

Remark 2: For a population of N objects $a_1, a_2, \ldots, a_N$, choosing $n(n \leq N)$ objects at random from these N objects would mean that all the combinations of n objects out of N objects are equally likely. This can be achieved provided the selection mechanism is such that objects are selected one after the other and at each draw all the objects still remaining in the population are equally likely to be chosen.

Let us now take up some more examples to illustrate the use of Formula (8) for calculating probabilities in the cases of selections with/without replacement.

Problem 1: Two balls are drawn at random from a bowl containing 6 balls, of which four balls are white and two are red. What is the probability that (a) both balls are white, (b) both balls are of the same colour, and (c) at least one of the balls is white?

Solution: Case 1: Balls drawn with replacement: Let us number the balls as 1,2,3,4,5 and 6 such that the balls numbered as 1-4 are white while the balls numbered as 5 and 6 are red. We are required to draw 2 balls randomly with replacement. For the first draw, we have 6 choices; since the ball chosen in the first draw is sent back to the bowl before the second ball is drawn, for the second draw also, we have 6 choices. Thus, the sample space $\zeta = \{(i,j)i = 1,2,\ldots,6; j = 1,2,\ldots,6\}$ has $6 \times 6 = 36$ sample points. Let E_j be the event that the j of the two balls selected are white, $j = 0, 1, 2$. Then

a) The required probability of both the balls being white will be $P(E_2)$. In order for E_2 to take place, the outcome (i,j) must be such that both i and j are numbers between 1 and 4. Since the ball chosen in the first draw is sent back to the bowl before the second ball is drawn, the number of sample points belonging to E_2 is $4 \times 4 = 16$. Hence $P(E_2) = 16/36 = 4/9$.

b) In this case the event we are interested in is $E_2 \cup E_0$ because E_2 is the event that both selected balls are white while E_0 is the event that none of the two selected ball is white implying that both must be red. Then the required probability is $P(E_2 \cup E_0) = P(E_2) + P(E_0)$, since these two events are mutually exclusive. In view of an argument of the previous case, the number of sample points belonging to E_0 is $2 \times 2 = 4$. Hence the required probability is $16/36 + 4/36 = 20/36 = 5/9$.

c) Now the event that at least one of the two balls drawn is white is complementary to E_0. Thus the probability that at least one of the two balls drawn is white is $1 - P(E_0) = 1 - 4/36 = 32/36 = 8/9$.

We shall now consider the case of finding probabilities without replacement.

Case 2: Balls drawn without replacement: Here, we are required to draw 2 balls randomly without replacement. For the first draw, we have 6 choices; since the ball chosen in the first draw is not being sent back to the bowl before the second ball is drawn, for the second draw, we have 5 choices. Thus, the sample space $\zeta = \{(i,j) | i = 1,2,\ldots,6; j = 1,2,\ldots,6; i \neq j\}$ has $6 \times 5 = 30$ sample points (combinations of 2 out of 6 balls, keeping in mind the order in which the balls are drawn). Let E_j be the event that the j of the two balls selected are white, j=0,1,2.

a) The required probability will be $P(E_2)$. In order for E_2 to take place, the outcome (i, j) must be such that both i and j are numbers between 1 and 4. Since the ball chosen in the first draw is not sent back to the bowl before the second ball is drawn, the number of sample points belonging to E_2 is $4 \times 3 = 12$. Hence $P(E_2) = 12/30 = 2/5$.

b) Here the event we are interested in is $E_2 \cup E_0$ and as such the required probability is $P(E_2 \cup E_0) = P(E_2) + P(E_0)$, since these two events are mutually exclusive. In view of an argument of the previous case, the number of sample points belonging to E_0 is $2 \times 1 = 2$. Hence the required probability is $12/30 + 2/30 = 14/30 = 7/15$.

c) The event that at least one of the two balls drawn is white is complementary to E_0. Thus the probability that at least one of the two balls drawn is white is $1 - P(E_0) = 1 - 2/30 = 28/30 = 14/15$.

——————×——————

Problem 2: Consider a workshop that employs three mechanics in shifts to repair the machines as and when these fail. During a given shift, 4 machines failed and the repair duties were assigned to the three mechanics through a system of lottery such that for each failed machine, all the three mechanics were equally likely to be chosen. What is the probability that none of the mechanics as well as machine remained idle in that shift?

Solution: Suppose we name the mechanics as 1,2 and 3 and designate the four machines as I, II, III and IV. The result of an allocation of mechanics to the four machines can be represented as (i,j,k,l) meaning that mechanic i is assigned to machine I, mechanic j is assigned to machine II, mechanic k is assigned to machine III and mechanic l is assigned to machine IV, where i,j k and l are each numbers between 1 and 3 designating the concerned mechanic. Note that there being only three mechanics and four machines, i,j,k,l cannot be all different.

Thus, $\zeta = \{(i, j, k, l) | i = 1,2,3; j = 1,2,3; k = 1,2,3; l = 1,2,3\}$. Since the allocation of the same mechanic to more than one machine is feasible and conversely one or two mechanics being idle is also feasible, each possible value of i can be combined with each possible value of j,k and l and since each of i,j,k and l can adopt three values, the total number of sample points in ζ must be $3 \times 3 \times 3 \times 3 = 3^4 = 81$.

Now if all the machines have to have work on that day then one of the mechanics must have been assigned two machines and the other two mechanics one each. This is the only way the event "no machine and no mechanics is idle" can occur. In terms of the sample point (i,j,k,l) this would mean that two of the numbers i,j, k and l must be equal and the other two must be different and different from each other. For example, a typical allocation leading to all the mechanics being engaged and no machine is idle on that day can be (1,1,2,3). To count all the possible allocations so that everybody is engaged, we have to choose two of the indices i,j,k and l that are to be identical and this can be done in ${}^4C_2 = 6$ ways. While these two indices are identical the other two indices can be switched around in 2 ways. Thus, for example, if we allocate machines I and II to mechanic 1, then machines III and IV can be allocated to mechanics 2 and 3 respectively or to mechanics 3 and 2 respectively. Since there are altogether 3 mechanics, the total number of ways the machines can be allocated to them so that all the three mechanics are engaged and no machine is idle will be $3 \times {}^4C_2 \times 2 = 36$. Thus the required probability $= 36/81 = 4/9$.

×

Why don't you try the following exercises now?

E6) Suppose that in a library, two of the six copies of a book are damaged. If the library assistant selects 2 out of the six copies at random, what is the probability that he will select (i) the two damaged copies? (ii) at least one of the two damaged copies?

E7) A student takes a multiple-choice test composed of 100 questions, each with six possible answers. If, for each question, he rolls a fair die to determine the answer to be marked, what is the probability that he answers 20 questions rightly?

We now discuss the rule for finding the probability of the union of any two events, disjoint or not.

2.3.2 Probability of Compound Events

We have seen that for two mutually exclusive events E and F, $P(E \cup F) = P(E) + P(F)$. But this formula cannot be used, for example, to find the probability that at least one of two friends will pass a language examination or the probability that a customer will buy a shirt, a sweater, a belt, or a tie at a departmental store. Both friends can pass the examination and customer of the departmental store can buy any number of these items. To find a formula for $P(E \cup F)$ which holds regardless of whether E and F are mutually exclusive or not, let us consider the following proposition

PROPOSITION 4: For any two events E and F,

$$P(E \cup F) = P(E) + P(F) - P(E \cap F). \qquad (9)$$

Let us draw the Venn diagram as in Fig.2.

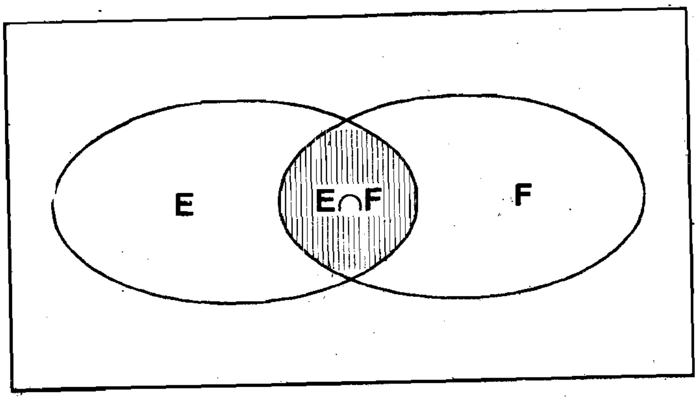

Fig.2

Note that each of the events E and F can be decomposed into mutually exclusive components as

$$E = (E \cap F^c) \cup (E \cap F)$$
$$F = (E^c \cap F) \cup (E \cap F)$$

Then by Formula (3)

$$P(E) = P(E \cap F^c) + P(E \cap F) \quad (10)$$
$$P(F) = P(E^c \cap F) + P(E \cap F) \quad (11)$$

Hence, by summing the two Eqns. (10) and (11), we have

$$\begin{aligned} P(E) + P(F) &= \{P(E \cap F^c) + P(E \cap F) + P(E^c \cap F)\} + P(E \cap F); \\ &= P(E \cup F) + P(E \cap F), \quad (12) \end{aligned}$$

since $(E \cap F^c) \cup (E \cap F) \cup (E^c \cap F) = E \cup F$, our result (9) follows from Eqn.(12).

Observe that the above result reduces to Formula 3 whenever E and F are mutually exclusive; this is so because if E and F are mutually exclusive, then $E \cap F = \phi$ and $P(\phi) = 0$.

PROPOSITION 5: For any k events $E_i, i = 1, 2, \ldots, k$, the above formula can be further generalised for calculation of the probability of occurrence of at least one of these events as follows:

$$P(\cup_{i=0}^{k} E_i) = S_1 - S_2 + S_3 - \ldots + (-1)^{k-1} S_k \quad (13)$$

where,

$$\begin{aligned} S_1 &= P(E_1) + P(E_2) + \ldots + P(E_k) \\ S_2 &= P(E_1 \cap E_2) + P(E_1 \cap E_3) + P(E_1 \cap E_4) + \ldots + P(E_2 \cap E_3) \\ &\quad + P(E_2 \cap E_4) + \ldots + P(E_{k-1} \cap E_k), \\ &\quad \ldots + P(E_{k-2} \cap E_{k-1} \cap E_k), \\ &\vdots \\ S_k &= P(E_1 \cap E_2 \cap E_3 \cap \ldots E_k). \end{aligned}$$

For example, whenever $k = 4$, Formula (13) will be as follows:

$$P(\cup_{i=0}^{4} E_i) = S_1 - S_2 + S_3 - S_4$$

where,

$$\begin{aligned} S_1 &= P(E_1) + P(E_2) + P(E_3) + P(E_4) \\ S_2 &= P(E_1 \cap E_2) + P(E_1 \cap E_3) \\ &\quad + P(E_1 \cap E_4) + P(E_2 \cap E_3) + P(E_2 \cap E_4) + P(E_3 \cap E_4), \\ S_3 &= P(E_1 \cap E_2 \cap E_3) + (E_1 \cap E_2 \cap E_4) + P(E_2 \cap E_3 \cap E_4), \\ S_4 &= P(E_1 \cap E_2 \cap E_3 \cap E_4). \end{aligned}$$

Given Proposition 4, Formula (13) can be proved by induction. We, however, do not discuss the proof here but shall illustrate it through a problem.

Problem 3: Suppose that a candidate has applied for admission to three management schools A,B and C. It is known that his chances of being selected by A,B,C,A and B both, A and C both, B and C both are 0.47, 0.29, 0.22, 0.08, 0.06, 0.07 respectively. It is also known that the chance of his being selected by all the three schools is 0.03. What is the probability that he will be selected by none?

Solution: Suppose A represents the event that he will be selected by the school A; events B and C are similarly defined. From above, we know that
$P(A) = 0.47, P(B) = 0.29, P(C) = 0.22, P(A \cap B) = 0.08,$
$P(A \cap C) = 0.06, P(B \cap C) = 0.07, P(A \cap B \cap C) = 0.03$. Hence,
$S_1 = P(A) + P(B) + P(C) = 0.47 + 0.29 + 0.22 = 0.98;$
$S_2 = P(A \cap B) + P(A \cap C) + P(B \cap C) = 0.08 + 0.06 + 0.07 = 0.21; S_3 = 0.03.$
Thus, as per Proposition 5, the probability that he will be selected by at least one of the

schools A,B and C will be $P(A \cup B \cup C) = S_1 - S_2 + S_3 = 0.98 - 0.21 + 0.03 = 0.80$. Since the event that he will be selected by none is complementary to the event that he will be selected by at least one of the schools A,B and C i.e. $A \cup B \cup C$, the required probability $P(A^c \cap B^c \cap C^c) = 1 - P(A \cup B \cup C) = 1 - 0.80 = 0.20$.

×

Based on Proposition 5, we may further develop an inequality, called Boole's In equality which provides an upper bound for the probability of occurrence of at least one out of k events in terms of the probabilities of these individual events. This inequality obviously provides a way to check probability computations in a situation where it would be computationally involved to apply Formula 13.

PROPOSITION 6: For any k events $E_i, i = 1, 2, \ldots k$ the above formula can be further generalised for calculating the probability of occurrence of at least one of these events as follows:

$$P(\cup_{i=1}^{k} E_i) \leq P(E_1) + P(E_2) + \ldots + P(E_k) \tag{14}$$

Let us define $B_j = \cup_{i=1}^{j} E_i, j = 1, 2, \ldots, k$. Then,

$$\begin{aligned} P(\cup_{i=1}^{k} E_i) &= P(B_k) = P(B_{k-1} \cup E_k) \\ &= P(B_{k-1}) + P(E_k) - P(B_{k-1} \cap E_k) \\ &\leq P(B_{k-1}) + P(E_k) \text{ (since probability of any event} \\ &\qquad\qquad \text{is non-negative)} \\ &\leq P(B_{k-2}) + P(E_{k-1}) + P(E_k) \\ &\leq P(B_{k-3}) + P(E_{k-2}) + P(E_{k-1}) + P(E_k) \\ &\ \vdots \\ &\leq P(E_k) + P(E_2) + \ldots + P(E_{k-2}) + P(E_{k-1}) + P(E_k) \end{aligned}$$

And now a few exercises for you.

E8) Two dice are thrown n times simultaneously. Find the probability that each of the six combinations $(1, 1), (2, 2), \ldots, (6, 6)$ appears at least once?

E9) How many times should an unbiased coin be tossed in order that the probability of observing at least one head be equal to or greater than 0.9?

We now introduce another concept which is important in the study of probability theory.

2.4 CONDITIONAL PROBABILITY AND INDEPENDENT EVENTS

Let ζ be the sample space corresponding to an experiment and E and F are two events of ζ. Suppose the experiment is performed and the outcome is known only partially to the effect that the event F has taken place. Thus there still remains a scope for speculation about the occurrence of the other event E. Keeping this additional piece of information confirming the occurrence of F in view, it would be appropriate to modify the probability of occurrence of E suitably. That such modifications would be necessary can be readily appreciated through two simple instances as follows:

Example 15: Suppose, E and F are such that $F \subset E$ so that occurrence of F would automatically imply the occurrence of E. Thus with the information that the event F has taken place in view, it is plausible to assign probability 1 to the occurrence of E irrespective of its original probability.

* * *

Example 16: Suppose, E and F are two mutually exclusive events and thus they cannot occur together. Thus whenever we come to know that the event F has taken place, we can rule out the occurrence of E. Therefore, in such a situation, it will be appropriate to assign probability 0 to the occurrence of E.

* * *

Example 17: Suppose a pair of balanced dice A and B are rolled simultaneously so that each of the 36 possible outcomes is equally likely to occur and hence has probability $\frac{1}{36}$. Let E be the event that the sum of the two scores is 10 or more and F be the event that exactly one of the two scores is 5.

Then $E = \{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}$ so that $P(E) = 6/36 = 1/6$.

Also, $F = \{(1,5),(2,5),(3,5),(4,5),(6,5),(5,1),(5,2),(5,3),(5,4),(5,6)\}$.

Now suppose we are told that the event F has taken place (note that this is only partial information relating to the outcome of the experiment). Since each of the outcome originally had the same probability of occurring, they should still have equal probabilities. Thus given that exactly one of the two scores is 5 each of the 10 outcomes of event F has probability $\frac{1}{10}$, while the probability of remaining 26 points in the sample space is 0. In the light of the information that the event F has taken place the sample points $(4,6),(6,4),(5,5)$ and $(6,6)$ in the event E must not have materialised. One of the two sample points $(5,6)$ or $(6,5)$ must have materialised. Therefore probability of E would no longer be $1/6$. Since all the 10 sample points in F are equally likely, the revised probability of E which given the occurrence of F, which occur through the materialization of one of the two sample points $(6,5)$ or $(5,6)$ should be $2/10 = 1/5$.

* * *

The probability just obtained is called the conditional probability that E occurs given that F has occurred and is denoted by $P(E|F)$. We shall now derive a general formula for calculating $P(E|F)$.

Consider the following probability table:

Table 2

Events	E	E^c
F	p	q
F^c	r	s

In Table 2, $P(E \cap F) = p, P(E^c \cap F) = q, P(E \cap F^c) = r$ and $P(E^c \cap F^c) = s$ and hence, $P(E) = P((E \cap F) \cup (E \cap F^c)) = P(E \cap F) + P(E \cap F^c) = p + r$ and similarly, $P(F) = q + s$.

Now suppose that the underlying random experiment is being repeated a large number of times, say N times. Thus, taking a cue from the long term relative frequency interpretation of probability, the approximate number of times the event F is expected to take place will be $NP(F) = N(q+s)$. **Under the condition that the event F has taken place**, the number of times the event E is expected to take place would be $NP(E \cap F)$ as both E and F must occur simultaneously. Thus, the long term relative frequency of E under the condition of occurrence of F, i.e. the probability of occurrence of E under the condition of occurrence of F, should be $NP(E \cap F)/NP(F) = P(E \cap F)/P(F)$. This is the proportion of times E occurs out of the repetitions where F takes place.

With the above background, we are now ready to define formally the conditional probability of an event given another.

Definition: Let E and F be two events from a sample space ζ. The **conditional probability** of the event E given the event F, denoted by $P(E|F)$, is defined as

$$P(E|F) = P(E \cap F)/P(F), \text{ whenever } P(F) > 0. \quad (15)$$

When $P(F) = 0$, we say that $P(E|F)$ is undefined. We can also write from Eqn.(15)

$$P(E \cap F) = P(E|F)P(F). \quad (16)$$

Referring back to Example 17, we see that $P(E) = 6/36, P(F) = 10/36$; since, $E \cap F = \{(5,6),(6,5)\}, P(E \cap F) = 2/36$.

From Result (15), $P(E|F) = (2/36)/(10/36) = 2/10 = 1/5$, which is same as that obtained in Example 17.

Result (16) can be generalised to k events $E_1, E_2, \ldots, E_k$, where $k \geq 2$. And now an exercise for you.

E10) Two fair dice are rolled simultaneously. What is the conditional probability that the sum of the scores on the two dice will be 7 given that (i) the sum is odd (ii) the sum is grater than 6, (iii) the outcome of the first die was odd, (iv) the outcome of the second die was even (v) the outcome of at least one of the dice was odd?

Let us now consider some more applications of Formula (16).

Problem 4: A blood disease is present in 12% of a population and is not present in the remaining 88%. An imperfect clinical test successfully detects the disease and with probability 0.90. Thus, if a person has the disease in the serious form, the probability is 0.9 that the test will be positive and it is 0.1, if the test is negative. Moreover, among the unaffected persons, the probability that the test will be positive is 0.05.

(a) A person selected at random from the population is given the test and the result is positive. What is the probability that this person has the disease?

(b) What is the probability that the test correctly detects the disease?

Solution: Let E be the event that a person has the disease and F be the event that the blood test is positive. From the given data, we note that
$P(E) = 0.12, P(F|E) = 0.90, P(F^c|E) = 0.10, P(F|E^c) = 0.05$

(a) We are required to compute $P(E|F)$.

By definition, $P(E|F) = \frac{P(E \cap F)}{P(F)}$.

We have, $P(E \cap F) = P(F|E)P(E) = (0.90)(0.12) = 0.108$

Also,

$$\begin{aligned}
P(F) &= P(F \cap E) + P(F \cap E^c) \\
&= P(F|E)P(E) + P(F|E^c)P(E^c) \\
&= (0.90)(0.12) + (0.05)(0.88) \\
&= 0.108 + 0.044 \\
&= 0.152
\end{aligned}$$

Hence the required conditional probability $P(E|F) = 0.108/0.152 = 0.7105$.

b) Let G be the event that the test correctly detects the disease. Then G will consists of all those who actually have the disease and their blood test is positive and also those who do not have the disease and their blood test is negative.

Then we can write $G = (E \cap F) \cup (E^c \cap F^c)$
so that $P(G) = P(E \cap F) + P(E^c \cap F^c)$
Now $P(E^c \cap F^c) = P(F^c|E^c)P(E^c) = \{1 - P(F|E^c)\} P(E^c) = (0.95)(0.88) = 0.836$
Thus $P(G) = 0.108 + 0.836 = 0.944$

—————×—————

Before we go further you may try these exercises.

E11) Consider a family with two children. Assume that each child is as likely to be a boy as it is to be a girl. What is the conditional probability that both children are boys given that (i) the elder child is a boy (ii) at least one of the children is a boy?

From the definition of conditional probability, you may **observe** that whenever E and F are two mutually exclusive events then $P(E|F) = 0$; similarly, if $F \subset E$, then $P(E|F) = 1$. In each of these cases, the knowledge of occurrence of F gives us a definite idea about the probability of occurrence of E. However, there are many situations where the knowledge of occurrence of some event F hardly have any bearing whatsoever on the occurrence or non-occurrence of another event E. To understand this let us once again consider Problem 1.

Suppose E is the event that the first ball drawn is white and F be the event that the second ball drawn is red. Let us consider the following two cases:

Case 1: Drawing balls with replacement: Arguing as in Problem 1, $P(E \cap F) = (4 \times 2)/36, P(E) = (4 \times 6)/36 = 4/6$ and $P(F) = (6 \times 2)/36 = 2/6$. Hence, $P(E|F) = P(E \cap F)/P(F) = 4/6$ which is the same as the unconditional probability of E.

In other words, the occurrence of F had no influence on the occurrence or non-occurrence of E. In this sense, therefore, the two events E and F are **independent**.

Case 2: Drawing balls without replacement: Arguing as in Problem 1, $P(E \cap F) = (4 \times 2)/30, P(E) = (4 \times 5)/30 = 4/6$ and $P(F) = (5 \times 2)/30 = 2/6$. Hence, $P(E|F) = P(E \cap F)/P(F) = 2/5$ which is different from the unconditional probability of E.

In other words, the occurrence of F had an influence on the occurrence or non-occurrence of E. In this sense, therefore, the two events E and F are **dependent**.

Definition Two events E and F from a given sample space are said to be **independent** if and only if

$$P(E \cap F) = P(E)P(F);$$

otherwise, they are said to be **dependent.**

Again, using the definition of conditional probability, one can equivalently say that two events E and F from a given sample space are independent if and only if

$$\begin{aligned} P(E|F) &= P(E), \text{and} \\ P(F|E) &= P(F). \end{aligned}$$

Consider the following example.

Example 18: Suppose an unbiased coin is tossed twice. Let F be the event that the first toss results in a head and E be the event that the second toss produces a head. The sample space in this case is $\zeta = \{(H,H), (H,T), (T,H), (T,T)\}$ and the coin being unbiased, all these outcomes are equally likely so that each has probability 0.25.

Here, $E = \{(H,H), (T,H)\}$ and $F = \{(H,H), (H,T)\}$ so that $E \cap F = \{(H,H)\}$.

As such, $P(E) = 2/4 = 0.50, P(F) = 2/4 = 0.50, P(E \cap F) = 0.25$.

Thus, $P(E \cap F) = 0.25 = (0.50)(0.50) = P(E)P(F)$.

Hence E and F are independent events.

* * *

And now some exercises for you.

E12) Kalpana (K) and Rahul (R) are taking a statistics course which has only 3 grades A, B and C. The probability that K gets a B is 0.3, the probability that R gets a B is 0.4. The probability that neither gets an A, but at least one gets a B is .42. The probability that K gets an A and R gets a B is 0.08. What is the probability that atleast one gets a B but neither gets a C, (assume the grades of the two students to be independent).

E13) The probability that at least one of two independent events occur is 0.5.

Probability that the first event occurs but not the second is 3/25. Also the probability that the second event occurs but not the first is 8/25. Find the probability that none of the two events occur.

E14) Suppose that A and B are independent events associated with an experiment. If the probability that A or B occurs equals 0.6 while the probability that A occurs equals 0.4, determine the probability that B occurs.

We now end this unit by giving a summary of what we have covered in it.

2.5 SUMMARY

In this unit we have covered the following.

1) Experiments whose outcomes cannot be precisely predicted primarily because the totality of the factors influencing the outcomes are either not identifiable or not controllable at the time of experimentation, even if known, are called **random experiments.**
2) Random experiment may comprise of a series of smaller sub-experiments called **trials**
3) Each outcome of an experiment is a **sample point** and a set of all possible sample points constitute the **sample space** of the experiment.
4) A specific collection or subset of sample points is called an **event.**
5) Two events E and F are **mutually exclusive** or disjoint if they do not have a common sample point i.e. $E \cup F = \phi$. Two mutually exclusive events cannot occur simultaneously.
6) If an experiment is repeated n number of times under identical conditions and an event E occurs $f_n(E)$ times in these n repetitions then $f_n(E)/n$ is the **relative frequency** of the event E in n repetitions of the experiment.
7) The **probability** P(E) of the event E is the proportion of times an event E takes place when the underlying random experiment is repeated a large number of times under identical conditions.
8) For two events E and F, the probability of an event E under the condition that F has already occurred is the **conditional probability** $P(E|F)$ of E.
9) Two events for which the occurrence of one has no influence on the occurrence or non-occurrence of the other are **independent** events, otherwise, they are **dependent.**

2.6 SOLUTIONS/ANSWERS

E1) a) $\zeta = \{g, d\}$, where g stands for "good" or "non-defective" and d stands for "defective"

b) Suppose we code the six colours by the numerals 1, 2, 3, 4, 5 and 6 and only one ball is drawn. Hence $\zeta = \{1,2,3,4,5,6\}$.

c) Let the six categories be coded as 1, 2, 3, 4, 5 and 6. Here, $\zeta = \{1,2,3,4,5,6\}$.

d) Suppose we code the ten cards by the numerals 1, 2, ..., 10. $\zeta = \{(x,y) | 1 \leq x \leq 10, 1 \leq y \leq 10x, y \text{ integers.}\}$

E2) Suppose i is the result on the first throw and j is the result on the second throw, $i = 1,2,\ldots,6$ and $j = 1,2,\ldots,6$. Thus, the sample space is as follows: $S = \{(i,j) : i = 1,2,\ldots,6 \text{ and } j = 1,2,\ldots,6\}$.

(i) Event that the maximum score is 6
$= \{(6, j) : j = 1, 2, \ldots, 6\} \cup \{(i, 6) : i = 1, 2, \ldots, 6\}$.

(ii) Event the total score is 9
$= \{(i, j) : i = 1, 2, \ldots, 6, j = 1, 2, \ldots, 6 \text{ and } i + j = 9\}$

(iii) Event that each throw results in an even score
$= \{(i, j) : i = 2, 4, 6 \text{ and } j = 2, 4, 6\}$

(iv) Event that each throw results in an even score larger than 2
$= \{(i, j) : i = 4, 6 \text{ and } j = 4, 6\}$

v) Event that the scores on the two throws differ by at least 2
$= \{(i, j) : i = 1, 2, \ldots, 6; j = 1, 2, \ldots, 6 \text{ and } |i - j| > 2\}$.

E3) (a) Consider $S = \{i : i = 1, 2, \ldots, 6\}$
Take$A = \{1, 2, 3, 4\}, B = \{3, 4, 5, 6\}$. Then
$(A \cap B) \cap (A \cup B) = \{3, 4\} \cap \{1, 2, 3, 4, 5, 6\} = \{3, 4\} \neq \phi$
so that the statement is false.

(b) $\omega \in (A^c \cap B^c)$ if and only if $\omega \notin (A^c \cap B^c)^c$
ie if and only if ω is neither in A^c nor in B^c
i.e. if and only if ω is in A or in B
i.e. if and only if $\omega \in A \cup B$. Thus, the statement is true.

(c) Take the example of (a) above.
$A^c \cap (A \cap B) = \{5, 6\} \cap \{3, 4\} = \phi \neq \{1, 2, 3, 4\} \cup \{1, 2\} = A \cup B^c$. So the statement is false.

(d) The statement is true, because $B \cup B^c = S$.

E4) In case of drawing with replacement, total number of sample points in the sample space is given by
$\zeta = \{(r_1, r_1)(r_1, r_2), (r_2, r_2)(r_1, w), (w, r_1), (r_2, w), (w, r_2), (w, w), (r_2, r_1)\}$

No. of sample points favouring the occurrence of event E is 4, so that
$P(E) = 4/9$.

E5) (a) Yes, because the non-negative quantities associated with the mutually exclusive and collectively exhaustive simple events add up to 1.

(b) $P(A) = P(0) + P(1) + P(2) = 0.58$.

(c) $P(B) = P(4) + P(3) = 0.42$; Note also that $B = A^c$. Hence
$P(B) = 1 - P(A)$.

E6) Let us code the books as 1, 2, 3, 4, 5 and 6. say, the damaged books coded as 1 and 2. Thus, P(he will elect the two damaged copies)
$\frac{2}{6} \times \frac{1}{5} = \frac{1}{15}$. Also,
P(he will elect at least one of the two damaged copies) = 1 – P(both the books he picks up are the undamaged ones) = $1 - \frac{4}{6} \times \frac{3}{5} = \frac{3}{5}$.

E7) Obviously, the total number of ways in which he can answer all the 100 questions is $6 \times 6 \times 6 \times \cdots \times 6 = 6^{100}$. He can choose the 20 questions to be answered correctly in ${}^{100}C_{20}$ ways. For wrong answers, he can answer each of the 80 questions in 5 ways while each of the 20 questions has one corrrect answer so that each such question can be answered correctly in only one way. Thus exactly 20 questions can be answered correctly and hence the remaining 80 are answered wrongly in

${}^{100}C_{20}\ 1^{20}\ 5^{80}$ ways. Therefore,
P (20 questions are answered correctly} = ${}^{100}C_{20}\ 1^{20}\ 5^{80}/6^{100}$
$= {}^{100}C_{20}(1/6)^{20}(5/6)^{80}$.

E8) There are 36 different results possible for each throw of a pair of dice. Thus, if the pair is thrown n times, the total number of possible results is 36^n. Again, there are 6 identical results on both dice each time the pair is thrown and thus these do not occur in a single throw in 30 different ways; therefore, these do not occur at all in n throws in 30^n ways. So, P(any of the six combinations (1,1), (2,2),..., (6,6) appears at least once} = $1-(30/36)^n$.

E9) Suppose n is the required number. The probability of no head at all in these n tosses is $(1/2)^n$, so that the probability of at least one head is $1 - (1/2)^n$. We need to determine n such that $1 - (1/2)^n \geq 0.9$ i.e. $(1/2)^n \leq 0.1$ i.e. $n\ln(1/2) \leq \ln(1/10)$ i.e. $-n\ln 2 \leq -1$ i.e. $n \geq (1/\ln 2) = 3.32$; Thus, we shall take n as 4.

E10) (i) Because, the events connected with different tosses are independent
P(sum is odd)
= P(score on one die is odd and the score on the other is even)
= 2 P(score on die I is odd and the score on die 2 is even)
= 2 P(score on die I is odd) P(the score on die II is even)
= 2 (3/6)(3/6) = 3/6.
P(sum of scores is 7|sum is odd)
= P(sum of scores is 7$\cap$ sum is odd)/ P(sum is odd)
= P(sum is 7)/ P(sum is odd)
$= (6/36)/(3/6) = 1/3.$

(ii) P(the sum is greater than 6) = 1 – P(sum is 6 or less)
= 1 – (15/36) = 21/36 = 7/12.
P(sum of scores is 7| the sum is greater than 6)
= P(sum is 7)/P(the sum is greater than 6) = (6/36)/(21/36) = 6/21 = 2/7.

(iii) P(the outcome of the first die was odd) = 3/6.
P(sum of scores is 7|the outcome of the first die was odd)
= P(scores of (1, 6) or (3, 4) or (5, 2))/P(the outcome of the first die was odd)
= (3/36)/(3/6) = 1/6.

(iv) P(the outcome of the second die was even)= (3/6)
P(sum of scores is 7| the outcome of the second die was even)
= P(scores of (5, 2) or (3, 4) or (1, 6))/ P(outcome of the second die was even) = (3/36)/(3/6) = 1/6.

(v) P(the outcome of at least one of the dice was odd)
$= 1 - \text{P(outcome of both even)} = 1 - (3\text{x}3/6\text{x}6) = \frac{3}{4}.$
P(sum of scores is 7| the outcome of at least one of the dice was odd)
= P(scores of (1, 6) or (2, 5) or (3, 4) or (4,3) or (5, 2) or (6, 1))/ (3/4)
= (6/36)/(3/4) = 2/9.

E11) (i) P(the elder child is a boy) = $\frac{1}{2}$.
P(both children are boys | the elder child is a boy)
$= \text{P(b, b)}/\{\text{P(b, b)} + \text{P(b. g)}\} = (1/4)/(1/4 + 1/4) = \frac{1}{2}.$

(ii) P(at least one of the children is a boy) = 1 – P(both girls) = 3/4.
P(both children are boys | at least one of the children is a boy)
= P(both children are boys)/ P(at least one of the children is a boy)
= (1/4)/(3/4) = 1/3.

E12) Let K_A be the event that Kalpana gets an A grade; the events K_B, K_C, R_A, R_B and R_C are similarly defined. Given

(a) $P(K_B) = 0.3$,

(b) $P(R_B) = 0.4$,

(c) $P((K_B \cap R_B) \cup (K_B \cap R_C) \cup (K_C \cap R_B)) = 0.42$
i.e. $P(K_B \cap R_B) + P(K_B \cap R_C) + P(K_C \cap R_B) = 0.42$
i.e $P(K_B)P(R_B) + P(K_B)P(R_C) + P(K_C)P(R_B) = 0.42$
i.e. $0.12 + .3\ P(R_C) + .4\ P(K_C) = 0.42$
i.e. $.3\ P(R_C) + .4\ P(K_C) = .30$ (17)

(d) $P(K_A \cap R_B) = .08$
i.e. $P(K_A)\ P(R_B) = .08$
i.e. $P(K_A) = 08/.4 = 0.2.$
From (17), $.3\ P(R_C) = .30 - .4\ P(K_C) = .30 - (.4)(1 - P(K_A) - P(K_B))$
$= .30 - (.4)(1 - 0.2 - 0.3) = 0.10$
so that $P(R_C) = 1/3$.
Therefore, from (a), (b), (c) & (d).
P(at least one gets a B, but neither gets a C)
$= P((K_B \cap R_B) \cup (K_B \cap R_A) \cup (K_A \cap R_B))$
$= P(K_B)P(R_B) + P(K_B)P(R_A) + P(K_A)P(R_B)$
$= (.3)(.4) + (.3)(.2) + (.2)(1/3) = 0.2467.$

E13) Let A and B be the two events. Given:
(a) $P(A \cup B) = 0.5$, (b) $P(A \cap B^c) = 3/25$, (c) $P(A^c \cap B) = 8/25$.

From (a), we have $P(A) + P(B) - P(A)\ P(B) = 0.5$; (18)

from (b) we have, $P(A)\{1 - P(B)\} = 3/25$, i.e. $P(A)\ P(B) = P(A) - 3/25$;
from (c), we have, $\{1 - P(A)\}P(B) = 8/25$, i.e. $P(A)P(B) = P(B) - 8/25$. (19)

Hence, $P(A) - 3/25 = P(B) - 8/25$, so that $P(A) = P(B) - 5/25$. Now, from (18) and (19), $P(B) - 5/25 + P(B) - \{P(B) - 8/25\} = 0.5$,
i.e. $P(B) = 0.5 - 3/25 = 0.38$ which implies $P(A) = P(B) - 5/25 = 0.38 - 0.2 = 0.18$.
Then P(none of the two events occur) $= P(A^c \cap B^c) = P(A^c)\ P(B^c)$
$= \{1 - P(A)\}\{1 - P(B)\} = (.82)(.62) = 0.5084.$

E14) Given : $0.6 = P(A \cup B) = P(A) + P(B) - P(A)\ P(B)$
$= 0.4 + P(B) - 0.4\ P(B)$, so that $P(B) = .2/.6 = 1/3$.