UNIT 1

REVIEW OF BASIC CONCEPTS OF PROBABILITY

Structure

1.1 INTRODUCTION

You know from the background of probability theory studied in earlier classes that it measures uncertainty in numbers. This act of measuring uncertainty in numbers may be done through different approaches to probability. Four approaches to probability which are generally discussed in a probability course are explained in Secs. 1.2 to 1.5. In school mathematics, you have studied conditional probability, addition law, multiplication law, total law of probability and Bayes' theorem with their applications. In this unit, we will give an overview of all these concepts. An overview of conditional probability is given in Sec. 1.6. A brief explanation of addition and multiplication laws of probability is given in Sec. 1.7. After understanding the multiplication law in Sec. 1.7 in general, you will study multiplication law for independent events in Sec. 1.8. Total law of probability is discussed in Sec. 1.9. Finally, Bayes' theorem and its applications are discussed in Sec. 1.10.

What we have discussed in this unit is summarised in Sec. 1.11. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit some more questions based on the entire unit are given in Sec. 1.12 under the heading Terminal

Questions. One of the characteristics of IGNOU study material is that it is self-contained. To make it self-contained regarding solutions of all the SAQ and Terminal Questions we have given solutions of all SAQs and Terminal Questions of this unit in Sec. 1.13. To best utilise this section, it is recommended that after going through all concepts discussed in the unit first, you should try solutions of all SAQs and Terminal Questions yourself. In the case, your answer does not match with the answer given in the solution of the same you may give it another try. If still not matching you may refer to the given solution and identify the concept(s) where you need to improve.

In the next unit, you will study what is a σ-algebra, probability function and probability space.

Expected Learning Outcomes

After completing this unit, you should be able to:

- ❖ explain different approaches to probability theory;
- ❖ explain conditional probability and apply laws of probability theory;
- ❖ identify whether two or more events are independent or not; and
- ❖ apply Bayes' theorem.

1.2 CLASSICAL APPROACH

The classical approach to probability is the oldest approach. To understand the classical definition of probability in a better way first of all we have to explain the meaning of some key terms involved in the definition. These key terms are explained as follows.

Random Experiment: An experiment is said to be random if we know all possible outcomes of the experiment in advance but we cannot predict with certainty which outcome will occur when we perform the experiment. … (1.1)

For example, tossing a coin is a random experiment because we know all possible outcomes (head and tail) in advance but we cannot predict that which one of two will occur when we will toss the coin with certainty. Similarly throwing a die is also a random experiment.

Sample Space: The set of all possible outcomes of a random experiment is called the sample space of the random experiment. It is denoted by Ω. … (1.2)

For example, when a coin is tossed and outcome head is denoted by H and outcome tail is denoted by T then sample space is $\Omega = \{H, T\}$. Similarly, when a die is thrown and six outcomes are denoted by numbers 1 to 6 then sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Event: Any subset of a sample space of a random experiment **of our interest** is known as an event. Keep in mind that the two words "our interest" are used intentionally and are highly important to make this definition valid for both discrete and continuous worlds of probability theory refer (2.2) in the next unit. Why they are so important you will get the answer of this question after completing first four units of this course. … (1.3)

For example, if a die is thrown then getting a multiple of 3 is an event and is given by E = {3, 6}. Full sample space itself and the empty set are also events.

Exhaustive cases: The number of elements in the sample space of a random experiment is known as exhaustive cases. … (1.4)

For example, in the random experiment of tossing a coin number of exhaustive cases is 2.

Favourable cases: Let Ω be the sample space of a random experiment and E be an event, i.e., a subset of Ω then the number of elements in Ω which favour the happening of the event E is known as favourable cases. … (1.5)

For example, favourable cases for the event getting a multiple of 3 in the random experiment of throwing a die are $n\left(\{3, 6\}\right) = 2$.

Mutually Exclusive Outcomes: Let Ω be the sample space of a random experiment and $\omega_1, \omega_2, \omega_3, \ldots, \omega_n$ be the possible outcomes of the random experiment then $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$. If the outcomes ω_i, $1 \le i \le n$ are such that happening of any one of them prevents the happening of any other then we say that outcomes are mutually exclusive or mutually exclusive cases. That is any two of them cannot occur simultaneously outcomes. … (1.6)

For example, in the coin tossing random experiment outcomes head and tail are mutually exclusive because if the head appears then it prevents the appearance of the tail and if the tail appears then it prevents the appearance of the head in the same toss. So, the head and tail both cannot appear simultaneously hence they are mutually exclusive.

Equally Likely Outcomes: Let Ω be the sample space of a random experiment and $\omega_1, \omega_2, \omega_3, \ldots, \omega_n$ be the possible outcomes of the random experiment then $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$. If the random experiment is such that chances of occurring of each ω_i, $1 \le i \le n$ are the same, i.e., we do not have any reason to expect any one of them in preference to any other then we say that outcomes ω_i, $1 \le i \le n$ are equally likely. … (1.7)

For example, when a fair coin is tossed then the outcomes head and tail are equally likely.

We have defined all key terms that will involve in the classical definition of probability. So, we can now define the classical approach to probability.

Classical Approach to Probability Theory

Suppose the sample space of a random experiment is $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$. Suppose the n possible outcomes ω_i, $1 \le i \le n$ are exhaustive, mutually exclusive and equally likely. Let $E \subseteq \Omega$ be any event having m favourable outcomes out of total n outcomes of Ω then the probability of the event E is denoted by $\mathcal{P}(E)$ and is defined as follows.

$$\mathcal{P}(E) = \frac{m}{n} = \frac{\text{Number of favourable outcomes for event E}}{\text{Number of exhaustive cases}} \qquad \ldots (1.8)$$

So, before applying the classical approach, we should make sure that outcomes should be (i) finite in numbers (ii) equally likely and (iii) mutually exclusive.

Let us apply this definition in the following example.

Example 1: Suppose two friends Anjali and Prabhat trying to meet for a date to have a lunch say between 1pm to 2pm. Suppose they follow the following rules for this meeting:

- Each of them will reach either on time or 15 minutes late or 30 minutes late or 45 minutes late or 1 hour late. Each of these arrival times are equally likely for both of them.
- Whoever of them reach first will wait for the other to meet only for 15 minutes. If within 15 minutes the other does not reach, he/she will leave the place and they will not meet.

Find the probability of their meeting.

Solution: Suppose the arrival time of Anjali is represented on the horizontal axis and the arrival time of Prabhat is represented on the vertical axis. As per the rules of the meeting, there are 25 possibilities for their arrival times in total. These 25 possibilities are represented by solid points in Fig. 1.1 (a). But they will meet only when difference between there arrival time is ≤ 15 minutes or ≤ 0.25 hours. Out of total 25 sample space points only 13 points satisfy this requirement which are circled in Fig. 1.1 (b). Hence, if E be the event that they meet then we can find probability of event E using classical approach as follows.

$$\mathcal{P}(\text{E}) = \frac{\text{Favourable cases for event E}}{\text{Number of exhaustive cases}} = \frac{13}{25} = 52\% \qquad \dots (1.9)$$

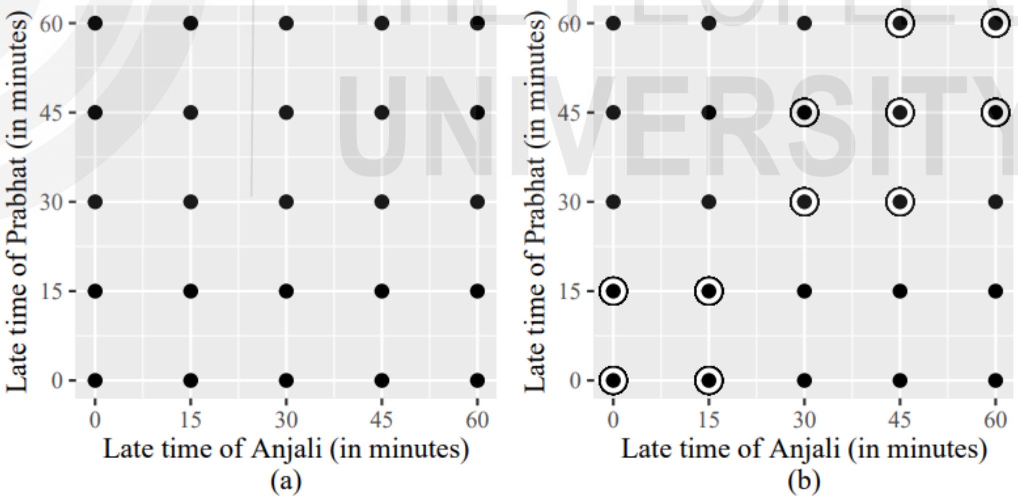

Fig. 1.1: Visualisation of (a) sample space points (b) meeting points are circled

Now, you can try the following Self-Assessment Question.

SAQ 1

Matching letters Problem: From sixth pay commission onwards designations of teachers in universities are assistant professor, associate professor and professor. Consider the two words ASSISTANT and ASSOCIATE. One letter is selected from each of the two words randomly. What is the probability that two selected letters are the same?

1.3 RELATIVE FREQUENCY APPROACH

Classical approach works fine if number of outcomes are finite and they are mutually exclusive and equally likely. But it will not work if any of these three requirements do not meet. So, for such situations, we need other approaches to measure uncertainty in numbers, i.e., probability. One such approach which does not demand for the requirement of equally likely cases is relative frequency approach. But like classical approach the requirement that total number of outcomes should be finite and mutually exclusive is also there in this approach. So, let us explain how this approach works.

You are familiar with frequency distribution from school mathematics. In general, if variable X takes values $x_1, x_2, x_3, \ldots, x_n$ with **absolute frequencies** $f_1, f_2, f_3, \ldots, f_n$ respectively. Then the following tabular form is known as frequency distribution of the variable X.

Variable (X)	x_1	x_2	x_3	…		x_n	Total
Frequencies (f)	f_1	f_2	f_3	…		f_n	N

Suppose $\sum_{i=1}^{n} f_i = N = \text{Sum of all frequencies}$ … (1.10)

Then the frequencies $\frac{f_1}{N}, \frac{f_2}{N}, \ldots, \frac{f_n}{N}$ are known as **relative frequencies**. (1.11)

and the tabular form representing values of X with their corresponding relative frequencies given as follows is known as **relative frequency table**. **…** (1.12)

Variable (X)	x_1	x_2	x_3	…		Total
Frequencies (f)	f_1 / N	f_2 / N	f_3 / N	…	f_n / N	1

This relative frequency table shares two important properties with the probability distribution of a random variable which you have already studied in school mathematics and will also study in Unit 4 of this course are:

- All relative frequencies are ≥ 0, i.e., $f_i / N \geq 0 \ \forall\, i, \ 1 \leq i \leq n$ … (1.13)
- Sum of all relative frequencies is 1, i.e., $\sum_{i=1}^{n} f_i / N = 1$ … (1.14)

Because of these two properties which are necessary for any probability distribution of a random variable, we can use the information available in a frequency distribution to estimate probabilities that variable X takes value $x_i, \ 1 \leq i \leq n$. But keep in mind that when we do so, we are assuming that data available in the form of frequency distribution is representative of the underlying population. We are making this assumption because when we talk about probability then there should be random/uncertainty component. This assumption automatically associates randomness with the data. So, the idea of **relative frequency approach** is related to estimate probabilities from a given frequency distribution. Hence, obtaining probabilities using recorded data is known as relative frequency approach to probability. The recorded data may be secondary or primary. Example 2 is based on this approach. … (1.15)

Statistical Approach: If we assume that an experiment is such that it is feasible to repeat it a large number of times n (say) where mathematically large means $n \to \infty$ under identical conditions then the probability of an event E (say) is defined as the value to which the following limit converges to

$$\lim_{n\to\infty} \frac{m}{n}, \qquad \dots (1.16)$$

where m is the frequency of occurring of the event E out of n trails.

For example, if you toss a fair coin n times and E be the event of getting a head then we know that $\lim_{n\to\infty} \frac{m}{n} \to \frac{1}{2}$, where m is the frequency of occurring head in n trails. You can verify it using simulation in R which is done as follows.

```
> sim<-sample(c("Head", "Tail"), 100, replace = TRUE)
> mean(sim == "Head")
[1] 0.47
> sim<-sample(c("Head", "Tail"), 1000, replace = TRUE)
> mean(sim == "Head")
[1] 0.5
> sim<-sample(c("Head", "Tail"), 10000, replace = TRUE)
> mean(sim == "Head")
[1] 0.5059
> sim<-sample(c("Head", "Tail"), 100000, replace = TRUE)
> mean(sim == "Head")
[1] 0.49931
```

Note that simulated results are close to 0.5. One time it is exactly 0.5. You should try it on your system, you will also observe similar close estimates

Example 2: From a city 100 couples are selected at random and ages at the time of marriage of both husband and wife were recorded. The data thus obtained are given as follows.

		Age of wife				
Age of husband		18-20	20-30	30-40	40-50	50-60
	21-25	7	2	0	0	0
	25-35	3	20	6	0	0
	35-45	0	5	24	3	0
	45-55	0	0	5	12	4
	55-65	0	0	0	4	5

A couple is selected at random then what is the probability that age of husband and wife satisfy the given requirement.

(a) Age of wife is in the interval 20-40?

(b) Age of wife is in the interval 18-30 and the age of husband is in the interval 25-35?

Solution: Here, data is of 100 couple so sum of absolute frequencies is 100. Since this problem is based on recorded data, so, we will use relative frequency approach to estimate probabilities of the given events.

(a) Let E be the event that selected couple is such that age of wife lies in 20-40. So, required probability is given by

$$\mathcal{P}(E) = \frac{2+20+5+6+24+5}{100} = \frac{62}{100} = 62\%$$

(b) Let F be the event that selected couple is such that age of wife lies in 18-30 and age of husband lies in the interval 25-35. So, required probability is given by

$\mathcal{P}(\mathrm{E})=\frac{3+20}{100}=\frac{23}{100}=23\%$

1.4 SUBJECTIVE APPROACH

If outcomes are not equally likely but data (secondary or primary) on required information of interest is available then we can apply relative frequency approach to obtain probability. There are situations where neither the requirements of classical approach are satisfied nor relative frequency approach are satisfied but we still want to measure uncertainty in number. One such approach is subjective approach to probability which is explained as follows.

In this approach, probability of an event of our interest is totally based on the experience, wisdom, intuition, perception, expertise and belief of an individual. He/she assess the probability from his/her own experiences. This approach is applicable in the situations where the events do not occur at all or occur only once or cannot be performed repeatedly under the same conditions. It is interpreted as a measure of degree of belief or as the quantified judgment of a particular individual. For example, a teacher may express his /her confidence that the probability for a particular student getting the first position in a test is 0.99 and that for a particular student getting failed in the test is 0.05. It is based on his/her personal beliefs.

You may notice here that since the assessment is purely subjective one and therefore, it will vary from person to person, depending on one's perception of the situation and past experience. Even when two persons have the same knowledge about the past, their assessment of probabilities may differ according to their personal prejudices and biases. … (1.17)

1.5 AXIOMATIC APPROACH

Classical approach solves lots of problems in probability theory and is commonly used and easily understood even by non-mathematics and statistics background people. But the two restrictions (i) finite number of outcomes in the sample space and of (ii) equally likely outcomes, were the main draw backs. So, there was a need of a definition of probability which go beyond the assumption of finite sample space and equally likely outcomes. Such a general approach was coined by A.N. Kolmogorov known as axiomatic approach which we are going to discuss here.

In Unit 3 of the course MST-011, you have studied what is σ-field and what is a measure (refer 2.10 and 2.23). In addition, if a measure $\mathcal{P}$ also satisfies the condition $\mathcal{P}(\Omega)=1$ then we say that it is a probability measure (refer 2.28 of this course). Here, we will use both σ-field and probability measure or probability function in defining axiomatic approach to probability. Axiomatic approach to probability is explained as follows.

Whenever you do an experiment and want to associate probability with outcomes (or collection of outcomes) of the experiment then first of all you have to specify its sample space (Ω). After specifying sample space, you have to assign probability to each:

- subset of Ω in the case of finite or countable sample space and
- member of some σ-field $\mathcal{F}$ on Ω in the case of uncountable sample space of the experiment

via some function known as probability measure or probability function. This probability measure or probability function should satisfy three axioms of probability theory known as (i) non-negativity, (ii) normalisation and (iii) countable additivity. A formal definition can be given as follows.

Let Ω be the sample space of a random experiment and let $\mathcal{F}$ be the σ-field of the subsets of Ω of our interest and $\mathcal{P}:\mathcal{F}\to[0,1]$ be a probability measure or probability function then axiomatic approach to probability theory coined by A. N. Kolmogorov in 1933 states that $\mathcal{P}$ should satisfy the following three conditions known as axioms of probability theory.

Axioms 1: **Non-negativity**: $0\le\mathcal{P}(E)\le 1$ or $\mathcal{P}(E)\in[0,1]$ … (1.18)

Axioms 2: **Normalisation**: $\mathcal{P}(\Omega)=1$ … (1.19)

Axioms 3: **Countable Additivity**: For any countably infinite sequence $E_1, E_2, E_3, \ldots$ of pairwise mutually disjoint events (i.e., $E_i \cap E_j = \phi,$ for $i \neq j$ where $E_i \in \mathcal{F}$) probability function $\mathcal{P}$ should satisfy

$$\mathcal{P}\left(\bigcup_{i=1}^{\infty} E_i\right)=\sum_{i=1}^{\infty}\mathcal{P}(E_i) \quad \ldots (1.20)$$

Now, we prove some results using these three axioms of probability theory.

Result 1: Prove that probability of empty set is zero using axioms of probability theory. … (1.21)

Proof: We know that

$\Omega\cup\phi=\Omega$

$\Rightarrow\mathcal{P}(\Omega\cup\phi)=\mathcal{P}(\Omega)$

$\Rightarrow\mathcal{P}(\Omega)+\mathcal{P}(\phi)=\mathcal{P}(\Omega)$ $\left[\begin{array}{l}\text{Using axiom (3) or (1.20) for finite events}\\ \text{refer (2.32) in the next unit}\end{array}\right]$

$\Rightarrow 1+\mathcal{P}(\phi)=1$ [Using axiom (2) or (1.19) in LHS and RHS]

$\Rightarrow P(\phi)=0$ … (1.22)

Result 2: Prove that $\mathcal{P}(E^c)=1-\mathcal{P}(E)$ for any event $E\subseteq\Omega.$ using axioms of probability theory. … (1.23)
where E^c or $\overline{E}$ both represents complement of the event E with respect to Ω

Proof: We know that

$E\cup E^c=\Omega$ [Refer Fig. 1.2(a)]

$\Rightarrow\mathcal{P}(E\cup E^c)=\mathcal{P}(\Omega)$ [Applying $\mathcal{P}$ on both sides]

$\Rightarrow\mathcal{P}(E)+\mathcal{P}(E^c)=\mathcal{P}(\Omega)$ $\left[\begin{array}{l}\text{Using axiom (3) or (1.20) for finite events}\\ \text{refer to (2.32) in the next unit}\end{array}\right]$

$\Rightarrow\mathcal{P}(E)+\mathcal{P}(E^c)=1$ [Using axiom (2) or (1.19) in RHS]

$\Rightarrow\mathcal{P}(E^c)=1-\mathcal{P}(E)$ … (1.24)

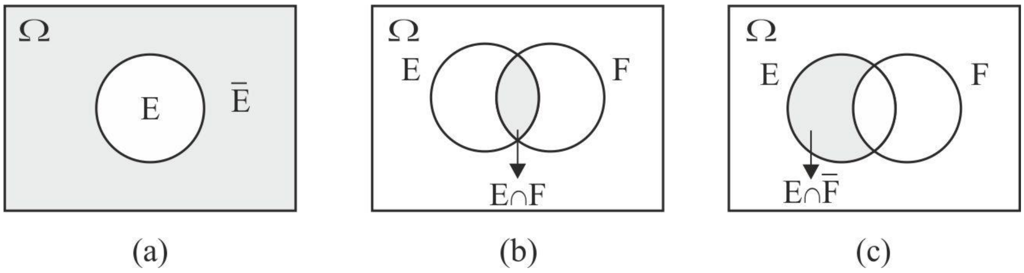

Fig. 1.2: Visualisation of some events (a) Complement of E (b) $E \cap F$, (c) $E \cap F^c$ or $E \cap \bar{F}$

Result 3: Prove that $\mathcal{P}(E \cap F^c) = \mathcal{P}(E) - \mathcal{P}(E \cap F)$ for any events $E, F \subseteq \Omega$.
using axioms of probability theory. … (1.25)
where F^c represents complement of the event F with respect to Ω

Proof: We know that any event E can be written as union of two disjoint events $E \cap F$ and $E \cap F^c$ refer to Fig. 1.2 (b) and (c).

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

$\Rightarrow \mathcal{P}((E \cap F) \cup (E \cap F^c)) = \mathcal{P}(E)$

$\Rightarrow \mathcal{P}(E \cap F) + \mathcal{P}(E \cap F^c) = \mathcal{P}(E) \quad [\text{Using axiom (3) or (1.20) in LHS}]$

$\Rightarrow \mathcal{P}(E \cap F^c) = \mathcal{P}(E) - \mathcal{P}(E \cap F)$ … (1.26)

After understanding various approaches to measure uncertainty in numbers (probability). We now discuss concept of conditional probability in the next section.

1.6 CONDITIONAL PROBABILITY

Let us first explain the idea of conditional probability with the help of an example. Consider the random experiment of throwing an unbiased die. If Ω denotes the sample space of this random experiment then $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let E and F be the events defined as follows.

E: the event of getting a composite number (a number greater than 1 and having at least 3 positive divisors), and

F: the event of getting even number

then, we have, E = {4, 6}, F = {2, 4, 6}, $E \cap F = \{4, 6\}$.

Now, probabilities of the events E, F and $E \cap F$ with respect to the sample space Ω are given by

$$\mathcal{P}(E) = \frac{n(E)}{n(\Omega)} = \frac{2}{6} = \frac{1}{3}, \ \mathcal{P}(F) = \frac{n(F)}{n(\Omega)} = \frac{3}{6} = \frac{1}{2}, \ \mathcal{P}(E \cap F) = \frac{n(E \cap F)}{n(\Omega)} = \frac{2}{6} = \frac{1}{3}.$$

To calculate these probabilities of the events E, F and $E \cap F$, we have the only information that an unbiased die is thrown. Now, suppose other than the information that an unbiased die is thrown you are given some **additional information** that an unbiased die is thrown and an even number is observed. That is the event F has occurred. Now, if with this additional information in hand, we want to calculate probability of event E then it is denoted by $\mathcal{P}(E|F)$, which is read as probability of event E given that event F has been happened or conditional probability of event E given F. Obviously, here our universe will be F because we know that an even number is observed and we are interesting to get a composite number out of the even numbers. Forget about the original sample space Ω. Now, you have new sample space F. So, the probability $\mathcal{P}(E|F)$ is given by

$$\mathcal{P}(E|F) = \frac{\text{No of favourable outcomes for event E in the sample space F}}{\text{Total number of exhaustive outcomes in the sample space F}}$$

$$= \frac{n(E \cap F)}{n(F)} = \frac{2}{3}$$

If you want to connect probability of the event E|F to the original sample space Ω then it can be done by dividing numerator and denominator by cardinality of the original sample space and the same is done as follows.

$$\mathcal{P}(E|F) = \frac{\dfrac{n(E \cap F)}{n(\Omega)}}{\dfrac{n(F)}{n(\Omega)}} = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)}$$

So, conditional probability simply reduces the sample space. Thus, $\mathcal{P}(E)$ is the probability of the event E in the original sample space Ω but $\mathcal{P}(E|F)$ represents probability of the event E in the reduced sample space F. In fact, after getting this idea of conditional probability, you can also interpret $\mathcal{P}(E)$ as conditional probability $\mathcal{P}(E|\Omega)$ but as a convenience we do not specify original sample space as conditioning event because it is always given to us so no need to specify it. Recall that you also follow such type of convention in mathematics. For example, if you want to write $5^{\frac{1}{2}}, 5^{\frac{1}{3}}, 5^{\frac{1}{4}}, 5^{\frac{1}{5}}$, etc. using radical signs then these are written as $\sqrt{5}, \sqrt[3]{5}, \sqrt[4]{5}, \sqrt[5]{5}$, etc. So, you see that when you have exponent 1/2 then you do not specify it. We only specify when we have other than 1/2 exponent. After explaining the idea of conditional probability through this example let us now define conditional probability formerly as follows. … (1.27)

Let Ω be the sample space of a random experiment and E, F be the events. If we are interested in the probability of the event E given that event F has already occurred then it is known as conditional probability of the event E given F and is defined as follows.

$$\mathcal{P}(E|F) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)}, \quad \mathcal{P}(F) \neq 0 \qquad \text{… (1.28)}$$

Similarly, probability of the event F given that event E has already occurred is denoted by $\mathcal{P}(F|E)$ and is given by

$$\mathcal{P}(F|E) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(E)}, \quad \mathcal{P}(E) \neq 0 \qquad \text{… (1.29)}$$

Now, you can try the following Self-Assessment Question.

SAQ 2

Two cards are drawn at random from a well shaffled pack of playing cards. What is the probability that both are kings given that at least one of them is a face card.

1.7 ADDITION AND MULTIPLICATION LAWS OF PROBABILITY

Here we will discuss two laws of probability:

(a) Addition law, and

(b) Multiplication law

Let us discuss these laws one at a time.

1.7.1 Addition Law

In addition law, we need to count number of elements in union of events. But

from set theory studied in earlier classes, you know how to count number of elements in unions of sets. You know this job is done using principle of inclusion-exclusion. You have used principle of inclusion-exclusion in school mathematics for two sets and three sets. Here, you will also see its general case.

(i) Principle of Inclusion-Exclusion for two sets A and B

$n(A \cup B) = n(A) + n(B)$ when sets A and B are mutually disjoint … (1.30)

$n(A \cup B) = n(A) + n(B) - n(A \cap B)$ when A and B are not disjoint … (1.31)

(ii) Principle of Inclusion-Exclusion for three sets A, B and C

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) \quad \ldots (1.32)$$

when sets A, B and C are mutually disjoint

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) - n(C \cap A) + n(A \cap B \cap C) \quad \ldots (1.33)$$

when sets A, B and C are not mutually disjoint

(iii) Principle of Inclusion-Exclusion for n sets $A_1, A_2, A_3, \ldots, A_n$

$$n(A_1 \cup A_2 \cup \ldots \cup A_n) = n(A_1) + n(A_2) + \ldots + n(A_n) = \sum_{i=1}^{n} n(A_i) \quad \ldots (1.34)$$

when sets $A_1, A_2, A_3, \ldots, A_n$ are mutually disjoint

$$n(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} n(A_i) - \sum_{1 \le i < j \le n}^{n} n(A_i \cap A_j) + \sum_{1 \le i < j < k \le n}^{n} n(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n+1} n(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n) \quad \ldots (1.35)$$

when sets $A_1, A_2, A_3, \ldots, A_n$ are not mutually disjoint

Now, if we divide on both sides of each relation by cardinality of the sample space, then each term will convert to probability of the corresponding event. Hence, using this idea of principle of inclusion-exclusion, the addition law of probability can be stated as follows.

(i) Addition law of probability for two events E_1 and E_2

When events E_1 and E_2 are mutually disjoint, then

$$\mathcal{P}(E_1 \cup E_2) = \mathcal{P}(E_1) + \mathcal{P}(E_2) \quad \ldots (1.36)$$

When events E_1 and E_2 are not mutually disjoint, then

$$\mathcal{P}(E_1 \cup E_2) = \mathcal{P}(E_1) + \mathcal{P}(E_2) - \mathcal{P}(E_1 \cap E_2) \quad \ldots (1.37)$$

(ii) Addition law of probability for three events E_1, E_2 and E_3

When events E_1, E_2 and E_3 are mutually disjoint, then

$$\mathcal{P}(E_1 \cup E_2 \cup E_3) = \mathcal{P}(E_1) + \mathcal{P}(E_2) + \mathcal{P}(E_3) \quad \ldots (1.38)$$

When events E_1, E_2 and E_3 are not mutually disjoint, then

$$\mathcal{P}(E_1 \cup E_2 \cup E_3) = \mathcal{P}(E_1) + \mathcal{P}(E_2) + \mathcal{P}(E_3) - \mathcal{P}(E_1 \cap E_2) - \mathcal{P}(E_2 \cap E_3) - \mathcal{P}(E_3 \cap E_1) + \mathcal{P}(E_1 \cap E_2 \cap E_3) \quad \ldots (1.39)$$

(iii) Addition law of probability for n events $E_1, E_2, E_3, \ldots, E_n$

When events $E_1, E_2, E_3, \ldots, E_n$ are mutually disjoint, then

$$\mathcal{P}(E_1 \cup E_2 \cup \ldots \cup E_n) = \mathcal{P}(E_1) + \mathcal{P}(E_2) + \ldots + \mathcal{P}(E_n) = \sum_{i=1}^{n} \mathcal{P}(E_i) \quad \ldots (1.40)$$

When sets $E_1, E_2, E_3, \ldots, E_n$ are not mutually disjoint, then

$$\mathcal{P}(E_1 \cup E_2 \cup \ldots \cup E_n) = \sum_{i=1}^{n} \mathcal{P}(E_i) - \sum_{1 \le i < j \le n}^{n} \mathcal{P}(E_i \cap E_j) + \sum_{1 \le i < j < k \le n}^{n} \mathcal{P}(E_i \cap E_j \cap E_k) - \ldots + (-1)^{n+1} \mathcal{P}(E_1 \cap E_2 \cap E_3 \cap \ldots \cap E_n) \quad \ldots (1.41)$$

Let us discuss one example based on the addition law of probability.

Example 3: A bag contains 4 red, 5 black and 6 white balls. Three balls are drawn at random. What is the probability that at least one colour is missing.

Solution: Let R, B and W be the event explained as follows.

R: red colour ball is missing among three drawn balls

B: black colour ball is missing among three drawn balls

W: white colour ball is missing among three drawn balls

Then we know that at least means union so required probability using the addition law of probability is given by

$$\mathcal{P}(R \cup B \cup W) = \mathcal{P}(R) + \mathcal{P}(B) + \mathcal{P}(W) - \mathcal{P}(R \cap B) - \mathcal{P}(B \cap W) - \mathcal{P}(W \cap R) + \mathcal{P}(R \cap B \cap W)$$

$$= \frac{\binom{11}{3}}{\binom{15}{3}} + \frac{\binom{10}{3}}{\binom{15}{3}} + \frac{\binom{9}{3}}{\binom{15}{3}} - \frac{\binom{6}{3}}{\binom{15}{3}} - \frac{\binom{4}{3}}{\binom{15}{3}} - \frac{\binom{5}{3}}{\binom{15}{3}} + 0$$

$$= \frac{3!12!}{15!}\left[\frac{11!}{3!8!} + \frac{10!}{3!7!} + \frac{9!}{3!6!} - \frac{6!}{3!3!} - \frac{4!}{3!1!} - \frac{5!}{3!2!}\right] \quad \left[\because \binom{n}{r} = \frac{n!}{r!\,(n-r)!}\right]$$

$$= \frac{12!}{15!}\left[\frac{11!}{8!} + \frac{10!}{7!} + \frac{9!}{6!} - \frac{6!}{3!} - \frac{4!}{1!} - \frac{5!}{2!}\right]$$

$$= \frac{1}{15 \times 14 \times 13}\left[11 \times 10 \times 9 + 10 \times 9 \times 8 + 9 \times 8 \times 7 - 6 \times 5 \times 4 - 24 - 60\right]$$

$$= \frac{1}{15 \times 14 \times 13}\left[990 + 720 + 504 - 120 - 24 - 60\right]$$

$$= \frac{2010}{15 \times 14 \times 13} = \frac{67}{91}$$

Now, we discuss multiplication law in the next subsection.

1.7.2 Multiplication Law

In set theory addition means union and multiplication means intersection. We have discussed the addition law of probability which gives probability of union of events. Similarly, the multiplication law deals with probability of intersection of events. In conditional probability, we have seen that

$$\mathcal{P}\left(E \mid F\right) = \frac{\mathcal{P}\left(E \cap F\right)}{\mathcal{P}(F)}, \quad \mathcal{P}(F) \neq 0 \quad \text{and} \quad \mathcal{P}\left(F \mid E\right) = \frac{\mathcal{P}\left(E \cap F\right)}{\mathcal{P}(E)}, \quad \mathcal{P}(E) \neq 0$$

which give $\mathcal{P}\left(E \cap F\right) = \mathcal{P}(F)\mathcal{P}\left(E \mid F\right)$ and $\mathcal{P}\left(E \cap F\right) = \mathcal{P}(E)\mathcal{P}\left(F \mid E\right)$ … (1.42)

The results shown in equation (1.42) are known as **multiplication law** of probability for **two events** E and F.

Similarly, the **multiplication law** for **three events** E_1, E_2 and E_3 which occur in succession is given by

$$\mathcal{P}\left(E_1 \cap E_2 \cap E_3\right) = \mathcal{P}(E_1)\mathcal{P}\left(E_2 \mid E_1\right)\mathcal{P}\left(E_3 \mid E_1 \cap E_2\right) \qquad \ldots (1.43)$$

In general the **multiplication law** for **n events** $E_1, E_2, \ldots, E_n$ which occur in succession is given by

$$\mathcal{P}\left(E_1 \cap E_2 \cap \ldots \cap E_n\right) = \mathcal{P}(E_1)\mathcal{P}\left(E_2 \mid E_1\right)\ldots\mathcal{P}\left(E_n \mid E_1E_2\ldots E_{n-1}\right) \qquad \ldots (1.44)$$

where $E_1E_2E_3\ldots E_j$ represents $E_1 \cap E_2 \cap \ldots \cap E_j$, etc.

Now, you can try the following Self-Assessment Question.

SAQ 3

In a bag there are 5 red, 2 blue and 8 white balls. Three balls are drawn one by one without replacement. Find the probability that two of them are red and one white.

1.8 INDEPENDENCE OF EVENTS

If happening or non-happening of an event does not provide any information about the happening or non-happening of another event then we say that two events are independent. For example, if E be the event of getting a multiple of 3 when a die is thrown and F be the event of getting an even number, then events E and F are independent because happening or non-happening of E does not provide any information about happening or non-happening of event F and vice-versa. But if E be the event of getting an odd number and F be the event of getting an even number then the two events are not independent because happening of E tells us that event F did not happen. So, by definition for independent two events for E and F, we should have

$$\mathcal{P}(E|F) = \mathcal{P}(E) \text{ and } \mathcal{P}(F|E) = \mathcal{P}(F) \quad \ldots (1.45)$$

But by conditional probability, we know that

$$\mathcal{P}(E|F) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)}, \ \mathcal{P}(F) \neq 0 \text{ and } \mathcal{P}(F|E) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(E)}, \ \mathcal{P}(E) \neq 0 \quad \ldots (1.46)$$

Using (1.45) in (1.46), we get

$$\mathcal{P}(E) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)}, \ \mathcal{P}(F) \neq 0 \text{ and } \mathcal{P}(F) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(E)}, \ \mathcal{P}(E) \neq 0$$

From both equations, we have

$$\mathcal{P}(E \cap F) = \mathcal{P}(E)\mathcal{P}(F) \quad \ldots (1.47)$$

This is the condition for independent of two events. This equation has one more advantage over equations (1.46) is that here we have no need to put the restrictions $\mathcal{P}(E) \neq 0$ and $\mathcal{P}(F) \neq 0$. Thus, in examples to check independence of two events we will check condition specify in equation (1.47).

Similarly, three events E_1, E_2, E_3 are said to be independent if following four requirements hold, three under (a) and one under (b):

(a) Events E_1, E_2, E_3 taken two at a time should be independent, i.e.,

- $\mathcal{P}(E_1 \cap E_2) = \mathcal{P}(E_1)\mathcal{P}(E_2)$
- $\mathcal{P}(E_2 \cap E_3) = \mathcal{P}(E_2)\mathcal{P}(E_3)$
- $\mathcal{P}(E_1 \cap E_3) = \mathcal{P}(E_1)\mathcal{P}(E_3)$

(b) Events E_1, E_2, E_3 taken all the three at a time should be independent, i.e.,

- $\mathcal{P}(E_1 \cap E_2 \cap E_3) = \mathcal{P}(E_1)\mathcal{P}(E_2)\mathcal{P}(E_3)$ … (1.48)

Similarly, n events $E_1, E_2, E_3, \ldots, E_n$ are said to be independent if any subset of these events is independent. That is events taken two at a time, three at a time, four at a time and so on n at a time all should be independent. … (1.49)

Let us study three important results regarding independent of events which we will use in this course and will also be used in some other courses.

Result 4: If events E and F are independent then prove that

(a) Events E and F^c are independent

(b) Events E^c and F are independent

(c) Events E^c and E^c are independent … (1.50)

Proof: We are given that events E and F are independent so, we have

$$\mathcal{P}(E \cap F) = \mathcal{P}(E)\mathcal{P}(F) \quad \ldots (1.51)$$

(a) We have to show that $\mathcal{P}(E \cap F^c) = \mathcal{P}(E)\mathcal{P}(F^c)$

We know that $\mathcal{P}(E \cap F^c) = \mathcal{P}(E) - \mathcal{P}(E \cap F)$ [Using (1.26)]

$$\therefore \ \mathcal{P}(E \cap F^c) = \mathcal{P}(E) - \mathcal{P}(E)\mathcal{P}(F) \quad \text{[Using (1.51)]}$$

$$= \mathcal{P}(E)(1 - \mathcal{P}(F))$$

$$= \mathcal{P}(E)\mathcal{P}(F^c) \quad \text{[Using (1.24)]}$$

(b) We have to show that $\mathcal{P}\left(E^c \cap F\right) = \mathcal{P}\left(E^c\right)\mathcal{P}(F)$

We know that $\mathcal{P}\left(E^c \cap F\right) = \mathcal{P}(F) - \mathcal{P}\left(E \cap F\right)$

$$= \mathcal{P}(F) - \mathcal{P}(E)\mathcal{P}(F) \qquad [\text{Using } (1.51)]$$

$$= \mathcal{P}(F)\left(1 - \mathcal{P}(E)\right)$$

$$= \mathcal{P}(F)\mathcal{P}(E^c) \qquad [\text{Using } (1.24)]$$

(c) We have to show that $\mathcal{P}\left(E^c \cap F^c\right) = \mathcal{P}\left(E^c\right)\mathcal{P}\left(F^c\right)$

Now, $\mathcal{P}\left(E^c \cap F^c\right) = \mathcal{P}\left(\left(E \cup F\right)^c\right) \qquad [\text{Using De Morgan's law}]$

$$= 1 - \mathcal{P}\left(E \cup F\right) \qquad [\text{Using } (1.24)]$$

$$= 1 - \left[\mathcal{P}(E) + \mathcal{P}(F) - \mathcal{P}(E \cap F)\right] \qquad \begin{bmatrix}\text{Using addition law of probability} \\ \text{refer } (1.37)\end{bmatrix}$$

$$= 1 - \mathcal{P}(E) - \mathcal{P}(F) + \mathcal{P}(E)\mathcal{P}(F) \qquad [\text{Using equation } (1.51)]$$

$$= \left(1 - \mathcal{P}(E)\right) - \mathcal{P}(F)\left(1 - \mathcal{P}(E)\right)$$

$$= \left(1 - \mathcal{P}(E)\right)\left(1 - \mathcal{P}(F)\right)$$

$$= \mathcal{P}\left(E^c\right)\mathcal{P}\left(F^c\right) \qquad [\text{Using } (1.24)]$$

Now, you can try the following Self-Assessment Question.

SAQ 4

Probabilities that students A, B and C solve a randomly selected problem from a book are 1/2, 1/3 and 1/4 respectively. If a randomly selected problem from this book is given to these students, then what is the probability that problem is solved.

After studying the addition and multiplication laws of probability, we now discuss the total law of probability in the next section.

1.9 TOTAL LAW OF PROBABILITY

Total law of probability is basically based on the idea of partition of a sample space. So, before defining total law of probability let us first define what we mean by partition of a set.

Partition of a set: Let X be any set then we say that a collection $\{E_1, E_2, E_3, \ldots, E_n\}$ of subsets of X form a partition of X if its members satisfy the following three conditions (refer Fig. 1.3):

(i) $E_i \neq \phi$, for $1 \leq i \leq n$ (ii) $E_i \cap E_j = \phi$ for $i \neq j$ and (iii) $\bigcup_{i=1}^{n} E_i = X$. … (1.52)

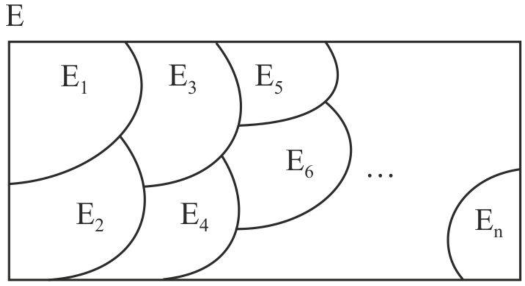

Fig. 1.3: Visualisation of partition of a set X by n subsets E_1, E_2, E_3, …, E_n

Now, we discuss total law of probability. In total law of probability to obtain probability of an event we use conditioning on n events which partition the sample space. Total law of probability is stated and proved as follows.

Statement: Let Ω be the sample space and events $E_1, E_2, \ldots, E_n$ partition Ω, i.e., (i) $E_i \neq \phi$, for $1 \leq i \leq n$ (ii) $E_i \cap E_j = \phi$ for $i \neq j$ and (iii) $\bigcup_{i=1}^{n} E_i = \Omega.$

Then probability of any event A where $A \subseteq \bigcup_{i=1}^{n} E_i$ can be obtained by conditioning on each E_i and is given by

$$\mathcal{P}(A) = \mathcal{P}(E_1)\mathcal{P}(A \mid E_1) + \mathcal{P}(E_2)\mathcal{P}(A \mid E_2) + \ldots + \mathcal{P}(E_n)\mathcal{P}(A \mid E_n)$$

$$= \sum_{i=1}^{n} \mathcal{P}(E_i)\mathcal{P}(A \mid E_i) \qquad \ldots (1.53)$$

Proof: Since $A \subseteq \Omega$ so, we have

$$A = A \cap \Omega = A \cap (E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_n) \qquad [\text{Using (iii) of the statement}]$$

$$A = (A \cap E_1) \cup (A \cap E_1) \cup \ldots \cup (A \cap E_n) \qquad \left[\begin{array}{l}\text{Using distributive}\\ \text{law of set theory}\end{array}\right]$$

$$\Rightarrow \mathcal{P}(A) = \mathcal{P}((A \cap E_1) \cup (A \cap E_1) \cup \ldots \cup (A \cap E_n))$$

Events $A \cap E_1, A \cap E_2, \ldots, A \cap E_n$ are disjoint so using addition for disjoint events, we have

$$\mathcal{P}(A) = \mathcal{P}(A \cap E_1) + \mathcal{P}(A \cap E_2) + \ldots + \mathcal{P}(A \cap E_n)$$

$$= \mathcal{P}(E_1)\mathcal{P}(A \mid E_1) + \mathcal{P}(E_2)\mathcal{P}(A \mid E_2) + \ldots + \mathcal{P}(E_n)\mathcal{P}(A \mid E_n)$$

$$[\text{Using multiplication law of probability, refer to (1.42)}]$$

Hence, $\mathcal{P}(A) = \sum_{i=1}^{n} \mathcal{P}(E_i)\mathcal{P}(A \mid E_i)$

Let us apply total law of probability in the following example.

Example 4: In a factory there are three machines X, Y, Z which produces 20%, 30% and 50% items respectively. Past experience shows that percentage of defective items produced by machines X, Y, Z are 10%, 5%, 2% respectively. If an item is selected at random from the production of these machines. What is the probability that it is a defective item.

Solution: Let E_1, E_2, E_3 be the events that selected item is produced by machine X, Y, Z respectively. Let A be the event that selected item is defective. Then in usual notations, we are given.

$$\mathcal{P}(E_1) = 20\% = \frac{20}{100},\ \mathcal{P}(E_2) = 30\% = \frac{30}{100},\ \mathcal{P}(E_3) = 50\% = \frac{50}{100},$$

$$\mathcal{P}(A \mid E_1) = 10\% = \frac{10}{100},\ \mathcal{P}(A \mid E_2) = 5\% = \frac{5}{100},\ \mathcal{P}(A \mid E_3) = 2\% = \frac{2}{100}$$

Using total law of probability required probability is given by

$$\mathcal{P}(A) = \mathcal{P}(E_1)\mathcal{P}(A \mid E_1) + \mathcal{P}(E_2)\mathcal{P}(A \mid E_2) + \mathcal{P}(E_3)\mathcal{P}(A \mid E_3)$$

$$= \frac{20}{100} \times \frac{10}{100} + \frac{30}{100} \times \frac{5}{100} + \frac{50}{100} \times \frac{2}{100} = \frac{200 + 150 + 100}{10000}$$

$$= \frac{450}{10000} = 0.045 = 4.5\%$$

Now, you can try the following Self-Assessment Question.

SAQ 5

Parking Problem: In a parking there are space of parking $n\ (\geq 2)$ cars. The design of the car parking is such that there is only one row for parking, i.e., the n car parking spaces are consecutive. Also, only one car can be parked in a parking space. Our old friends Prabhat and Anjali reach in this empty parking to park their cars. If both Prabhat and Anjali park their cars randomly in this parking, then what is the probability that there is at the most one empty parking space between their cars?

1.10 BAYES' THEOREM

Bayes' theorem is also known as inverse probability theorem because it reverses conditioning event, i.e., it connects $\mathcal{P}(E\,|\,F)$ and $\mathcal{P}(F\,|\,E)$. Using this theorem, we can also revise probability after getting any additional information. You will see it in Example 5. Let us first state and prove this theorem.

Statement: Let Ω be the sample space and events $E_1, E_2, \ldots, E_n$ partition Ω, i.e., (i) $E_i \neq \phi$, for $1 \leq i \leq n$ (ii) $E_i \cap E_j = \phi$ for $i \neq j$ and (iii) $\bigcup_{i=1}^{n} E_i = \Omega$. Then for any event A with P(A) > 0, conditional probability for each E_i given A is given by

$$\mathcal{P}(E_i\,|\,A) = \frac{\mathcal{P}(E_i)\mathcal{P}(A\,|\,E_i)}{\sum_{j=1}^{n}\mathcal{P}(E_j)\mathcal{P}(A\,|\,E_j)},\quad 1 \leq i \leq n \qquad \ldots (1.54)$$

Proof: Since $A \subseteq \Omega$ so, we have

$$A = A \cap \Omega = A \cap (E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_n) \quad [\text{Using (iii) of the statement}]$$

$$A = (A \cap E_1) \cup (A \cap E_1) \cup \ldots \cup (A \cap E_n) \quad \left[\begin{array}{l}\text{Using distributive}\\ \text{law of set theory}\end{array}\right]$$

$$\Rightarrow \mathcal{P}(A) = \mathcal{P}((A \cap E_1) \cup (A \cap E_1) \cup \ldots \cup (A \cap E_n))$$

$$= \mathcal{P}(A \cap E_1) + \mathcal{P}(A \cap E_2) + \ldots + \mathcal{P}(A \cap E_n)$$

$$[\text{Using addition law of probability, refer (1.40)}]$$

$$= \mathcal{P}(E_1)\mathcal{P}(A\,|\,E_1) + \mathcal{P}(E_2)\mathcal{P}(A\,|\,E_2) + \ldots + \mathcal{P}(E_n)\mathcal{P}(A\,|\,E_n)$$

$$[\text{Using multiplication law of probability, refer (1.42)}]$$

Hence, $$\mathcal{P}(A) = \sum_{i=1}^{n}\mathcal{P}(E_i)\mathcal{P}(A\,|\,E_i) \qquad \ldots (1.55)$$

Now,

$$\mathcal{P}(E_i\,|\,A) = \frac{\mathcal{P}(E_i \cap A)}{\mathcal{P}(A)} \quad [\text{By definition of conditional probability}]$$

$$= \frac{\mathcal{P}(E_i)\mathcal{P}(A\,|\,E_i)}{\mathcal{P}(A)} \quad \left[\begin{array}{l}\text{Using multiplication law of}\\ \text{probability in numerator}\end{array}\right]$$

$$= \frac{\mathcal{P}(E_i)\mathcal{P}(A | E_i)}{\sum_{i=1}^{n}\mathcal{P}(E_i)\mathcal{P}(A | E_i)} \quad [\text{Using equation (1.55)}]$$

Note that denominator in this final expression is exactly the same as we obtained in total law of probability.

Remark 1: Bayes' theorem actually **revises the probabilities** $\mathcal{P}(E_i),\ 1 \le i \le n$ after getting some additional information that event A has happened and the **revised probabilities** are denoted by $\mathcal{P}(E_i | A),\ 1 \le i \le n.$ So, we call the probabilities $\mathcal{P}(E_i),\ 1 \le i \le n$ as **prior probabilities** and probabilities $\mathcal{P}(E_i | A),\ 1 \le i \le n$ are called as **posterior probabilities**. … (1.56)

Let us apply Bayes' theorem of probability in the following example.

Example 5: [Extended statement of Example 4] In a factory there are three machines X, Y, Z which produces 20%, 30% and 50% items respectively. Past experience shows that percentage of defective items produced by machines X, Y, Z are 10%, 5%, 2% respectively. An item from the production of these machines is selected at random and it is found defective. What is the probability that it is produced by (i) machine X (ii) machine Y (iii) machine Z.

Solution: Events E_1, E_2, E_3 and A are the same as defined in Example 4. So, complete solution of Example 4 will be used here and so we are not doing that again. Assuming those calculation in hand let us apply Bayes' theorem to obtain required probability.

(i) $$\mathcal{P}(E_1 | A) = \frac{\mathcal{P}(E_1)\mathcal{P}(A | E_1)}{\sum_{j=1}^{3}\mathcal{P}(E_j)\mathcal{P}(A | E_j)} \quad [\text{Using Bayes' theorem}]$$

$$\mathcal{P}(E_1 | A) = \frac{\frac{20}{100} \times \frac{10}{100}}{\frac{450}{10000}} \quad \left[\begin{array}{l}\text{Using calculated value of}\\ \text{Example 4 in denominator}\end{array}\right]$$

$$= \frac{200}{450} = \frac{4}{9} \simeq 0.4444 = 44.44\% \qquad \text{... (1.57)}$$

(ii) Similarly, $$\mathcal{P}(E_2 | A) = \frac{\frac{30}{100} \times \frac{5}{100}}{\frac{450}{10000}} \quad \left[\begin{array}{l}\text{Using calculated value of}\\ \text{Example 4 in denominator}\end{array}\right]$$

$$= \frac{150}{450} = \frac{1}{3} \simeq 0.3333 = 33.33\% \qquad \text{... (1.58)}$$

(iii) $$\mathcal{P}(E_2 | A) = \frac{\frac{50}{100} \times \frac{2}{100}}{\frac{450}{10000}} \quad \left[\begin{array}{l}\text{Using calculated value of}\\ \text{Example 4 in denominator}\end{array}\right]$$

$$= \frac{100}{450} = \frac{2}{9} \simeq 0.2222 = 22.22\% \qquad \text{... (1.59)}$$

Remark 2: To see the effect of additional information that drawn ball is found to be defective, you can compare the prior and posterior probabilities shown in first and second rows of the following table.

$P(E_1) = 20\%$	$P(E_2) = 30\%$	$P(E_3) = 50\%$	**prior**
$P(E_1 \mid A) = 44.44\%$	$P(E_2 \mid A) = 33.33\%$	$P(E_3 \mid A) = 22.22\%$	**posterior**

Note that prior to this information probability that a randomly selected item is produced by machine X was 20% while after getting this information it is revised to 44.44%. Similarly, you can compare these probabilities for machines Y and Z. Think why a jump from 20% to 44.44% happened it is because machine X produces more defective items in percentage compare to machines Y and Z. Similar reasoning works for the machine Y. But for machine Z there is a decrease from 50% to 22.22%. Its reason is that machine Z produces fewer defective items in percentage compare to machines X and Y so if selected item is found to be defective then there will be less chances that it is produced by machine Z.

1.11 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- **Random Experiment**: An experiment is said to be random if we know all the possible outcomes of the experiment in advance but we cannot predict with certainty that which outcome will occur when we will perform the experiment.
- **Sample Space**: Set of all possible outcomes of a random experiment is called sample space of the random experiment.
- **Event**: Any subset of a sample space of a random experiment of our interest is known as event.
- **Exhaustive cases**: The number of elements in the sample space of a random experiment is known as exhaustive cases.
- **Favourable cases**: Let Ω be the sample space of a random experiment and E be an event then the number of elements in Ω which favour the happening of the event E are known as favourable cases.
- **Mutually Exclusive Outcomes**: Let Ω be the sample space of a random experiment and $\omega_1, \omega_2, \omega_3, \ldots, \omega_n$ be the possible outcomes of the random experiment then $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$. If the outcomes $\omega_i,\ 1 \le i \le n$ are such that happening of any one of them prevents the happening of any other then we say that outcomes are mutually exclusive.
- **Equally Likely Outcomes**: Let Ω be the sample space of a random experiment and $\omega_1, \omega_2, \omega_3, \ldots, \omega_n$ be the possible outcomes of the random experiment then $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$. If the random experiment is such that chances of occurring of each $\omega_i,\ 1 \le i \le n$ are the same, i.e., we do not have any reason to expect any one of them in preference to any other then we say that outcomes $\omega_i,\ 1 \le i \le n$ are equally likely.
- **Classical Approach to Probability Theory**: Suppose sample space of a random experiment is $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$. Suppose the n possible outcomes $\omega_i,\ 1 \le i \le n$ are exhaustive, mutually exclusive and equally likely.

Let $E \subseteq \Omega$ be any event having m favourable outcomes out of total n outcomes of Ω then probability of the event E is denoted by $\mathcal{P}(E)$ and defined as follows. $\mathcal{P}(E) = \frac{m}{n}$

- The idea of **relative frequency approach** to probability is related to estimate probabilities from a given frequency distribution.
- **Statistical Approach**: If we assume that an experiment is such that it is feasible to repeat it a large number of times n (say) where mathematically large means $n \to \infty$ under identical conditions then probability of an event E (say) is defined as the value to which the following limit converges to

  $\lim_{n\to\infty} \frac{m}{n}$, where m is the frequency of occurring of the event E out of n trails.
- In **subjective approach** probability of an event of our interest is totally based on the experience, wisdom, intuition, perception, expertise and belief of an individual. He/she assess the probability from his/her own experiences.
- **Axiomatic approach to probability** states that probability measure or probability function defined on the sample space should satisfy three axioms of probability theory known as (i) non-negativity, (ii) normalisation and (iii) countable additivity.
- **Conditional probability** of the event E given an event F is defined as follows.

  $$\mathcal{P}(E \mid F) = \frac{\mathcal{P}(E \cap F)}{\mathcal{P}(F)}, \quad \mathcal{P}(F) \neq 0$$
- **Addition law of probability** for n events $E_1, E_2, E_3, \ldots, E_n$ is given by

  $$\mathcal{P}(E_1 \cup E_2 \cup \ldots \cup E_n) = \mathcal{P}(E_1) + \mathcal{P}(E_2) + \ldots + \mathcal{P}(E_n) = \sum_{i=1}^{n} \mathcal{P}(E_i)$$

  when events $E_1, E_2, E_3, \ldots, E_n$ are mutually disjoint

  $$\mathcal{P}(E_1 \cup E_2 \cup \ldots \cup E_n) = \sum_{i=1}^{n} \mathcal{P}(E_i) - \sum_{1\le i<j\le n}^{n} \mathcal{P}(E_i \cap E_j) + \sum_{1\le i<j<k\le n}^{n} \mathcal{P}(E_i \cap E_j \cap E_k) - \ldots + (-1)^{n+1} \mathcal{P}(E_1 \cap E_2 \cap E_3 \cap \ldots \cap E_n)$$

  when sets $E_1, E_2, E_3, \ldots, E_n$ are not mutually disjoint
- **Multiplication law** for **n events** $E_1, E_2, \ldots, E_n$ which occur in succession is given by

  $$\mathcal{P}(E_1 \cap E_2 \cap \ldots \cap E_n) = \mathcal{P}(E_1)\mathcal{P}(E_2 \mid E_1)\mathcal{P}(E_3 \mid E_1E_2)\ldots\mathcal{P}(E_n \mid E_1E_2\ldots E_{n-1})$$

  where $E_1E_2E_3\ldots E_j$ represents $E_1 \cap E_2 \cap \ldots \cap E_j$, etc.
- Two events E and F are said to be **independent** if $\mathcal{P}(E \cap F) = \mathcal{P}(E)\mathcal{P}(F)$.
- In **total law of probability** to obtain probability of an event we use conditioning on n events which partition the sample space.

- **Bayes' theorem** is also known as **inverse probability theorem** because it reverses conditioning event, i.e., it connects $\mathcal{P}(E|F)$ and $\mathcal{P}(F|E)$. Using this theorem, we can also revise probability after getting any additional information.

1.12 TERMINAL QUESTIONS

1. **Monty Hall Problem**: Monty hosted a TV show "**Let's make a deal**". Rules of this show were as follows.

   - There were three doors one of them has a car behind it and the remaining two doors have goats behind them. One car and two goats are placed at random behind three doors but Monty knows where is car and where are the two goats.
   - Contestant has to choose one of the three doors. After the choice of one door by the contestant Monty opens one of the remaining two doors which does not have prize (car) behind it.
   - After this Monty again offers to the contestant would you like to switch your initial choice with the other unopened door or you will be stuck with your initial choice.

   Suppose you get this opportunity will you stick with initial choice or switch? Justify your decision by calculating probabilities of wining in both the cases (i) when you decide to stick and (ii) you decide to switch.

2. **Monty Hall Problem Continue**: Suppose in Monty Hall problem instead of three doors we have 100 doors. Further, suppose that Monty Hall opens 98 doors other than the door you initially pick. Like three doors problem you are now given the opportunity to stick or switch to the single unopened door other than you picked initially. What are the probabilities of winning the prize in both the cases you stick and switch.

3. **Birthday Problem**: Suppose in a party there are n (< 365) persons. Assume that each of them is born in a non-leap year. Also, assume that birth rate is uniform throw-out the 365 days of a year (It means their birthdays are independent, i.e., no twin case or any other kind of dependency). Find the probability of at least one sharing birthday. Also, find value of this probability when n = 1, 2, 3, 4, 22, 23. By observing values of the probabilities at these values of n as n increases comment on the behaviour of the general function giving these probabilities. Comment on the probability at n = 23 which is more than 50%. Give mathematical justification why probability is more than 50% for such a small value 23 of n compare to 365 number of days in a year.

4. **Birthday Problem Continue**: If you are not familiar with making graph in R then it is suggested that solve this problem after completing the course MSTL-011.

   Continue with the previous problem. In the previous problem you noted that as n increases then probability of at least one sharing birthday increases as n increases. Intuitively, this is obvious because as the number of persons in the party will increase then possibility of at least one sharing birthday will also increase. To visualise it make a graph by taking the number of persons on the x-axis and probability of sharing a birthday on the y-axis. Also, mention how many persons are required to achieve at least a probability of 50%, 90% and 99.

5. **Revisit to the Meeting Problem of two Friends Anjali and Prabhat**: Suppose two friends Anjali and Prabhat trying to meet for a date to have lunch say between 1pm to 2pm. Suppose they follow following rules for this meeting:

   - Each of them will reach any time from 1pm to 2pm. Assume that all possibilities of their reaching time from 1pm to 2pm are equally likely for both of them
   - Whoever of them reach first will wait for the other to meet only for 15 minutes. If within 15 minutes the other does not reach, he/she leaves the place and they will not meet.

   Find the probability of their meeting.

6. A pair of two fair tetrahedral dice is thrown. Sum of the two outcomes is noted in each throw. Find the probability that a 4 will be observed before observing a 5 in this random experiment.

1.13 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. Total number of letters in the word ASSISTANT is 9. Also, total number of letters in the word ASSOCIATE is 9. So, total number of ways of selecting one letter from each is $\binom{9}{1}\times\binom{9}{1} = 9\times 9 = 81.$ Let E be the event of getting the same letter from each of the two words ASSISTANT and ASSOCIATE. To count favourable cases for event E let us form the following table.

	Matching letters			
	A	S	I	T
Count in ASSISTANT	2	3	1	2
Count in ASSOCIATE	2	2	1	1
Favourable cases for event E	$\binom{2}{1}\times\binom{2}{1}$	$\binom{3}{1}\times\binom{2}{1}$	$\binom{1}{1}\times\binom{1}{1}$	$\binom{2}{1}\times\binom{1}{1}$
Total favourable cases for event E	$2\times 2+3\times 2+1\times 1+2\times 1=13$			

So, probability of event E using classical approach is given by

$$\mathcal{P}(E) = \frac{\text{No of favourable outcomes for event E}}{\text{Number of exhaustive cases}} = \frac{13}{81}$$

2. Let E and F be the events defined as follows.

   E: the event of getting two kings, and

F: the event of getting two cards having at least one face card.

Here, event F is our reduced sample space. So, required conditional probability is given by

$$\mathcal{P}(E\,|\,F) = \frac{\binom{4}{2}}{\binom{52}{2} - \binom{40}{2}} = \frac{\frac{4!}{2!2!}}{\frac{52!}{2!50!} - \frac{40!}{2!38!}} = \frac{6}{\frac{52\times 51}{2} - \frac{40\times 39}{2}}$$

$$= \frac{12}{2652 - 1560} = \frac{12}{1092} = \frac{1}{91}$$

3. Let R_i be the event that i^{th} drawn ball is red, i = 1, 2, 3 and W_i be the event that i^{th} drawn ball is white, i = 1, 2, 3.

Now, one white and two red balls can be drawn in three mutually exclusive ways mentioned as follows.

$W_1 \cap R_2 \cap R_3$: First drawn ball is white and the other two are red

$R_1 \cap W_2 \cap R_3$: Second drawn ball is white and the other two are red

$R_1 \cap R_2 \cap W_3$: Third drawn ball is white and the other two are red

If we denote the intersection $W_1 \cap R_2 \cap R_3$ by $W_1 R_2 R_3$ then required probability is given by

$$\mathcal{P}(W_1 R_2 R_3 \cup R_1 W_2 R_3 \cup R_1 R_2 W_3) = \mathcal{P}(W_1 R_2 R_3) + \mathcal{P}(R_1 W_2 R_3) + \mathcal{P}(R_1 R_2 W_3)$$

[Using addition law for mutually exclusive events]

$$= \mathcal{P}(W_1)\mathcal{P}(R_2\,|W_1)\mathcal{P}(R_3\,|W_1R_2) + \mathcal{P}(R_1)\mathcal{P}(W_2\,|R_1)\mathcal{P}(R_3\,|R_1W_2) + \mathcal{P}(R_1)\mathcal{P}(R_2\,|R_1)\mathcal{P}(W_3\,|R_1R_2)$$

$$= \frac{\binom{4}{1}}{\binom{11}{1}} \times \frac{\binom{5}{1}}{\binom{10}{1}} \times \frac{\binom{4}{1}}{\binom{9}{1}} + \frac{\binom{5}{1}}{\binom{11}{1}} \times \frac{\binom{4}{1}}{\binom{10}{1}} \times \frac{\binom{4}{1}}{\binom{9}{1}} + \frac{\binom{5}{1}}{\binom{11}{1}} \times \frac{\binom{4}{1}}{\binom{10}{1}} \times \frac{\binom{4}{1}}{\binom{9}{1}}$$

$$= 3 \times \frac{4}{11} \times \frac{5}{10} \times \frac{4}{9} = \frac{8}{33}$$

4. Let $E_1, E_2,$ and E_3 be the events that problem is solved by students A, B and C respectively. Then, we are given

$$\mathcal{P}(E_1) = \frac{1}{2},\;\; \mathcal{P}(E_2) = \frac{1}{3},\;\; \mathcal{P}(E_3) = \frac{1}{4}$$

We know that problem will be solved if at least one of the three students A, B and C solve the problem. But in set theoretical notations at least means union. So, required probability is given by

$$\mathcal{P}(E_1 \cup E_2 \cup E_3) = 1 - \mathcal{P}\left((E_1 \cup E_2 \cup E_3)^c\right) \quad \left[\because \mathcal{P}(E) = 1 - \mathcal{P}(E^c)\right]$$

$$= 1 - \mathcal{P}\left(E_1^c \cap E_2^c \cap E_3^c\right) \quad \left[\text{Using De Morgan's Law}\right]$$

$$= 1 - \mathcal{P}\left(E_1^c\right)\mathcal{P}\left(E_2^c\right)\mathcal{P}\left(E_3^c\right) \quad \left[\begin{array}{l} \because E_i\text{'s are independent} \\ \Rightarrow E_i^c\text{'s are independent} \\ \text{and using (1.48)} \end{array}\right]$$

$$= 1 - \left(1 - \mathcal{P}(E_1)\right)\left(1 - \mathcal{P}(E_2)\right)\left(1 - \mathcal{P}(E_3)\right)$$

$$= 1 - \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{4}\right) = 1 - \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} = 1 - \frac{1}{4} = \frac{3}{4}$$

5. Let us first visualise n consecutive parking spaces in Fig. 1.4.

1	2	3	…	n – 2	n – 1	n

Fig. 1.4: Visualisation of n consecutive parking spaces

Let E_i, $i = 1, 2, 3, \ldots, n$ be the event that Prabhat parks his car at i^{th} parking space. Let A be the event that there is at the most one empty car parking space between their cars. Since each of them park their cars randomly in the n empty parking spaces, so, we have

$$\mathcal{P}(E_1) = \frac{1}{n}, \ \mathcal{P}(E_2) = \frac{1}{n}, \ \mathcal{P}(E_3) = \frac{1}{n}, \ldots, \mathcal{P}(E_n) = \frac{1}{n}$$

$$\mathcal{P}(A \mid E_1) = \frac{\mathcal{P}(A \cap E_1)}{\mathcal{P}(E_1)} = \frac{\frac{2}{n-1}}{\frac{1}{n}} \quad \left[\begin{array}{l} \because \text{If Prabhat parks his cars in the first} \\ \text{parking space then Anjali has only two} \\ \text{options second or third parking spaces} \\ \text{out of available } n-1 \text{ empty spaces} \end{array}\right]$$

Similar situation is when Prabhat parks his car at the n^{th} parking space then Anjali has only two options $(n-2)^{th}$ or $(n-1)^{th}$ parking spaces to park her car, therefore

$$\mathcal{P}(A \mid E_n) = \frac{\mathcal{P}(A \cap E_n)}{\mathcal{P}(E_n)} = \frac{\frac{2}{n-1}}{\frac{1}{n}}$$

If Prabhat parks his car at the second parking space, then in this case Anjali has 3 options 1^{st} or 3^{rd} or 4^{th} parking spaces out of available $n-1$ empty parking spaces to park her car, therefore

$$\mathcal{P}(A \mid E_2) = \frac{\mathcal{P}(A \cap E_2)}{\mathcal{P}(E_2)} = \frac{\frac{3}{n-1}}{\frac{1}{n}}$$

Similar situation is when Prabhat parks his car at the $(n-1)^{th}$ parking space then Anjali has 3 options n^{th} or $(n-3)^{th}$ or $(n-4)^{th}$ parking spaces to park her car, therefore

$$\mathcal{P}(A \mid E_{n-1}) = \frac{\mathcal{P}(A \cap E_{n-1})}{\mathcal{P}(E_{n-1})} = \frac{\frac{3}{n-1}}{\frac{1}{n}}$$

If Prabhat parks his car in any of the parking spaces from 3^{rd} to $(n-2)^{th}$ then in each of these cases Anjali has 4 options to park her car out of the available $n-1$ empty parking spaces, therefore

$$\mathcal{P}(A \mid E_k) = \frac{\mathcal{P}(A \cap E_k)}{\mathcal{P}(E_k)} = \frac{\frac{4}{n-1}}{\frac{1}{n}}, \quad k = 3, 4, 5, \ldots, n-3, n-2$$

Now, using total law of probability required probability is given by

$$\mathcal{P}(A) = \mathcal{P}(E_1)\mathcal{P}(A \mid E_1) + \mathcal{P}(E_2)\mathcal{P}(A \mid E_2) + \ldots + \mathcal{P}(E_n)\mathcal{P}(A \mid E_n)$$

$$= \frac{\frac{2}{n-1}}{\frac{1}{n}} + \frac{\frac{3}{n-1}}{\frac{1}{n}} + \underbrace{\frac{\frac{4}{n-1}}{\frac{1}{n}} + \frac{\frac{4}{n-1}}{\frac{1}{n}} + \ldots + \frac{\frac{4}{n-1}}{\frac{1}{n}}}_{(n-4)\text{ times}} + \frac{\frac{3}{n-1}}{\frac{1}{n}} + \frac{\frac{2}{n-1}}{\frac{1}{n}}$$

$$= \frac{1}{n}\left(\frac{2}{n-1} + \frac{3}{n-1} + \underbrace{\frac{4}{n-1} + \frac{4}{n-1} + \ldots + \frac{4}{n-1}}_{(n-4)\text{ times}} + \frac{3}{n-1} + \frac{2}{n-1}\right)$$

$$= \frac{1}{n}\left(\frac{2 \times 2}{n-1} + \frac{3 \times 2}{n-1} + \frac{4(n-4)}{n-1}\right) = \frac{4 + 6 + 4n - 16}{n(n-1)} = \frac{4n-6}{n(n-1)} = \frac{2(n-3)}{n(n-1)}$$

Terminal Questions

1. First, we solve this problem by considering all possible cases in the following table.

Door number Initially pick by the contestant	Door number behind which prize lies	Decision of the contestant as stick or switch and after that its result as win or lose		Explanation
		Stick	Switch	
1	1	win	lose	Pick d1, prize d1, • win if stick • lose if switch
1	2	lose	win	Pick d1, prize d2, • lose if stick • win if switch
1	3	lose	win	Pick d1, prize d3, • lose if stick • win if switch
2	1	lose	win	Similar explanation works for these cases. Where d1, d2, d3 represent door number 1, 2, 3 respectively
2	2	win	lose	
2	3	lose	win	
3	1	lose	win	
3	2	lose	win	
3	3	win	lose	
Total		3 win and 6 lose	6 win and 3 lose	

Here, we note that total number of possibilities are 9 out of which a contestant wins 3 times if he/she goes with stick option and wins 6 times if goes with switch option. Hence, probability of wining the prize in the case of stick is 3/9 or 1/3 while probability of wining the prize in the case of switch is 6/9 or 2/3. Thus, to increase probability of wining the prize (car) the contestant should switch. In fact, switching strategy makes probability of wining the prize double to the probability of wining in the case of stick. Fig. 1.5 gives a look of the situation when Monty Hall opens a door.

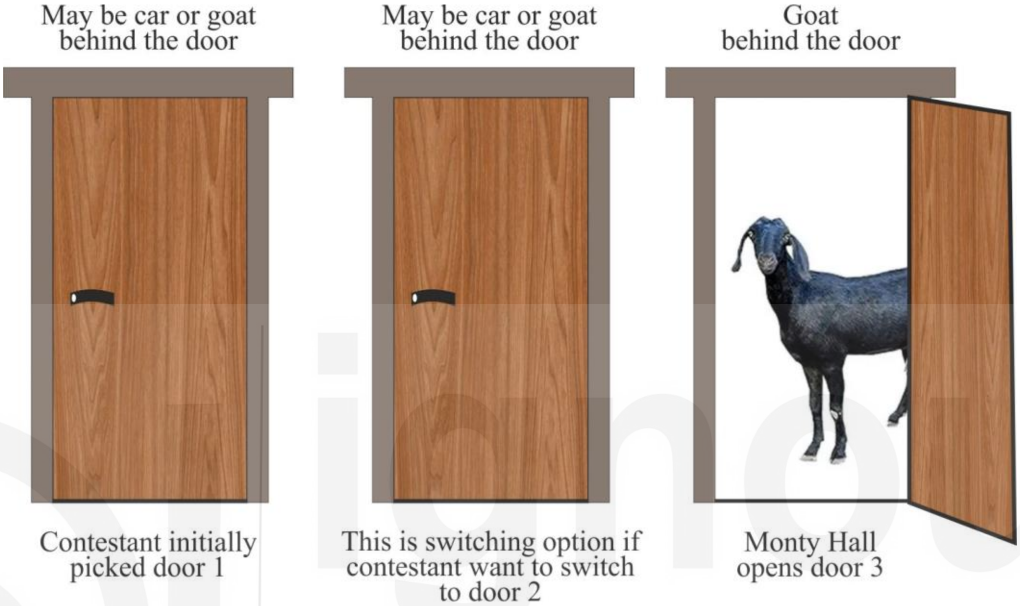

Fig. 1.5: Visualisation of the problem just after the moment Monty Hall opens door 3 and contestant initially picked door 1. So, in this situation prize will be either behind door 1 or door 2

Intuitive solution of this problem: Initially there are three doors and two goats and one car are placed randomly behind the three doors. So, by classical approach to probability theory probability of getting car behind a randomly selected door is 1/3. So, probability of getting a car if you stick with initial choice is 1/3.

We, now proceed to calculate probability in the case you switch. Without loss of generality suppose you initially pick door 1. If you are not familiar with the meaning of 'Without loss of generality' then remember it has simple meaning that our supposition does not affect the genelisability of the result. For example, if contestant pick door 1 or door 2 or door 3 our logic of proving the result will remain the same. So, we have used without loss of generality. After this explanation let us return to the intuitive solution of the problem. After picking door 1 and you are adopting the strategy of switching. It means you will win the prize in both the two cases the car is behind door 2 or door 3. It is so because if car is behind door 2 then Monty Hall will definitely open door 3 as per rules of the game and so you will win on switching. Similarly, if car is behind door 3 then Monty Hall will definitely open door 2 and so again you will win on switching. So, indirectly switching strategy gives you chances of winning the car if it is either behind door 2 or door 3. So, probability of winning car if you go for switching is 2/3. This completes the intuitive solution of the problem.

2. We will not consider all the cases like three doors problem due to large number of doors. So, we will attack the problem intuitively.

   Initially there are 100 doors and 99 goats and one car are placed randomly behind these 100 doors. So, by classical approach to probability theory probability of getting car behind a randomly selected door is 1/100. So, probability of getting a car if you stick with initial choice is 1/100.

   We, now proceed to calculate probability in the case you switch. Without loss of generality suppose you initially pick door 1. After picking door 1 and you are adopting the strategy of switching. It means you will win the prize if the car is behind any of the door 2 to door 100. It is so because if car is behind door 2 then Monty Hall will definitely open all doors 3 to 100 as per rules of the game and so you will win on switching. Similarly, if car is behind door 3 then Monty Hall will definitely open door 2 and all the doors 4 to 100 and so again you will win on switching and so on. So, indirectly switching strategy gives you chances of winning the car if it is either behind the door 2 or door 3 or door 4 and so on up to door 100. So, out of total 100 doors you are going to win in 99 cases. So, probability of winning car if you go for switching is 99/100. This completes the intuitive solution of the problem.

3. Let p(n) denote the required probability that at least two persons among the n persons share their birthday.

$$\therefore\ p(n) = \mathcal{P}\begin{pmatrix}\text{At least two among n persons}\\ \text{share their birthday}\end{pmatrix}$$

$$= 1 - \mathcal{P}\begin{pmatrix}\text{None of the n persons}\\ \text{share their birthday}\end{pmatrix} \quad \begin{bmatrix}\because \text{ at least and none are}\\ \text{complement of each other}\\ \text{and so using (1.19)}\end{bmatrix}$$

$$= 1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \ldots \times \frac{(365-(n-1))}{365} \quad \begin{bmatrix}\because \text{ their birthdays are}\\ \\ \text{independent}\end{bmatrix}$$

$$\Rightarrow p(n) = 1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \ldots \times \frac{(366-n))}{365} \qquad \ldots\ (1.60)$$

Putting n = 1, 2, 3, 4, 22 and 23, we get

$p(1) = 1 - \frac{365}{365} = 1 - 1 = 0$ as expected, because if there is only one person in the party then there is no possibility of a sharing birthday.

$$\left.\begin{aligned}
p(2) &= 1 - \frac{365}{365}\times\frac{364}{365} \simeq 1 - 0.997260 = 0.00274\\
p(3) &= 1 - \frac{365}{365}\times\frac{364}{365}\times\frac{363}{365} \simeq 1 - 0.991796 = 0.008204\\
p(4) &= 1 - \frac{365}{365}\times\frac{364}{365}\times\frac{363}{365}\times\frac{362}{365} \simeq 1 - 0.983644 = 0.016356\\
p(22) &= 1 - \frac{365}{365}\times\frac{364}{365}\times\frac{363}{365}\times\ldots\times\frac{344}{365} \simeq 1 - 0.524305 = 0.475695\\
p(23) &= 1 - \frac{365}{365}\times\frac{364}{365}\times\frac{363}{365}\times\ldots\times\frac{343}{365} \simeq 1 - 0.492703 = 0.507297
\end{aligned}\right\} \quad (1.61)$$

Birthday problem is a famous problem in the world of probability theory. So, we have a built-in function “pbirthday()” in R to calculate such probabilities. Screenshot of R console where we have obtained these probabilities in R is given as follows. You should try it on your system to feel in real time in front of your eyes. … (1.62)

```
> pbirthday(n = 1, classes = 365, coincident = 2)
[1] 0
> pbirthday(n = 2, classes = 365, coincident = 2)
[1] 0.002739726
> pbirthday(n = 3, classes = 365, coincident = 2)
[1] 0.008204166
> pbirthday(n = 4, classes = 365, coincident = 2)
[1] 0.01635591
> pbirthday(n = 22, classes = 365, coincident = 2)
[1] 0.4756953
> pbirthday(n = 23, classes = 365, coincident = 2)
[1] 0.5072972
```

Now, come to point of mathematical justification why p(23) is > 50% for such a small value 23 of n. When n = 23 then let us count the number of possible pairs of sharing a birthday. Number of such pairs is

$\binom{23}{2} = \frac{23!}{2!(23-2)!} = \frac{23\times 22}{2} = 253.$ The number 253 is big compare to

23. So, mathematical justification is as n increases then number of pairs of persons taken two at a time increases and much more than the value of n. This is the reason behind the magic of 23 in this birthday problem.

4. Required graph is shown in Fig. 1.6. Note that graph verify the observation of problem 3. Number of persons that should be present in the party to achieve at least 50%, 90% and 99% probability are 23, 41 and 57 respectively. … (1.63)

5. If the arrival time of Anjali is represented on the horizontal axis and the arrival time of Prabhat is represented on the vertical axis then their possible arrival time points constitute the complete unit square shown in blue colour in Fig. 1.7 (a). Hence, in this case sample space is the entire portion of the unit square shown in blue colour in Fig. 1.7 (a). If E be the event that Anjali and Prabhat meet, then as per the rules of the meeting,

green area shown in Fig. 1.7 (b) indicates favourable area for event E and red area shown in Fig. 1.7 (b) indicates non favourable area for the event E. Hence, probability of the event E is given by

$$\mathcal{P}(\text{E}) = \frac{\text{Area of green portion in Fig. 1.7 (b)}}{\text{Area of unit square}}$$

$$= \frac{1 - \text{Area of two red right angled triangles shown in Fig. 1.7 (b)}}{1}$$

$$= 1 - 2\left(\text{Area of one red right angled triangles shown in Fig. 1.7 (b)}\right)$$

$$= 1 - 2\left(\frac{1}{2} \times \frac{3}{4} \times \frac{3}{4}\right) = 1 - \frac{9}{16} = \frac{7}{16} = 43.75\% \qquad \text{... (1.64)}$$

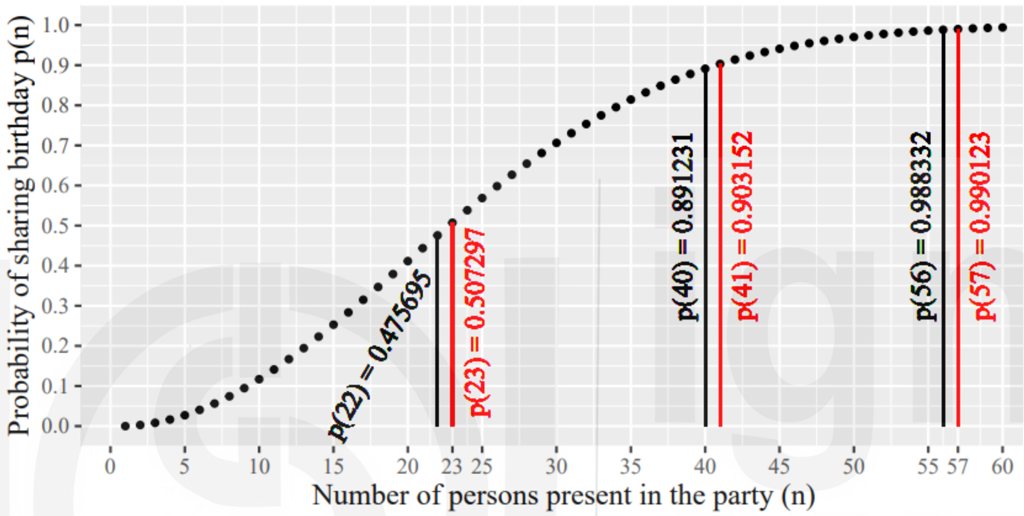

Fig. 1.6: Visualisation of the probability of sharing birthday when n persons are present in a party

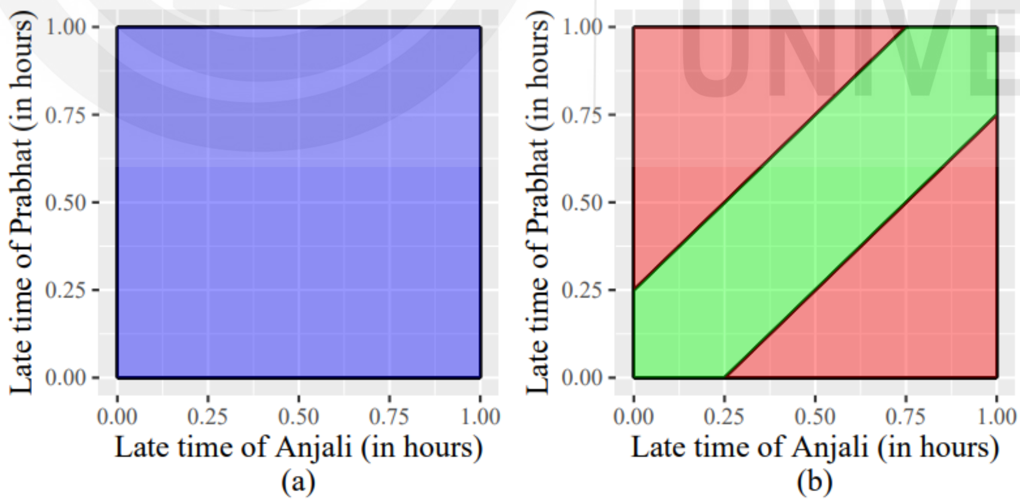

Fig. 1.7: Visualisation of (a) sample space in blue colour (b) meeting portion of the sample in green colour and non-meeting portion in red colour

6. We know that sample space of this random experiment contains 16 equally likely outcomes given as follows.

$$\Omega = \left\{\begin{matrix}(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (2, 4), \\ (3, 1), (3, 2), (3, 3), (3, 4), (4, 1), (4, 2), (4, 3), (4, 4)\end{matrix}\right\}$$

Let us define some events of our interest.

Let E be the event of getting 4 as sum and F be the event of getting 5 as sum.

$\therefore$ $E = \{(1, 3), (3, 1), (2, 2)\}$ and $F = \{(1, 4), (4, 1), (2, 3), (3, 2)\}$ and so

$$\mathcal{P}(E) = \frac{n(E)}{16} = \frac{3}{16} \quad \left[\begin{array}{l}\text{Since all the 16 outcomes are equally likely, so,}\\ \text{using classical approach to probability theory}\end{array}\right]$$

Similarly, $\mathcal{P}(F) = \dfrac{n(F)}{16} = \dfrac{4}{16} = \dfrac{1}{4}$.

Now,

$$\mathcal{P}(\text{getting sum neither 4 nor 5}) = \mathcal{P}\left(\bar{E} \cap \bar{F}\right) = \mathcal{P}\left(\overline{E \cup F}\right) \left[\begin{array}{l}\text{Using De}\\ \text{Morgan's law}\end{array}\right]$$

$$= 1 - \mathcal{P}(E \cup F) \quad \left[\begin{array}{l}\mathcal{P}(A) + \mathcal{P}\left(\bar{A}\right) = 1\\ \Rightarrow \mathcal{P}\left(\bar{A}\right) = 1 - \mathcal{P}(A)\end{array}\right]$$

$$= 1 - \left(\mathcal{P}(E) + \mathcal{P}(F)\right) \quad [\because E \cap F = \phi]$$

$$= 1 - \left(\frac{3}{16} + \frac{4}{16}\right) = 1 - \frac{7}{16} = \frac{9}{16}$$

Let E_k, $k = 1, 2, 3, 4, \ldots$ be the event of getting sum neither 4 nor 5 in the first k – 1 trials and a sum of 4 in k^{th} trial.

$$\therefore \ \mathcal{P}(E_k) = \left(\frac{9}{16}\right)^{k-1} \times \frac{3}{16} \qquad \ldots (1.65)$$

Now, required probability is given by

$$\mathcal{P}\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mathcal{P}(E_k) \quad \left[\begin{array}{l}\text{Using Axiom 3 of probability theory.}\\ \text{You may refer to (1.20)}\end{array}\right]$$

$$= \sum_{k=1}^{\infty} \left(\frac{9}{16}\right)^{k-1} \times \frac{3}{16} \quad [\text{Using (1.65)}]$$

$$= \frac{3}{16} \sum_{k=1}^{\infty} \left(\frac{9}{16}\right)^{k-1} = \frac{3}{16}\left[1 + \left(\frac{9}{16}\right) + \left(\frac{9}{16}\right)^2 + \left(\frac{9}{16}\right)^3 + \ldots\right]$$

$$= \frac{3}{16}\left[\frac{1}{1 - 9/16}\right] \quad \left[\begin{array}{l}\because \text{sum of infinite GP } a + ar + ar^2 + \ldots\\ = \dfrac{a}{1-r}, \text{ provided } |r| < 1\end{array}\right]$$

$$= \frac{3}{7}$$