## 3. Probability Theory

## Introduction

Probability is a concept that most people understand naturally, since words such as "chance", "likelihood", "possibility" and "proportion" (and indeed probability itself) are used as part of everyday speech. For example, most of the following, which might be heard in any business situation, are in fact statements of probability:
i. There is a $30 \%$ chance that this job will not be finished in time.
ii. There is every likelihood that the business will make a profit next year.
iii. Nine times out of ten he arrives late for his appointments.
iv. There is no possibility of delivering the goods before Tuesday.

All the above are expressions indicating a degree of uncertainty. A very important branch of mathematics called the theory of probability provides a numerical measure of uncertainty. The probability describes certainty by 1 , impossibility by 0 and the various grades of uncertainties by fractions or decimals in between 0 and 1 .

## Definitions of terms used in probability theory

The terms, which are used in probability theory under different situations, are discussed below.

## Random Experiment

An experiment is, any repeatable process from which an outcome, measurement or result is obtained. When the outcomes cannot be predicted with certainty, then the experiment is a random experiment.

## Examples:

i. Throwing a die.
ii. Tossing of a coin.
iii. Inspection of an item to determine whether it is defective or non defective.

## Sample space and sample points

The set of all possible outcomes of a random experiment is called the sample space. It is usually denoted by S . The elements of sample space are called sample points.

## Examples:

In throwing of two dice the sample space is,
$\mathrm{S}=\{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5)$,
$(2,6),(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,1),(4,2),(4,3),(4,4),(4,5)$,
$(4,6),(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),(6,1),(6,2),(6,3),(6,4)$,
$(6,5),(6,6)\}$
This experiment has 36 possible outcomes. Therefore it has 36 sample points.

## Events

An Event is any collection of outcomes of an experiment: that is an event is any subset of the sample space.

## Example:

Consider throwing of two dice and $\mathrm{A}, \mathrm{B}, \mathrm{C}$ are the three events defined as follows.
$\mathrm{A}=$ the sum of the numbers shown by two dice is odd.
$B=$ the sum of the numbers of the two dice is 7 .
$\mathrm{C}=$ Two dice has the same number.

$$
\text { Then } \begin{aligned}
& A=\{(1,2),(1,4),(1,6),(2,1) \ldots(6,5)\} \\
& B=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \\
& C=\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}
\end{aligned}
$$

## Rules of Event Operations

## (a).Intersection of Events

Intersection of two events A and B results in another event containing all such elements that are common to A and B . It is denoted by,


## Example:

$S=\{1,2,3,4,5,6\}$
$A=\{x / x$ is odd numbers $\} \quad B=\{x / x$ is numbers less than 4$\}$
$A=\{1,3,5\}$,
$B=\{1,2,3\}$
$\therefore \quad \mathrm{A} \cap B=\{1,3\}$

## (b).Union of events.

Union of two events $\mathrm{A} \& \mathrm{~B}$ results in another event containing all elements, which belong to A or to B or to both. It is denoted by, $A \cup B=\{x / x \in A$ or $x \in B\}$


## Example:

Consider the example given under intersection of two events.

$$
\therefore A \cup B=\{1,2,3,5\}
$$

## (c). Difference of two events

Difference of an event A from another event B denoted by A B is defined as the event consisting of those elements of $A$ which does not belong to $B$.


Venn diagram for difference of events

## Example:

Consider the example given under intersection of two events.

$$
A \backslash B=\{5\}
$$

## (d).Compliment of an event

Compliment of an event A denoted by $A^{\prime}$ is the event containing all the elements of S that are not elements of A .

$$
A^{\prime}=\{x / x \in S \text { and } x \notin A\}
$$

## Example:

Consider the example discussed under the intersection of two events. Find $A^{\prime}$ ?

$$
\therefore A^{\prime}=\{2,4,6\}
$$

## Collectively Exhaustive Events

If $A_{1}, A_{2}, \ldots A_{n}$ are events of sample space $S$ and if
$A_{1} \cup A_{2} \cup \ldots . . \cup A_{n}=S$ then $A_{1}, A_{2}, \ldots, A_{n}$ are said to be collectively exhaustive events.

## Example:

$S=\{1,2,3,4,5,6\} \quad A=\{1,3,5\} \quad B=\{2,4,6\} \quad C=\{1,2,6\}$
$A \cup B=S \therefore$ A and B are collectively exhaustive events.
$A \cup C=\{1,2,3,5,6\} \neq \mathrm{S} \quad \therefore \mathrm{A}$ and C are not collectively exhaustive events.

## Mutually Exclusive Events

Events are said to be mutually exclusive or disjoint if two or more of them cannot occur simultaneously in a single trial of an experiment. If $A$ and $B$ are mutually exclusive events then $A \cap B=\phi$

## Examples:

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

A and B are not mutually exclusive. Because $\mathrm{A} \cap \mathrm{B} \neq \phi$.
A and C are mutually exclusive. Because $\mathrm{A} \cap \mathrm{C}=\phi$.

## Equally likely Events

The outcomes of a random experiment are said to be equally likely if after taking into consideration all the relevant evidences there is no reason to expect one in preference to the other.

## Example:

In throwing an unbiased die, all the six faces are equally likely to come.

## Counting Techniques

Counting techniques are used to determine the number of possible outcomes of a particular experiment or the number of elements in a particular set without direct enumeration. Here we discussed three counting techniques.

1. Fundamental principle of counting.
2. Permutations.
3. Combinations.

## Fundamental principle of counting

If same procedure can be performed in $\mathrm{n}_{1}$ different ways, and if, following this procedure, a second procedure can be performed in $\mathrm{n}_{2}$ different ways, and if, following this second procedure, a third procedure can be performed in $n_{3}$ different ways, and so forth; then the number of ways the procedures can be performed in the order is the product n1.n2.n3.....

## Example:

Suppose a car number plate contains two distinct letters followed by three digits with the first digit not zero. How many different number plates can be printed?

The first letter can be printed in 26 different ways, the second letter in 25 different ways (since the letter printed first cannot be chosen for the second letter), the first digit in 9 ways and each of other two digits in 10 ways. Hence,

Number of different plates $=26 \cdot 25 \cdot 9 \cdot 10 \cdot 10=585000$

## Factorial Notation

The product of the positive integers from 1 to $n$ inclusive occurs very often in mathematics and hence denoted by the special symbol $n$ ! (read " $n$ " factorial)

$$
1.2 .3 \ldots n=n!\quad \text { i.e. } n!=n .(n-1) \cdot(n-2)
$$

It also convenient to define $0!=1$

## Examples:

1. $2!=2 \times 1=2$
2. $3!=6$
3. $4!=24$
4. $5!=5 \times 4!=5 \times 24=120$
5. $6!=720$
6. $\frac{8!}{6!}=56$

## Permutations and Combinations

## Introduction

Each of the arrangement, which can be made by taking some or all of a number of things, is called a permutation. Each of the groups or selections, which can be made by taking some or all of a number of things, is called a combination.

## Permutations

An arrangement of a set of $n$ objects in a given order is called a permutation of the objects (taken all at a time). An arrangement of any $r \leq n$ of these objects in a given order is called an $r$ - permutation or a permutation of the $n$ objects taken $r$ at a time.

## Example:

Consider the set of letters $a, b$, and $c$. Then,
(i) abc, acb, bac, bca, cab, cba are the all permutations of the 3 letters (taken all at a time)
(ii) $\mathrm{ab}, \mathrm{ac}, \mathrm{ba}, \mathrm{bc}, \mathrm{ca}, \mathrm{cb}$ are the all permutations of the 3 letters taken 2 at time.

The number of permutations of n objects taken r at a time will be denoted by ${ }^{n} p_{r}$ and it can be obtained using the formula, $\quad{ }^{n} p_{r}=\frac{n!}{(n-r)!}$

## Examples:

Let's say we have 8 contestants, how many ways can we award a 1st, 2 nd and 3rd place prize among eight contestants? (Gold / Silver / Bronze)

$$
{ }^{8} P_{3}=8 * 7 * 6
$$

## Combinations

As mentioned earlier, a combination is a selection of objects in any order. The number of combinations of $n$ objects taken $r$ at a time will be denoted by ${ }^{n} C_{r}$.

$$
{ }^{n} C_{r}=\frac{n!}{r!(n-r)!}
$$

## Examples:

1 How many committees of 3 persons can be selected from a group of 9 ?
2 In how many ways can form a committee of 4 men and 2 women be selected from 10 men and 8 women?

$$
{ }^{9} C_{3}=9 * 8 * 7 /(3 * 2 * 1) \quad{ }^{10} C_{4} *{ }^{8} C_{2}=\frac{10 * 9 * 8 * 7}{4 * 3 * 2 * 1} * \frac{8 * 7}{2 * 1}
$$

## Summary of Counting Procedures

Number of ways of permuting (ordering) a set of n distinct objects $=\mathrm{n}$ !

| n- number of distinct items to <br> choose from <br> k- number of items chosen | Ordered | Un-ordered |
| :--- | :--- | :--- |
| Without <br> replacement | $n^{n} P_{k}$ | ${ }^{n} C_{k}$ |


| With <br> replacement | $n^{k}$ | ${ }^{n+k-1} C_{k}={ }^{n+k-1} C_{n-1}$ |
| :--- | :--- | :--- |

## Probability Definitions of Probability

In this section we shall discuss four approaches to the definition and interpretation of probability: Classical approach, relative frequency approach, subjective approach and axiomatic approach.

## Classical approach

The classical approach of defining probability is the oldest and simplest. It originated in eighteenth century in problems pertaining to game of chance, such as throwing of coins, dice etc.

## Definition

If a random experiment result in $n$ mutually exclusive and equally likely outcomes and out of which $m$ one favourable to the occurrence of event A , then the probability of occurrence of event A , usually denoted by $\mathrm{P}(\mathrm{A})$ is given by

$$
\mathrm{P}(\mathrm{~A})=\frac{\text { Number of favourable cases }}{\text { Total number of equally likelycases }}=\frac{n(A)}{n(S)}=\frac{m}{n}
$$

The classical approach is based an abstract reasoning. Classical probability is often called a prior probability because if we keep, using examples of unbiased dice, fair coin etc. we can state the answer in advance without rolling a die or tossing a coin etc.

## Examples:

1. A coin is tossed twice. What is the probability of getting?
a) Two heads?
b) Two tails?
c) A head and a tail?
$1 / 2 * 1 / 2$
$1 / 2^{*} 1 / 2$
$1 / 2 * 1 / 2+1 / 2 * 1 / 2$
2. Consider throwing of two dice and find the probability of the following events.
a. A = The sum of the numbers shown by two dice is odd
b. $\quad \mathrm{B}=$ The sum of the numbers of the two dice is 7
c. $\mathrm{C}=$ two dice has same number

A is set with two pairs with one odd and one even. $n(A)=2 * 3 \mathrm{C} 1 * 3 \mathrm{C} 1=18$
$\mathrm{P}(\mathrm{A})=18 / 36$
$n(B)=6 \therefore P(B)=6 / 36 \quad n(C)=6 \therefore P(C)=6 / 36$
3. A bag contains 3 red, 6 white and 7 blue balls. What is the probability that two balls drawn are white and blue?
Without replacement, 6/16*7/15*2
The basic assumption underling the classical approach is that the outcomes are equally likely. This requirement restricts the application of probability only to such experiment, which gives rise to symmetrical outcomes. Therefore, this definition is different or impossible to apply as soon as we divert from the fields of coins, dice, cards, and other simple games of chance.

## The Relative Frequency Approach

Suppose that an event A occurs $m$ times in $n$ repetition of a random experiment. Then the ratio $\frac{m}{n}$ gives the relative frequency of the event A . The value which is approached by $\frac{m}{n}$ when $n$ becomes infinity defined as the probability of occurrence of event A. i.e. $\mathrm{P}(\mathrm{A})==_{\mathrm{n} \rightarrow \infty}^{\operatorname{Lim}} \frac{m}{n}$
Theoretically, we can obtain the probability of an event as given by the above limit. However, in practice we can only try to have a close estimate of $\mathrm{P}(\mathrm{A})$ based on large number of observations. For practical convenience, the estimate of $\mathrm{P}(\mathrm{A})$ can be written as if it were actually $\mathrm{P}(\mathrm{A})$ and the relative frequency definition of probability may be expressed as: $P(A)=\frac{m}{n}$. The probability obtained by above relative frequency definition is called a posterior or empirical probability.

## Examples:

1. If, out of 60 orders received so far this year, 12 were not completely satisfied; the proportion $12 / 60$ or 0.2 is the (empirical) probability that the next order received will not be completely satisfied
2. A number of families of a particular type were measured by the number of children they contain to give the following frequency distribution

| No. of children | 0 | 1 | 2 | 3 | 4 | 5 or more |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| No. of families | 12 | 28 | 22 | 8 | 2 | 2 |

Use this information to calculate the probability that another family of this type of will have,
a) 2
b) 3 or more
c) Less than 2 children.
a) $22 / 74$
b) $8 / 74+2 / 74+2 / 74$
c) $12 / 74+28 / 74$

## Subjective approach

The Subjective probability is defined, as the probability assigned to an event by an individual based on whatever evidence is available. Hence such probabilities are based on the knowledge, past experience, beliefs of the person making the probability statements.

Subjective probability assignments are frequently found when events occur only once or at most a very few times. Say that it is your job to interview and select a new social services caseworker. You have narrowed your choice to three people. Each has an attractive appearance, a high level of energy, abounding self-confidence, a record of past accomplishments and a state of mind that seems to welcome challenges. What are the chances each will relate to clients successfully? Answering this question and choosing among the three will require you to assign a subjective probability to each person's potential.

Main shortcoming of this approach is that two reasonable people faced with the same evidence could easily come up with quite different subjective probabilities for the same event.

## Axiomatic Approach to Probability

The Russian mathematician A.N.Kolmagerov introduced the axiomatic approach to probability in the year 1933. He introduced probability as a set function. When this approach is followed, no precise definition of probability is given, rather it gives certain axioms on which probability calculations are based.

## Definition

Given a sample space of a random experiment, the probability of the occurrence of any event A is defined as a probability function $\mathrm{P}(\mathrm{A})$ satisfying the following axioms.
i) The probability of an event exists is real and non negative, $\mathrm{P}(\mathrm{A}) \geq 0$
ii) The probability of the entire sample space is one, $\mathrm{P}(\mathrm{S})=1$
iii) If $\mathrm{A}_{1}, \mathrm{~A}_{2}, \ldots \mathrm{~A}_{3}$ are mutually exclusive events, then, $\mathrm{P}\left(\mathrm{A}_{1} \cup \mathrm{~A}_{2} \cup \ldots \mathrm{U} \mathrm{A}_{\mathrm{n}}\right)=\mathrm{P}\left(\mathrm{A}_{1}\right)+\mathrm{P}\left(\mathrm{A}_{2}\right)+\ldots+\mathrm{P}\left(\mathrm{A}_{\mathrm{n}}\right)$
That is, $\mathrm{P}\left(\cup_{i=1}^{n} \mathrm{~A}_{\mathrm{i}}\right)=\sum_{i=1}^{P}\left(A_{i}\right)$
Using the axioms of probability, it's possible to derive many theorems, which play an important role in applications.

Theorem (1) If $\phi$ is the empty set, then $\mathrm{P}(\phi)=0$
Theorem (2) If $A^{\prime}$ is the complement of an event $A^{\prime}$ then $\mathrm{P}\left(A^{\prime}\right)=1-\mathrm{P}(\mathrm{A})$
Theorem (3) If A and B are any two events then,

$$
P(A \backslash B)=P(A)-P(A \cap B) \quad \text { Or } P\left(A \cap B^{\prime}\right)=P(A)-P(A \cap B)
$$

Theorem (4) If A and B are any two events, then,
$P(A \cup B)=P(A)+P(B)-P(A \cap B)$
This theorem is called addition low of probability.

## Examples:.

The results of an examination conducted in two parts I and II for 50 candidates were recorded as follows. 20 passed in part I; and 15 pass in part II; and 18 failed in both part I and Part II. If out of these candidates one is selected at random, find the probability that the candidate:
(i) Passed in both part I and II.
(ii) Failed only in part I.
(iii) Failed in only one of the parts.

## Conditional Probability

Let A be any event, where $\mathrm{P}(\mathrm{A})>0$. The probability that an event B occurs subject to the condition that A has already occurred is known as the conditional Probability of occurrence of the event B on the assumption that event A has already occurred and is denoted,

$$
\mathrm{P}(\mathrm{~B} / \mathrm{A})=\frac{P(A \cap B)}{P(A)}
$$

Example:
In a group of 200 laptop buyers, 80 bought antivirus software, 60 purchased screen guards, and 40 purchased an antivirus software and a screen guard. If a laptop buyer bought an antivirus software, what is the probability they also bought a screen guard?
$\mathrm{P}(\mathrm{A})=0.4$ and $\mathrm{P}(\mathrm{A} \cap \mathrm{B})=0.2$

$$
\mathrm{P}(\mathrm{~B} \mid \mathrm{A})=\mathrm{P}(\mathrm{~A} \cap \mathrm{~B}) / \mathrm{P}(\mathrm{~A})=0.2 / 0.4=0.5 .
$$

Example:

|  | Have pets | Don't have pets | Total |
| :--- | :--- | :--- | :--- |
| Male | 41 | 8 | 49 |
| Female | 45 | 6 | 51 |
| Total | 86 | 14 | 100 |

1. What is the probability that a randomly selected person is male, given that they own a pet?
2. What is the probability that a randomly selected Female doesn't have a pet?
$M=$ event that a male is selected and $P=$ event that a pet owner is selected
3. $\mathrm{P}(\mathrm{M} \mid \mathrm{P})=\mathrm{P}(\mathrm{M} \cap \mathrm{P}) / \mathrm{P}(\mathrm{P})=0.41 / 0.86=0.477$
4. $\mathrm{P}\left(\mathrm{P}^{\prime} \mid \mathrm{M}^{\prime}\right)=\mathrm{P}\left(\mathrm{M}^{\prime} \cap \mathrm{P}^{\prime}\right) / \mathrm{P}\left(\mathrm{M}^{\prime}\right)=0.6 / 0.51=0.118$

## Multiplication Rule of Probability

If we cross multiply the above equations we can obtain the following results.
i. $\quad \mathrm{P}(\mathrm{B} / \mathrm{A})=\frac{P(A \cap B)}{P(A)} \quad P(A \cap B)=P(A) \cdot P(B / A)$
ii. $\quad \mathrm{P}(\mathrm{A} / \mathrm{B})=\frac{P(A \cap B)}{P(B)} \quad P(A \cap B)=P(B) \cdot P(A / B)$

This is called multiplication rule of probability. That is,

$$
\mathrm{P}(\mathrm{~A} \cap B)=\mathrm{P}(\mathrm{~A}) \cdot \mathrm{P}(\mathrm{~B} / \mathrm{A}) \text { and } \mathrm{P}(\mathrm{~A} \cap \mathrm{~B})=\mathrm{P}(\mathrm{~B}) \cdot \mathrm{P}(\mathrm{~A} / \mathrm{B})
$$

Example:
At a job interview, 2 people are randomly selected from a group of 7 tall and 3 short applicants. Tall and short applicants are selected as security officers and laborers respectively. Find the probability that both randomly selected applicants are laborers.
$A=$ event that first person is a laborer and $B=$ event that second person is a laborer.
$\mathrm{P}(\mathrm{A})=3 / 10 \mathrm{P}(\mathrm{B} \mid \mathrm{A})=2 / 9$
$P(A \cap B)=3 / 10 * 2 / 9=0.067$

## Example:

A bag has 4 white cards and 5 blue cards. We draw two cards from the bag one by one with replacement. Find the probability of getting two different colored cards.
$A=$ event that first card is white and $B=$ event that second card is white.
$P(A)=4 / 9 P(B \mid A)=4 / 9$
$P\left(A^{\prime}\right)=5 / 9 P\left(B^{\prime} \mid A^{\prime}\right)=5 / 9$
$P(A \cap B)=4 / 9 * 4 / 9=0.197 \quad P(A \cap B)=5 / 9 * 5 / 9=0.308$
Ans $=1-0.197-0.308=0.495$

## Independence of two events

When events A and B have no influence on one another, we say that events A and B are independent. If the events A and B are independent, the probability of occurrence of any one of them does not depend upon that of the other.

If A and B are independent then, $\quad \mathrm{P}(\mathrm{B} / \mathrm{A})=\mathrm{P}(\mathrm{B}) \quad$ and $\quad \mathrm{P}(\mathrm{A} / \mathrm{B})=\mathrm{P}(\mathrm{A})$ Now substituting $\mathrm{P}(\mathrm{B})$ for $\mathrm{P}(\mathrm{B} / \mathrm{A})$ in the multiplication rule of probability,

$$
\mathrm{P}(\mathrm{~A} \cap B)=\mathrm{P}(\mathrm{~A}) \cdot \mathrm{P}(\mathrm{~B} / \mathrm{A})
$$

Therefore, events A and B are independent if $\mathrm{P}(\mathrm{A} \cap B)=\mathrm{P}(\mathrm{A}) \cdot \mathrm{P}(\mathrm{B})$; otherwise they are dependent.

## Examples:

1. In an examination $30 \%$ of the students are failed in Mathematics, $20 \%$ of the students are failed in Economics and $12 \%$ are failed in both Mathematics \& Economics. A student is selected at random.
i. What is the probability that the student has failed in Mathematics, if it is known that he has failed Economics?
A- The event of randomly selected student is failed in Mathematics
B- The event of randomly selected student is failed in Economics

$$
\begin{array}{ll}
\mathrm{P}(\mathrm{~A})=0.3 & \mathrm{P}(\mathrm{~B})=0.2 \\
\mathrm{P}(\mathrm{~A} \mid \mathrm{B})=\mathrm{P}(\mathrm{~A} \cap \mathrm{~B}) / \mathrm{P}(\mathrm{~B})
\end{array}
$$

ii. What is the probability he has failed only in Mathematics?

$$
\mathrm{P}(\mathrm{~A} \backslash \mathrm{~B})=\mathrm{P}(\mathrm{~A})-\mathrm{P}(\mathrm{~A} \cap \mathrm{~B})
$$

iii. What is the probability he has failed in one of the subjects?

$$
\begin{aligned}
& \mathrm{P}((\mathrm{~A} U \mathrm{~B}) \backslash(\mathrm{A} \cap \mathrm{~B}))=\mathrm{P}(\mathrm{~A} U \mathrm{~B})-\mathrm{P}(\mathrm{~A} \cap \mathrm{~B})=\mathrm{P}(\mathrm{~A})+\mathrm{P}(\mathrm{~B})-2 \mathrm{P}(\mathrm{~A} \cap \mathrm{~B}) \\
& \text { Or } \mathrm{P}((\mathrm{~A} \backslash \mathrm{~B}) \mathrm{U}(\mathrm{~B} \backslash \mathrm{~A}))=\mathrm{P}(\mathrm{~A} \backslash \mathrm{~B})+\mathrm{P}(\mathrm{~B} \backslash \mathrm{~A})
\end{aligned}
$$

2. An article made up of three parts A, B \& C manufactured by a company. The probabilities of those components being defective are respectively $0.03,0.02$ and 0.05 . What is the probability that assemble article will be defective?
$\mathrm{A}, \mathrm{B}$ and C be the events of three parts $\mathrm{A}, \mathrm{B}$ and C are being defective respectively.
W be the event of the article is not defective. i.e. $\mathrm{W}=\mathrm{A}^{\prime} \cap \mathrm{B}^{\prime} \cap \mathrm{C}^{\prime}$
We are interested in finding $\mathrm{P}\left(\mathrm{W}^{\prime}\right)$
$\mathrm{P}\left(\mathrm{W}^{\prime}\right)=1-\mathrm{P}\left(\mathrm{A}^{\prime} \cap \mathrm{B}^{\prime} \cap \mathrm{C}^{\prime}\right)=1-\mathrm{P}\left(\mathrm{A}^{\prime}\right) \mathrm{P}\left(\mathrm{B}^{\prime}\right) \mathrm{P}\left(\mathrm{C}^{\prime}\right)$

## Total Probability Law

If $A_{1}, A_{2}, A_{3}, \ldots A_{n}$ are mutually exclusive and exhaustive events, where $A_{i}>0$ for all $i=1,2, \ldots, \mathrm{n}$ and if B is any random event

Then

$$
\mathrm{P}(\mathrm{~B})=\sum_{i=1}^{n} P\left(A_{i}\right) \cdot P\left(B / A_{i}\right)
$$

Proof


Diagram for proof of total Probability Law

By the multiplication rule
$P\left(A_{i} \cap B\right)=P\left(A_{i}\right) \cdot P\left(B / A_{i}\right) ; \mathrm{i}=1,2, \ldots . n$.

$$
P(B)=\mathrm{P}\left(\mathrm{~A}_{1}\right) \cdot \mathrm{P}\left(\mathrm{~B} / \mathrm{A}_{1}\right)+\mathrm{P}\left(\mathrm{~A}_{2}\right) \cdot \mathrm{P}\left(\mathrm{~B} / \mathrm{A}_{2}\right)+\mathrm{P}\left(\mathrm{~A}_{3}\right) \cdot \mathrm{P}\left(\mathrm{~B} / \mathrm{A}_{3}\right)+\ldots \ldots \ldots .+\mathrm{P}\left(\mathrm{~A}_{\mathrm{n}}\right) \cdot \mathrm{P}\left(\mathrm{~B} / \mathrm{A}_{\mathrm{n}}\right)
$$

$$
P(B)=\sum_{\mathrm{i}=1}^{n} P\left(A_{i}\right) \cdot P\left(B / A_{i}\right)
$$

## Bayes' Theorem

Conditional Probability takes into accounts information about the occurrence of one event to predict the probability of another event. This concept can be extended to revise probabilities based on new information and to determine the probability that a particular effect was due to a specific cause. The procedure for revising these probabilities is known as "Bayes' Theorem".

If $A_{1}, A_{2}, \ldots \ldots ., A_{n}$ are mutually exclusive events whose union is the sample space $S$, where $P(A) \neq 0$ for all $\mathrm{i}=1, \ldots . \mathrm{n}$, and if B is any random event (for which $\mathrm{P}(\mathrm{B}) \neq 0$ ), thus for all i ,

$$
\begin{aligned}
& P\left(A_{i} / B\right)=\frac{P\left(B / A_{i}\right) P\left(A_{i}\right)}{P(B)} \\
& P\left(A_{i} / B\right)=\frac{P\left(B / A_{i}\right) P\left(A_{i}\right)}{P\left(A_{1}\right) P\left(B / A_{1}\right)+P\left(B / A_{2}\right) P\left(A_{2}\right)+\ldots .+P(A / B) P\left(A_{n}\right)}
\end{aligned}
$$

## Example:

Two groups of candidates are competing for the position of the board of Directors of a company. The probabilities that the first and the second groups will win are 0.6 and 0.4 respectively. If the first group wins, the probability of introducing a new product is 0.4 and corresponding probability if the second group wins is 0.75 .
(i) What is the probability that the new product will be introduced?

A1-event of first group wins A2-event of second group wins
$\mathrm{P}(\mathrm{A} 1)=0.6 \quad \mathrm{P}(\mathrm{A} 2)=0.4$
B -event of introducing a new product

$$
\mathrm{P}(\mathrm{~B} \mid \mathrm{A} 1)=0.4 \quad \mathrm{P}(\mathrm{~B} \mid \mathrm{A} 2)=0.75
$$

$$
\therefore \mathrm{P}(\mathrm{~B})=\mathrm{P}(\mathrm{~B} \cap \mathrm{~A} 1)+\mathrm{P}(\mathrm{~B} \cap \mathrm{~A} 2)=\mathrm{P}(\mathrm{~A} 1) \mathrm{P}(\mathrm{~B} \mid \mathrm{A} 1)+\mathrm{P}(\mathrm{~A} 2) \mathrm{P}(\mathrm{~B} \mid \mathrm{A} 2)
$$

(ii) If the product was introduced what is the probability that the first group won?

$$
\begin{aligned}
\mathrm{P}(\mathrm{~A} 1 \mid \mathrm{B}) & =\mathrm{P}(\mathrm{~A} 1 \mathrm{~B}) / \mathrm{P}(\mathrm{~B}) \\
& =\mathrm{P}(\mathrm{~A} 1) \mathrm{P}(\mathrm{~B} \mid \mathrm{A} 1) / \mathrm{P}(\mathrm{~A} 1) \mathrm{P}(\mathrm{~B} \mid \mathrm{A} 1)+\mathrm{P}(\mathrm{~A} 2) \mathrm{P}(\mathrm{~B} \mid \mathrm{A} 2)
\end{aligned}
$$

Example At a certain university, $4 \%$ of men are over 6 feet tall and $1 \%$ of women are over 6 feet tall. The total student population is divided in the ratio $3: 2$ in favour of women. If a student is selected at random from among all those over six feet tall, what is the probability that the student is a woman?

## Solution

Let $M=\{$ Student is Male $\}, F=\{$ Student is Female $\}$, (note that $M$ and $F$ partition the sample space of students), $T=\{$ Student is over 6 feet tall $\}$. We know that $P(M)=2 / 5, P(F)=3 / 5$, $P(T \mid M)=4 / 100$ and $P(T \mid F)=1 / 100$. We require $P(F \mid T)$. Using Bayes' Theorem we have:

$$
\begin{aligned}
P(F \mid T) & =\frac{P(T \mid F) P(F)}{P(T \mid F) P(F)+P(T \mid M) P(M)} \\
& =\frac{\frac{1}{100} \times \frac{3}{5}}{\frac{1}{100} \times \frac{3}{5}+\frac{4}{100} \times \frac{2}{5}} \\
=\frac{3}{11} &
\end{aligned}
$$

Example A factory production line is manufacturing bolts using three machines, $A, B$ and $C$. Of the total output, machine $A$ is responsible for $25 \%$, machine $B$ for $35 \%$ and machine $C$ for the rest. It is known from previous experience with the machines that $5 \%$ of the output from machine $A$ is defective, $4 \%$ from machine $B$ and $2 \%$ from machine $C$. A bolt is chosen at random from the production line and found to be defective. What is the probability that it came from
(a) machine $A$
(b) machine $B$
(c) machine $C$ ?

## Solution

Let $D=\{$ bolt is defective $\}, A=\{$ bolt is from machine $A\}, B=\{$ bolt is from machine $B\}, C=\{$ bolt is from machine $C\}$. We know that $P(A)=0.25, P(B)=0.35$ and $P(C)=0.4$. Also $P(D \mid A)=0.05, P(D \mid B)=0.04, P(D \mid C)=0.02$. A statement of Bayes' Theorem for three events $A, B$ and $C$ is

$$
\begin{aligned}
P(A \mid D) & =\frac{P(D \mid A) P(A)}{P(D \mid A) P(A)+P(D \mid B) P(B)+P(D \mid C) P(C)} \\
& =\frac{0.05 \times 0.25}{0.05 \times 0.25+0.04 \times 0.35+0.02 \times 0.4} \\
& =0.362
\end{aligned}
$$

Similarly

$$
\begin{aligned}
P(B \mid D) & =\frac{0.04 \times 0.35}{0.05 \times 0.25+0.04 \times 0.35+0.02 \times 0.4} \\
& =0.406 \\
P(C \mid D) & =\frac{0.02 \times 0.4}{0.05 \times 0.25+0.04 \times 0.35+0.02 \times 0.4} \\
& =0.232
\end{aligned}
$$

An engineering company advertises a job in three papers, $A, B$ and $C$. It is known that these papers attract undergraduate engineering readerships in the proportions 2:3:1. The probabilities that an engineering undergraduate sees and replies to the job advertisement in these papers are $0.002,0.001$ and 0.005 respectively. Assume that the undergraduate sees only one job advertisement.
(a) If the engineering company receives only one reply to it advertisements, calculate the probability that the applicant has seen the job advertised in place $A$.
i. in $A$
ii. in $B$
iii. in $C$
(b) If the company receives two replies, what is the probability that both applicants saw the job advertised in paper $A$ ?

Let $A=\{$ Person is a reader of paper $A\}, B=\{$ Person is a reader of paper $B\}, C=\{$ Person is a reader of paper $C\}$ and let $R=\{$ Reader applies for the job $\}$. We have the probabilities
(a) $\quad P(A)=1 / 3 \quad P(R \mid A)=0.002 \quad P(B)=1 / 2 \quad P(R \mid B)=0.001$

$$
P(C)=1 / 6 \quad P(R \mid C)=0.005
$$

$$
\begin{aligned}
P(A \mid R) & =\frac{P(R \mid A) P(A)}{P(R \mid A) P(A)+P(R \mid B) P(B)+P(R \mid C) P(C)} \\
& =\frac{1}{3}
\end{aligned}
$$

Similarly

$$
P(B \mid R)=\frac{1}{4} \quad \text { and } \quad P(C \mid R)=\frac{5}{12}
$$

(b) Now, assuming that the replies and readerships are independent

$$
\begin{aligned}
\mathrm{P}(\text { Both applicants read paper } A) & =P(A \mid R) \times P(A \mid R) \\
& =\frac{1}{3} \times \frac{1}{3} \\
& =\frac{1}{9}
\end{aligned}
$$

