CHAPTER FIVE

5. Introduction to Probability Theory

5.1 Introduction

Probability and statistics are concerned with events which occur by chance. Examples include occurrence of accidents, errors of measurements, production of defective and non-detective items from a production line, and various games of chance, such as drawing a card from a well-mixed deck, flipping a coin, or throwing a symmetrical six-sided die. In each case we may have some knowledge of the likelihood of various possible results, but we cannot predict with any certainty the outcome of any particular trial. Probability and statistics are used throughout engineering. In electrical engineering, signals and noise are analyzed by means of probability theory. Civil, mechanical, and industrial engineers use statistics and probability to test and account for variations in materials and goods. Chemical engineers use probability and statistics to assess experimental data and control and improve chemical processes. It is essential for today’s engineer to master these tools.

Probability is an area of study which involves predicting the relative likelihood of various outcomes. It is a mathematical area which has developed over the past three or four centuries. One of the early uses was to calculate the odds of various gambling games. Its usefulness for describing errors of scientific and engineering measurements was soon realized. Engineers study probability for its many practical uses, ranging from quality control and quality assurance to communication theory in electrical engineering. Engineering measurements are often analyzed using statistics, and a good knowledge of probability is needed in order to understand statistics.

Probability: ***can be defined as a measure of the likelihood that a particular event will occur or it is a science of decision making with calculated risk in face of uncertainty.***

- It is a numerical measure with a value between 0 and 1 of such likelihood. Where the probability of zero indicates that the given event cannot occur and the Probability of one assures certainty of such an occurrence.

❖ **Mathematical models**

A mathematical model is a description of a system using mathematical concepts and language. Mathematical models are used not only in the natural sciences and engineering disciplines but also in the social sciences. Physicists, engineers, statisticians, operations research analysts and economists use mathematical models most extensively.

Mathematical models can be classified in different ways such as: Linear vs. nonlinear, Deterministic vs. probabilistic (stochastic), Static vs. dynamic and Discrete vs. Continuous:

5.1.1 Deterministic vs. non-deterministic (probabilistic) models

A deterministic model is one in which every set of variable states is uniquely determined by parameters in the model and by sets of previous states of these variables. Hypothesize exact relationships and it will be suitable when prediction error is negligible.

In a non-deterministic (stochastic model), randomness is present, and variable states are not described by unique values, but rather by probability distributions. Hence, there will be a defined pattern or regularity appears to construct a precise mathematical model. Hypothesize two components, which is deterministic and random error.

Example 5.1: Some examples of deterministic models (formulas)

a. Energy contained in a body moving in a vacuum with a speed of light $E = mc^2$

b. If the price of an item increases, then the demand for that item will decrease.

c. Body mass index (BMI) is measure of body fat $BMI = \frac{weight\ in\ kg}{(height\ in\ meters)^2}$

d. Systolic blood pressure of newborns is 6 Times the Age in days + Random Error

$$SBP = 6\ age\ (d) + \varepsilon$$

5.2 Review of set theory

Definition: *Set* is a collection of well-defined objects. These objects are called elements. Sets usually denoted by capital letters and elements by small letters. Membership for a given set can be denoted by $\in$ to show belongingness and $\notin$ to say not belong to the set.

- **Description of sets:**

Sets can be described by any of the following three ways. That is the complete listing method (all element of the set are listed), the partial listing method (the elements of the set can be indicated by listing some of the elements of the set) and the set builder method (using an open proposition to describe elements that belongs to the set).

Example 5.2:

The possible outcomes in tossing a six-side die

S = {1, 2, 3, 4, 5, 6} or

S = {1, 2 . . . 6} or

S = {x: x is an outcome in tossing a six-side die}

❖ Types of set

Universal set: is a set that contains all elements of the set that can be considered the objects of that particular discussion.

Empty or null set: is a set which has no element, denoted by {} or ϕ

Finite set: is a set which contains a finite number of elements. (eg. {x: x is an integer, 0 < x < 5})

Infinite set: is a set which contains an infinite number of elements. (eg. {x: x $\in \Re$, x > 0})

Sub set: If every element of set A is also elements of set B, set A is called sub sets of B, and denoted by A $\subseteq$ B.

Proper subset: For two sets A and B if A is subset of B and B is not sub set of A, then A is said to be a proper subset of B. Denoted by A $\subset$ B.

Equal sets*:* two sets A and B are said to be equal if elements of set A are also elements of set B.

Equivalent sets: Two sets A and B are said to be equivalent if there is a one to one correspondence between elements of the two sets.

❖ Basic properties of the set operations

Let U be the universal set and sets A, B, C are sets in the universe, the following properties will hold true.

1. $A \cup B = B \cup A$ (Union of sets is commutative)
2. $A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$ (Union of sets is associative)
3. $A \cap B = B \cap A$ (Intersection of sets is commutative)
4. $A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C$ (Intersection of sets is associative)

5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (union of sets is distributive over Intersection)
6. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Intersection of sets is distributive over union)
7. $A - B = A \setminus B = A \cap B^{`}$
8. If $A \subseteq B$, then $B^{`} \subseteq A^{`}$ or if $A \subset B$ then $B^{`} \subset A$
9. $A \cup \phi = A$ and $A \cap \phi = \phi$
10. $A \cup U = U$ and $A \cap U = A$
11. $(A \cup B)^{`} = A^{`} \cap B^{`}$ De Morgan's first rule
12. $(A \cap B)^{`} = A^{`} \cup B^{`}$ De Morgan's second rule
13. $A = (A \cap B) \cup (A \cap B^{`})$

❖ **Some of the corresponding statement in set theory and probability**

Set theory	**Probability theory**
Universal set, U	- Sample space S (sure event), P(S) = 1
Empty set ϕ	- Impossible event
Elements a, b…	- Sample point a, b, c… (or simple events)
Set A, B, C,. .	- Event A, B, C,.
Set A	- Event A occur
$A^{`}$	- Event A doesn\`t occur, $P(A') = 1 - P(A)$
$A \cup B$	- At least one of event A or B occur (the event containing all the elements that belong to A or B or both).
$A \cap B$	- Both events A and B occur (the event containing all elements that are common to A and B).
$A \subseteq B$	- The occurrence of A necessarily implies the occurrence of B
$A \cap B = \phi$	- A and B are mutually exclusive (That is, they cannot occur simultaneously)

❖ Several results that follow from the foregoing definitions, which may easily be verified by means of Venn diagrams, are as follows: ***(For any events A and B)***

1. $A \cap ø = ø$.
2. $A \cup ø = A$.
3. $A \cap A' = ø$.
4. $A \cup A' = S$.
5. $S' = ø$.
6. $ø' = S$.
7. $(A')' = A$.
8. $(A \cap B)' = A' \cup B'$.
9. $(A \cup B)' = A' \cap B'$.
10. $A' = 1 - A$

5.3 <u>Definitions of Some Basic Probability Terms</u>

1. **Random Experiment**: An experiment is any activity that generates *outcome(s)*. Results of experiments may not be the same even through conditions which are identical. Such experiments are called *random experiments*. For example, tossing of a fair coin and rolling a die are considered as a random experiment.
2. **Outcome:** is the result of an experiment.

<u>Example 5.3:</u>

- If an experiment consists of measuring "lifetimes" of electric light bulbs produced by a company, then the result of the experiment is a time t in hours that lies in some interval say, $0 \leq t \leq 4000$ where we assume that no bulb lasts more than 4000 hours.

<u>Example 5.4:</u>

Experiment	**Outcomes**
Tossing of a fair coin	Head, tail
Rolling a die	1, 2, 3, 4, 5, or 6
Selecting an item from a production lot	defective (faulty), non-defective (good)
Introducing a new product	Success, failure

3. **Sample space**: A sample space is the collection of all possible outcomes of an experiment. For example, there are two possible outcomes of a toss of a fair coin, which are a head and a tail. Then the sample space, for this experiment denoted by S would be: S = {H, T}. Each possible outcome in the sample space is called *sample point*.

4. **Event** is a subset of the sample space or it is a set containing sample points of a certain sample space under consideration. They are denoted by capital letters. For examples, getting two heads in the trial of tossing three fair coins simultaneously would be an event.

   Example 5.5:

   Considering the experiment of rolling a die, let A be the event of odd numbers, B be the event of even numbers, and C be the event of number 8.

   ⇒ **A**= {1, 3, 5} **B**= {2, 4, 5} **C** = {} *or empty space or impossible event*

5. **Elementary event** (simple event) is a single possible outcome of an experiment.

6. **Complement of an Event:** the complement of an event A means non- occurrence of A and is denoted by$A' or A^c or \bar{A}$ contains those points of the sample space which don't belong to A.

7. **Composite** (compound) **event** is an event having two or more elementary events in it. For example, rolling a die sample space = {1,2,3,4,5,6) an event having { 5} is simple event where as having even number={2,4,6} is compound (composite)event.

8. **Mutually exclusive events**: Two events are said to be mutually exclusive, if both events cannot occur at the same time as outcome of a single experiment. In other word two events E_1 and E_2 said to be mutually exclusive evens if there is no sample

point in common to both events E_1 and E_2. For example, if we roll a fair dice, then the experiment is rolling the die and Sample space (S) is

$S = \{1,2,3,4,5,6\}$

If we are interested the outcome of event E_1 getting even numbers and E_2 odd numbers

$E_1 = \{2, 4, 6\}$

$E_2 = \{1, 3, 5\}$

Clearly $E_1 \cap E_2 = \emptyset$. Thus E_1 and E_2 are mutually exclusive events.

9. **Exhaustive Events:** Events are said to be exhaustive if their union equals the sample space. For instance, when a die is rolled, the event of getting even numbers {2; 4; 6} and the event of getting odd numbers {1; 3; 5} are exhaustive events as the union of the events are equal to the sample space {1,2,3,4,5,6}. When two coins are tossed the event that at least one head will come up {HH, HT, TH} and the event that at least one tail will come up {TT, TH, HT} are exhaustive events as the union of the events are equal to the sample space {HH, HT, TH, TT}.

10. **Favorable Event:** Favorable event is an event about which the experimenter is concerned or interested. A favorable outcome is the outcome of interest. For instance, one can define a favorable outcome in the flip of a coin as a tail.

11. **Independent Events:** Two events A and B are said to be ***independent events*** if the occurrence of event A has no influence (bearing) on the occurrence of event B. ***<u>For example</u>***, if two fair coins are tossed, then the result of one toss is totally independent of the result of the other toss. The probability that a head will be the outcome of any one toss will always be ½, irrespective of whatever the outcome is of the other toss.

Hence, ***these two events are independent***. On the other hand, consider drawing two cards from a pack of 52 playing cards. The probability that the second card will be an ace would depend up on whether the first card was an ace or not. Hence these two events are ***not independent events.***

Another example a bag contains balls of two different colours say yellow and white. Two balls are drawn successively. First ball is drawn from a bag and replaced after notes its colour. Let us assume that it is yellow and denote this event by A. Another ball is drawn from the same bag and its colour is noted let this event denoted by B. Clearly, the result of first draw has no effect on the result of the second draw. Hence, the ***events A and B are independent events.***

12. Finite and infinite sample space

If a sample space has finite number of points, it is called a *finite sample space*. If it has as many points as natural numbers1, 2, 3…it is called a *countable infinite sample space*. If it has as many points as there are in some interval the *x*-axis, such as $0 < x < 1$, it is called a *non-countably infinite sample space*. A sample space which is finite or countably infinite is often called a *discrete sample space* while a set which is non countably infinite is called *non discrete* or *continuous sample space*.

Example 5.6: a) The result of the experiment making bolts observing defective. Thus, the outcome will be a member of the set {defective, non-defective}.

b) The lifetime of a bulb in example 5.3.

13. Equally likely outcomes: Equally likely outcomes are outcomes of an experiment which has equal chance (equally probable) to appear. In most case it is commonly assumed finite or countable infinite sample space is equally likely.

If we have n equally likely outcomes in the sample space then the probability of the i^{th} sample point x_i is p $(x_i) = \frac{1}{n}$, where x_i can be the first, second... or the n^{th} outcome.

Example 5.7:

In an experiment tossing a fair die, the outcomes are equally likely (each outcome are equally probable. Hence,

$$P(x_i = 1) = P(x_i = 2) = P(x_i = 3) = P(x_i = 4) = P(x_i = 5) = P(x_i = 6) = \frac{1}{6}$$

5.4 Counting techniques

If the number of possible outcomes in an experiment is small, it is relatively easy to list and count all possible events. When there are large numbers of possible outcomes an enumeration of cases is often difficult, tedious, or both. Therefore, to overcome such problems one can use various counting techniques or rules.

In order to calculate probabilities, we have to know

- The number of elements of an event
- The number of elements of the sample space.

That is in order to judge what is **probable**, we have to know what is **possible.**

- In order to determine the number of outcomes (possibilities), one can use several rules of counting. These are;
  - The addition rules
  - The multiplication rules
  - Permutation rule
  - Combination rule

5.4.1 Addition rule

Suppose that a procedure designated by 1, can be performed in n_1 ways. Assume that second procedure designated by 2 can be performed in n_2 ways. Suppose furthermore that it is not possible both procedures 1 and 2 are performed together. The number of ways in which we can perform 1 or 2 procedures is $n_1 + n_2$ ways.

This can be generalized as follows if there are k procedure and i^{th} procedures may be performed in n_i ways, i=1, 2… k, then the number of ways in which we perform procedure 1 or 2 or … or k is given by $n_1+n_2+\ldots+n_k=\sum_{i=1}^{k} n_i$, assuming that no two procedures performed together.

Example 5.8:

Suppose that we are planning a trip and are deciding between bus and train transportation. If there are 3 bus routes and 2 train routes to go from A to B, find the available routes for the trip. There are 3+2 = 5 possible routes for someone to go from A to B.

5.4.2 Multiplication Rule [*Fundamental Principle of counting (mn Rule)*]

Rule 1: Suppose that procedure 1 can be performed in n_1 ways. Let us assume procedure 2 can be performed in n_2 ways. Suppose also that each way of doing procedure 2 may be followed by any way of doing procedure 1, then the procedure consisting of n_1 followed by n_2 may be performed by $n_1 * n_2$ ways.

Example 5.9:

An airline has 6 flights from A to B, and 7 flights from B to C per day. If the flights are to be made on separate days, in how many different ways can the airline offer from A to C?

Soln.

In operation 1 there are 6 flights from A to B, 7 flights are available to make flight from B to C. Altogether there are 6*7 = 42 possible flights from A to C.

Rule 2: If an operation can be performed in n1 ways, and if for each of these a second operation can be performed in n2 ways, and for each of the first two a third operation can be performed in n3 ways, and so forth, then the sequence of k operations can be performed in n1n2 · · · nk ways.

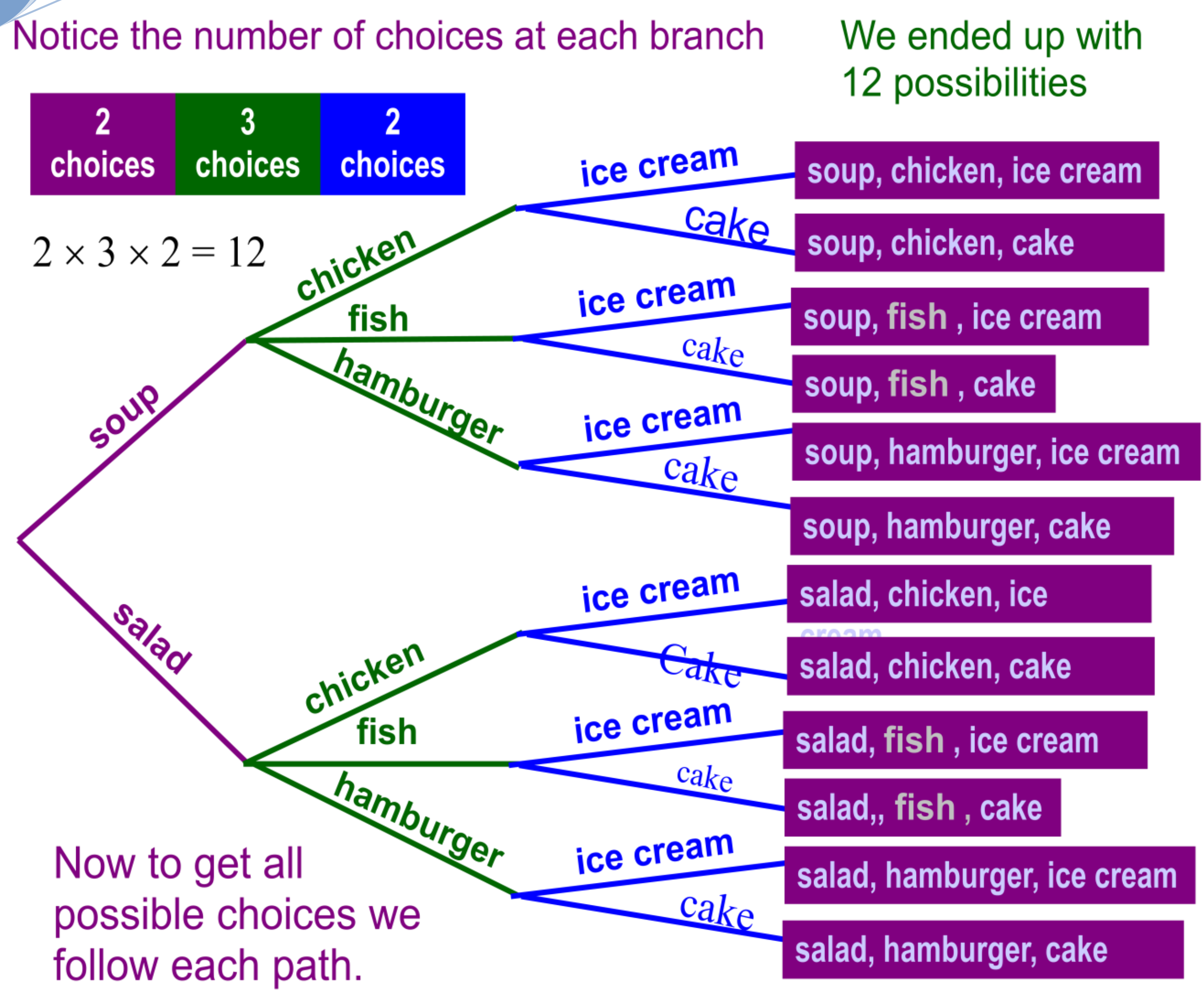
Example 5.10:

Suppose that in a medical study patients are classified according to their blood type as A, B , AB, and O; according to their RH factors as + or - and according to their blood pressure as high, normal or low, then in how many different ways can a patient be classified ?

Soln.

The1st classification has done in 4 ways; the 2nd in 2 ways, and the 3rd in 3 ways. Thus, patient can be classified in 4*2*3 = 24 different ways.

- The following diagram also shows the number of possibility to chice an items from supermarket which follows a multiplication rule.



Example 5.11:

How many even four-digit numbers can be formed from the digits 0, 1, 2, 5, 6, and 9 if each digit can be used only once?

Solution: Since the number must be even, we have only n1 = 3 choices for the unit's position. However, for a four-digit number the thousands position cannot be 0. Hence, we consider the units position in two parts, 0 or not 0. If the units' position is 0 (i.e., n1 = 1), we have n2 = 5 choices for the thousands position, n3 = 4 for the hundreds position, and n4 = 3 for the tens position. Therefore, in this case we have a total of:

$n_1n_2n_3n_4 = (1)(5)(4)(3) = 60$

even four-digit numbers. On the other hand, if the units' position is not 0 (i.e., n1 = 2), we have n2 = 4 choices for the thousands position, n3 = 4 for the hundreds position, and n4 = 3 for the tens position. In this situation, there are a total of

$n_1n_2n_3n_4 = (2)(4)(4)(3) = 96$

even four-digit numbers.

Since the above two cases are mutually exclusive, the total number of even four-digit numbers can be calculated as **60 + 96 = 156*.***

5.4.3 <u>Permutation Rule</u>

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

Permutation is an arrangement of all or parts of a set of objects with regard to **order.**

<u>Rule 1:</u> The number of permutations of n distinct objects taken all together is n! **or** In particular, the number of permutations of n objects taken n at a time is

$nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$ Where n! = n (n-1)(n-2)---(n-n)!..................*in definition 0! = 1! = 1*

***<u>Example 5.12</u>*:**

In how many ways 4 people are lined up to get on a bus (or to sit for photo graph)?

In 4! = 4*3*2*1 = 24 Ways.

<u>Rule-2</u>: A permutation of n different objects taken r at a time is an arrangement of r out of the n objects, with attention given to the order of arrangement. The number of permutations of n objects taken r at a time is denoted by $_nP_r$, or P (n,r) and is given by

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

Example 5.13:

The number of permutations of letters a, b & c taken two at a time is

$_3P_2 = 6$. These are ab, ba, ac, ca, bc & cb.

Rule-3: The number of permutations of n objects taken all at a time, when n_1 objects are alike of one kind, n_2 objects are alike of second kind, …, n_k objects are alike of k^{th} kind is given by:

$$\frac{n!}{n_1!\,n_2!\,n_3!,\ldots,n_k!} = \frac{(\sum_{i=1}^{k} n_i)!}{\prod_{i=1}^{k}(n_i)!}$$

Example 5.14

The total number of arrangements of the letters of the word *STATISTICS* taken all at a time is given by $\frac{10!}{3!3!1!2!1!}$ = 50,400 since there are 3s's, 3t's, 1a, 2i's and 1c.

Rule 4: The number of arrangements of n distinct objects around circular object (table) is (n-1)! And when the method of selection or arrangement of r objects from n objects with repetition the possible numbers of arrangements are n^r.

Example 5.15

Adama Science and Technology University Registrar Office want to give identity number for students by using 4 digits. The number should be considered by the following numbers only: {0, 1, 2, 3, 4, 5, and 6}. Hence, how many different ID Numbers could be preferred by the Registrar

a. Without repeating the number?

b. With repetition of numbers?

We have 7 possible numbers for 4 digits. But the required number of digits for ID number is 4. Hence n = 7 & r = 4. The possible number of ID numbers given for students without repeating the digit is

$$_{n}P_{r,} = \frac{n!}{(n-r)!} = \frac{7!}{(7-4)!} = 7*6*5*4 = 840.$$

☞ The possible number of ID numbers given for students with repeating the digit is

$$n^r = 7^4 = 7*7*7*7 = 2401$$

Exercise:

1. Suppose we have a letters A, B, C, D

   a) How many permutations are there taking all the four?

   b) How many permutations are there two letters at a time?

2. How many different permutations can be made from the letters in the word "COORRECTION"?

3. How many different arrangements are possible for sitting 10 people around a circular table?

5.4.4 Combinations Rule

Combination is the selection of objects without regarding order of arrangement. A combination of n different objects taken r at a time is a selection of r out of n objects, with no attention given to the order of arrangement. The total number of combinations of r objects selected from n (also called the combinations of n things taken r at a time) is denoted by $_{n}C_{r}$ or $\binom{n}{r}$ or C_{r}^{n}.

$$_nC_r = \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!} = \frac{_nP_r}{r!}$$

Example 5.16

The number of combinations of letters a, b& c taken two at a time is 3C2= $\frac{3!}{2!1!}$ = 3.

These are ab, ac and bc. Note that *ab* is the same combination as *ba*, but not the same permutation.

Example 5.17:

Suppose in the box 3 red, 3 white and 5 black equal sized balls are there. We want to draw 3 balls at a time. How many ways do we have from each type?

Solution

$$\binom{3}{1}\binom{3}{1}\binom{5}{1} = 3(3)5 = 45 \text{ ways.}$$

Example 5.18:

A bag contains 3 red, 6 white and 7 blue balls. What is the probability that two balls drawn are white and blue? ***Solution*:** Total number of balls = 3 + 6 + 7 = 16

Now, out of 16 balls, 2 can be drawn in $_{16}C_2$ ways.

$\Rightarrow$ Exhaustive number of cases = $^{16}C_2 = \frac{16*15}{2} = 120$

Out of 6 white balls, 1 ball can be drawn in $^6C_1$ ways and out of 7 blue balls, one can be drawn is $^7C_1$ ways. Since each of the former case is associated with each of the latter case, therefore total number of favorable cases are $^6C_1$ * $^7C_1$ = 6 * 7 = 42.

$$\Rightarrow \text{required probability} = \frac{42}{120} = \frac{7}{20}$$

Exercise:

1. Among 15 clocks there are two defectives. In how many ways can an inspector chose three of the clocks for inspection so that:
   a) There is no restriction.
   b) None of the defective clock is included.
   c) Only one of the defective clocks is included.
   d) Two of the defective clocks is included.

5.5 Different Approaches to probability

In any random experiment there is always uncertainty as to whether a particular event will or will not occur. As a measure of the chance, or probability, with which we can expect the event to occur, it is convenient to assign a number between 0 and 1. If we are sure or certain that the event will occur, we say that its probability is 100% or 1, but if we are sure that the event will not occur, we say that its probability is zero.

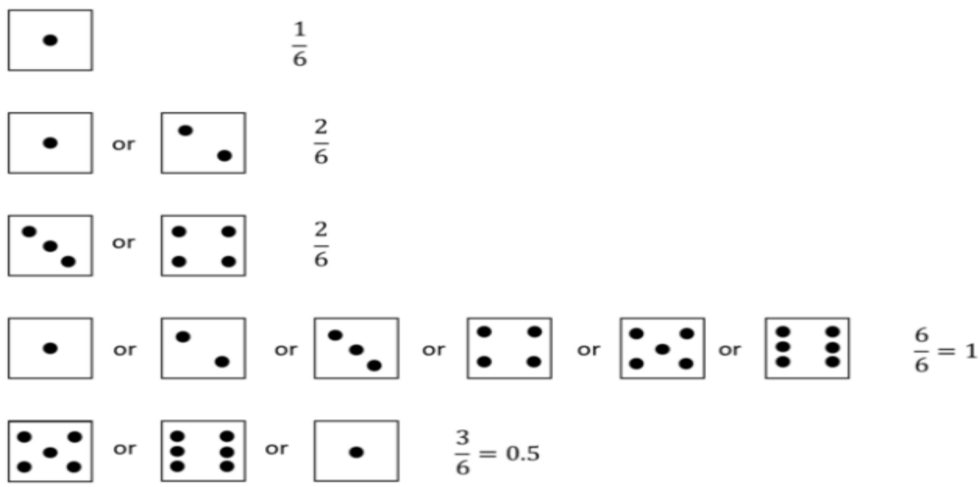
There are different approaches by means of which we can define or estimate the probability of an event. These approaches are discussed below:

1. Classical Approach

This approach traces back to the field where probability was first systematically employed, which is gambling (flipping coins, tossing dice and so forth). Gambling problems are characterized by random experiments which have n possible outcomes, equally likely to occur. It means that none of them is more or less likely to occur than other ones, hence they are said to be in a symmetrical position. The idea of the classical approach is that, given a collection of k elements out of n (where $0 \leq k \leq n$), the probability of occurrence of the event E represented by that collection is equal to:

$$\frac{k}{n}$$

To give you the intuition, let's imagine you are tossing a dice and you want to predict the probability of the following collection of outcomes:



We know that the n possible outcomes are 6. The event "one" is 1 out of 6 outcomes; hence its probability is 1/6. Similarly, the event "five or six or one" (that is, the event in which one of those numbers turns out) represents 3 outcomes out of 6, hence the probability will be 3/6=0.5.

The classical approach is pretty intuitive; nevertheless, it suffers from some pitfalls:

- The assumption of symmetry is far too strong and irrealistic. Namely, imagine you want to know the probability of the event "tomorrow I will have a car accident". The possible outcomes of this scenario are two: having a car accident or not having a car accident. Given that k=having a car accident, the probability of that event is 1/2, which, besides being a bit worrying, is not representative of the real likelihood of the event.
- In this approach, there is no space for the concept of information, which is strictly related to probability. Let's think about the previous example of the dice. Imagine

you are told this dice is loaded and, instead of having the number “one”, it has two “six” (so the faces will be 2, 3, 4, 5, 6, 6). Provided with this information, which probability would you attribute to the event “one”? Since it is impossible, the probability is equal to zero and not 1/6. Hence, probability does depend on the available information (the intuition will be clearer in the subjective approach)

Example 5.19:

In rolling 2 *dice* and summing the 2 numbers on top. What is the sample space? What are the probabilities of P (2), P (7) and P (10)?

2. **Frequency-based (or empirical) approach**

This approach was formally introduced in the field of natural science, where the assumption of symmetric position poorly fails. Instead, the idea on which this approach is based is that several experiments can be run under certain conditions considered as equivalent. Each experiment might lead either to success or to failure.

Here in this approaches Probabilities are assigned on the basis of experimentation or historical data. Formally, Let A be an event of interest, and assume that you have performed the same experiment n times so that n is the number of times A could have occurred. Further, let n_A be the number of times that A did occur. Now, consider the relative frequency $\frac{n_A}{n}$**.** Then, in this method, we “attempt” to define P (A) as:

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

The above can only be viewed as an attempt because it is not physically feasible to repeat an experiment an infinite number of times. Another important issue with this definition is that two sets of n experiments will typically result in two different ratios. However, we expect the discrepancy to converge to 0 for large n. Hence, for large n, the ratio $\frac{n_A}{n}$ may be taken as a reasonable approximation for P (A).

For instance, if we roll a die we can have a sample space: S = {1, 2, · · · , 6}

Probabilities: Roll the given die 100 times (say) and supposes the number of times the outcome 1 is observed is 15. Thus, A = {1}, n_A = 15, and n = 100. Therefore, we can say that P(A) is approximately equal to 15/100 = 0.15.

Let's see the difference between the frequency-based and classical approach with the following example. Imagine you want to know the probability of the outcome of your tossed coin being "head". You start with your classical approach: since the possible n outcomes are two (head or tail), the probability of "head" is 1/2=0.5.

Now you decide to follow the empirical approach, and you start tossing your coin several times, let's say 100. Out of your attempts, you obtained 55 "head" and 45 "tail". Hence, the frequency of the event "head" is 55/100=0.55, and it can approximate the probability of the event "head".

As you can see, we obtained two different probabilities (0.5 vs. o.55) for the same event. The key difference is the role of information: after 100 experiments, you gathered empirical evidence that "head" occurred more often than "tail": it might be that your coin is not perfect, and you can incorporate this information while formulating your conclusions.

This approach is not lacking of criticisms though:

- Again, there is one big assumption which is the convergence property of the frequency, whose limit might not exist
- Repeating experiments under equivalent conditions might not be possible
- There are events extremely rare, for which is impossible to run many simulations (think about extreme natural events like tsunami).

Example 5.19:

Bits & Bytes Computer store tracks the daily sales of desktop computers in the past 30 days. The resulting data are:

Desktops Sold	No. of Days
0	1
1	2
2	10
3	12
4	5
5 or more	0

The approximate probabilities are:

Desktops Sold	No. of Days	Probability
0	1	$^{1}/_{30} = 0.03$
1	2	$^{2}/_{30} = 0.07$
2	10	$^{10}/_{30} = 0.33$
3	12	$\mathbf{^{12}/_{30} = 0.40}$
4	5	$^{5}/_{30} = 0.17$
5 or more	0	0

Thus, for example, there is a 40% chance that the store will sell 3 desktops on any given day.

Example 5.21:

What about randomly selecting a student and observing their gender? S = {Male, Female}. Are these probabilities ½?

3. Subjective approach

Developed by probabilist B. de Finetti, this is the most intuitive definition of probability. Indeed, according to that approach, the probability of an event is the degree of belief a person attaches to that event, based on his/her available information. This reasoning holds only under the assumption of rationality, which assumes that people act coherently.

Let's provide a more specific definition. Imagine a lottery where you can win an amount of money equal to S if event E occurs. To participate, you have to buy one ticket. Now, which is the price you would be willing to pay to participate in the lottery? If you indicate that price as π (E, S), the probability of event E is given by:

$$P(E) = \frac{\pi(E,S)}{S}$$

Imagine you want to predict the probability that your favourite football team will win the match tomorrow. You have the possibility to participate in a lottery where, if the team wins, you obtain a prize of 1000birr, otherwise you gain nothing. Which is the price you would be willing to pay to participate? Let's say you are very confident about your team capabilities and you are willing to pay 700birr. Hence, the probability your team wins the match tomorrow is:

$$P("win") = \frac{\pi("win", 1000)}{1000} = \frac{700}{1000} = 0.7$$

This last approach does not count serious criticisms, since it resolves some pitfalls of the previous approaches (like the impossibility of repeating experiments under equivalent conditions, because of the uniqueness of many events) and, at the same time, does not contrast with other theories. Indeed, the evaluator who has to decide the price of the lottery is not prevented from running experiments, compute the frequency of successes and use this information to propose a price.

❖ Axioms of Probability

Now that the probability of an event has been defined, we can collect the assumptions that we have made concerning probabilities into a set of axioms that the probabilities in any random experiment must satisfy. The axioms ensure that the probabilities assigned in an experiment can be interpreted as relative frequencies and that the assignments are consistent with our intuitive understanding of relationships between relative frequencies. For example, if event A is contained in event B, we should have $P(A) \leq P(B)$. The axioms do not determine probabilities; the probabilities are assigned based on our knowledge of the system under study. However, the axioms enable us to easily calculate the probabilities of some events from knowledge of the probabilities of other events.

- Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties:

  If S is the sample space and E is any event in a random experiment,

  i. $P(S) = 1$
  ii. $0 \leq P(E) \leq 1$
  iii. For two events E_1 and E_2 with $E_1 \cap E_2 = \emptyset$

     $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

❖ ADDITION RULES

Joint events are generated by applying basic set operations to individual events. Unions of events, such as $A \cup B$; intersections of events, such as $A \cap B$; and complements of events, such as A’, are commonly of interest. The probability of a joint event can often be determined from the probabilities of the individual events that comprise it. Basic set operations are also sometimes helpful in determining the probability of a joint event.

- The probability $P(A \cup B)$ is interpret E_1ed as the probability of A or B and to compute that the following general addition rule applies.

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

- If A and B are mutually exclusive events,

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

- For three or more events:

  More complicated probabilities, such as $P(A \cup B \cup C)$, can be determined by repeated use of addition rule and by using some basic set operations. For example,

$$\boldsymbol{P(A \cup B \cup C)}$$

$$\boldsymbol{= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)}$$

- If A, B and C are mutually exclusive events,

$$\boldsymbol{P(A \cup B \cup C) = P(A) + P(B) + P(C)}$$

<u>Example:</u> If A, B and C are mutually exclusive events, is it possible to have P (A) = 0.3, P (B) = 0.4, and P(C) = 0.5? if not why?

- A collection of events, $E_1, E_2, E_3, \ldots, E_k$ is said to be mutually exclusive if for all pairs,

$$E_i \cap E_j = \emptyset$$

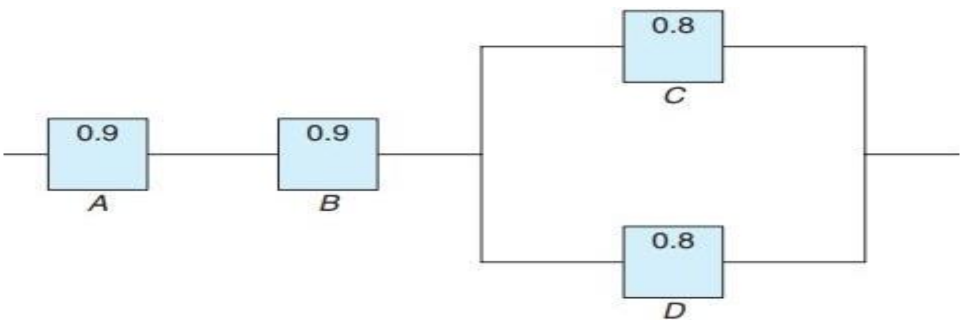
- For a collection of mutually exclusive events,

$$\boldsymbol{P(E_1 \cup E_2 \cup E_3 \ldots \cup E_K) = P(E_1) + P(E_2) + P(E_3) \ldots + P(E_k)}$$

Example 5.22:

An electrical system consists of four components as illustrated in Figure 2.9. The system works if components *A* and *B* work and either of the components *C* or *D* works. The reliability (probability of working) of each component is also shown in Figure 2.9. Find the probability that

a) The entire system works and
b) The component *C* does not work, given that the entire system works. Assume that the four components work independently.



Soln:

In this configuration of the system, *A*, *B*, and the subsystem *C* and *D* constitute a serial circuit system, whereas the subsystem *C* and *D* itself is a parallel circuit system.

(a) Clearly the probability that the entire system works can be calculated as follows:

$$P[A \cap B \cap (C \cup D)] = P(A)P(B)P(C \cup D) = P(A)P(B)[1 - P(C' \cap D')]$$

$$= P(A)P(B)[1 - P(C')P(D')]$$

$$= (0.9)(0.9)[1 - (1 - 0.8)(1 - 0.8)] = 0.7776.$$

The equalities above hold because of the independence among the four components.

(b) To calculate the conditional probability in this case, notice that

$$P = \frac{P(\text{the system works but } C \text{ does not work})}{P(\text{ the system works})} = \frac{P(A \cap B \cap C' \cap D)}{P(\text{ the system works})}$$

$$= \frac{(0.9)(0.9)(1 - 0.8)(0.8)}{0.7776} = 0.1667$$

5.6 Conditional Probability

For two events A and B, the probability of an event B occurring when it is known that some event A has occurred is called a **conditional probability** and is denoted by P(B|A). The symbol P (B|A) is usually read "the probability that B occurs given that A occurs" or simply "the probability of B, given A.

Let there are two events A and B, then the probability of event A given that the outcome of event B is given by:

$$P\,[A|B] = \frac{P(A\cap B)}{P(B)} \text{Provided } P(B) > 0$$

Where P [A |B] is interpreted as the probability of event A on the condition that event B has occurred. In this case P [A n B] is the joint probability of event A and B.

Example 5.23:

120 employees of a certain factory are given a performance test and are divided in to two groups as those with good performance(G) and those with poor performance (P) the result is given below

	Good performance (G)	Poor performance(P)	Total
Male (M)	60	20	80
Female (F)	25	15	40
Total	85	35	120

The probability of a person to be male given that it has a good performance is

$$P\,(M|G) = \frac{P(M\cap G)}{P(G)} = \frac{60/_{120}}{85/_{120}} = \frac{60}{85} = \mathbf{0.706}$$

The probability of a person to be female given that it has a poor performance is

$$P\,(F|P) = \frac{P(M\cap P)}{P(P)} = ?\ \textbf{Exercise!}$$

Example 5.24:

A jar contains black and white marbles. Two marbles are chosen without replacement. The probability of selecting a black marble and a white marble is 0.34, and the probability of selecting a black marble on the first draw is 0.47. What is the probability of selecting white marble on the second draw, given that the first marble drawn is black?

$$P(White \mid Black) = \frac{P(Black\ and\ White)}{P(Black)} = \frac{0.34}{0.47} = 0.72$$

Example5.25:

The probability that it is Friday and that a student is absent is 0.03. Since there are 5 schooldays in a week, the probability that it is Friday is 0.2. What is the probability that a student is absent given that today is Friday?

$$P(Absent \mid Friday) = \frac{P(Friday\ and\ Absent)}{P(Firday)} = \frac{0.03}{0.2} = 0.15$$

Example5.26:

Suppose that an office has 100 calculating machines. Some of them use electric power (E) while others are manual (M) and some machines are well known (N) while others are used (U). The table below gives numbers of machines in each category. A person enter the office picks a machine at random and discovers that it is new. What is the probability that it is used with electric power?

(**Exercise!)**

	E	M	Total
N	40	30	70
U	20	10	30
Total	60	40	100

Probability of Independent Events

Two events A and B are independent if and only if the following equivalent statements is true

i. P (A∩B) = P (A) * P (B).
ii. P (A/B) = P (A) and
iii. P(B/A) = P(B)

Example 5.27:

A box contains four black and six white balls. What is the probability of getting two black balls in drawing one after the other under the following conditions?

a. The first ball drawn is not replaced
b. The first ball drawn is replaced

Solution; Let A= first drawn ball is black, B = second drawn is black

Required P ($A \cap B$)

a. P (A∩B) = P (A). P (B|A) = 4/10 * 3/9 = 2/15

b. P (A∩B) = P (A). P (B) = 4/10 * 4/10 = 4/25

5.7 Multiplication Rule, Total Probability Rules, and Bayes' Theorem

❖ **Multiplication Rul**

The definition of conditional probability can be rewritten to provide a general expression for the probability of the intersection of two events. This formula is referred to as a **multiplication rule** for probabilities. And expressed as;

$$\boldsymbol{P(A \cap B) = P(A/B) * P(B) = P(B/A) * P(A)}$$

Example 5.28; The probability that an automobile battery subject to high engine compartment temperature suffers low charging current is 0.7. The probability that a battery is subject to high engine compartment temperature is 0.05. Let C denote the event that a battery suffers low charging current, and let T denote the event that a battery is subject to high engine compartment temperature. The probability that a battery is subject to low charging current and high engine compartment temperature is;

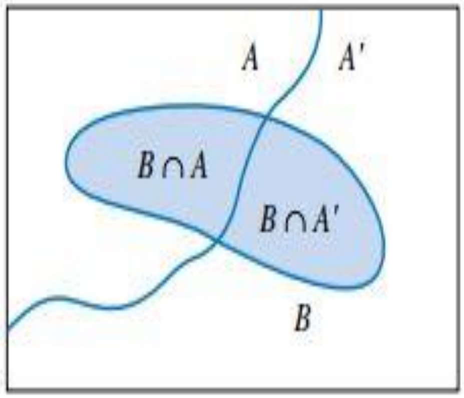
$$P(C \cap T) = P(C/T) * P(T) = 0.7 * 0.05 = 0.035$$

❖ Total Probability Rule

The multiplication rule is useful for determining the probability of an event that depends on other events. For example, suppose that in semiconductor manufacturing the probability is 0.10 that a chip that is subjected to high levels of contamination during manufacturing causes a product failure. The probability is 0.005 that a chip that is not subjected to high contamination levels during manufacturing causes a product failure. In a particular production run, 20% of the chips are subject to high levels of contamination. What is the probability that a product using one of these chips fails? Clearly, the requested probability depends on whether or not the chip was exposed to high levels of contamination. We can solve this problem by the following reasoning. For any event *B*, we can write *B* as the union of the part of *B* in *A* and the part of *B* in A'. That is,

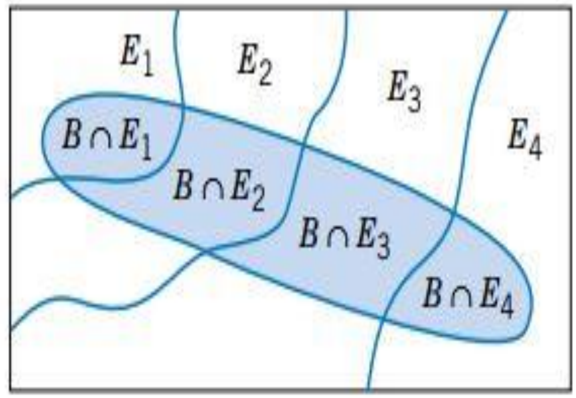
$$B = (\boldsymbol{A} \cap \boldsymbol{B}) \cup (\boldsymbol{A'} \cap \boldsymbol{B})$$

This result is shown in the Venn diagram bellow. Because *A* and A' are mutually exclusive, A∩B and A'∩B are mutually exclusive. Therefore, from the probability of the union of mutually exclusive events and the Multiplication Rule, the following **total probability rule** is obtained.



$B = (\boldsymbol{A} \cap \boldsymbol{B}) \cup (\boldsymbol{A'} \cap \boldsymbol{B})$

Partitioning an event in to two mutually exclusive subsets



$B = (\boldsymbol{B} \cap \boldsymbol{E_1}) \cup (\boldsymbol{B} \cap \boldsymbol{E_2}) \cup (\boldsymbol{B} \cap \boldsymbol{E_3}) \cup (\boldsymbol{B} \cap \boldsymbol{E_4})$

Partitioning an event in to several mutually exclusive subsets

- For any two events **A** and **B** the total probability can be stated as:

$$\mathrm{P(B) = P(B \cap A) + P(B \cap A')}$$

$$\mathrm{= P(B|A)P(A) + P(B|A')P(A')}$$

- **Total Probability Rule for Multiple Events;**

Assume$E_1, E_2, E_3 \ldots . E_k$ are k mutually exclusive and exhaustive sets. Then

$$P(B) = P(B \cap E_1) \cup P(B \cap E_2) \cup P(B \cap E_3) \ldots \cup P(B \cap E_k)$$

$$= P(B \cap E_1) + P(B \cap E_2) + P(B \cap E_3) \ldots + P(B \cap E_k)$$

$$= P(B|E_1)P(E_1) + P(B|E_2)P(E_2) \ldots + P(B|E_k)P(E_k)$$

Example 5.29: In a certain assembly plant, three machines, *B*1, *B*2, and *B*3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective.

Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

Soln

- Consider the following events:

A: the product is defective,

B1: the product is made by machine B1,

B2: the product is made by machine B2,

B3: the product is made by machine B3.

Applying the rule of total probability, we can write

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\ &= 0.02(0.3) + 0.03(0.45) + 0.02(0.25) \\ &= 0.006 + 0.0135 + 0.005 \\ &= 0.0245 \end{aligned}$$

❖ Bayes' Theorem

In some examples, we do not have a complete table of information like in the case of conditional probability above. We might know one conditional probability but would like to calculate a different one.

From the definition of conditional probability,

$$\mathrm{P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)}$$

Now considering the second and last terms in the expression above, we can write:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \text{ for } P(B) > 0$$

This is a useful result that enables us to solve for P(A|B) in terms of P(B|A).

In general, if P(B) in the denominator of above equation is written using the Total Probability Rule, we obtain the following general result, which is known as **Bayes 'Theorem.**

- If $E_1, E_2, E_3 \ldots . E_k$ are k mutually exclusive and exhaustive events and B is any event,

$$P(E_1|B) = \frac{P(B|E_1)P(E_1)}{P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \cdots + P(B|E_k)P(E_k)}$$

Example 2.30: Because a new medical procedure has been shown to be effective in the early detection of an illness, a medical screening of the population is proposed. The probability that the test correctly identifies someone with the illness as positive is 0.99, and the probability that the test_correctly identifies someone without the illness as negative is 0.95. The incidence of the

illness in the general population is 0.0001. You take the test, and the result is positive. What is the probability that you have the illness?

Soln

- Let D denote the event that you have the illness, and
- Let S denote the event that the test signals positive.

The probability requested can be denoted as $P(D|S)$. The probability that the test correctly signals someone without the illness as negative is 0.95. Consequently, the probability of a positive test without the illness is;

$$P(S|D') = 0.5$$

From Bayes' Theorem,

$$P(D|S) = \frac{P(S|D)P(D)}{P(S|D)P(D) + P(S|D')P(D')}$$

$$= \frac{0.99(0.0001)}{0.99(0.0001) + 0.05(1 - 0.0001)} = \frac{0.000099}{0.5} = 0.0002$$

Example 5.31: While watching a game of Champions League football in a cafe, you observe someone who is clearly supporting Manchester United in the game. What is the probability that they were actually born within 25 miles of Manchester? Assume that:

- the probability that a randomly selected person in a typical local bar environment is born within 25 miles of Manchester is 1/20, and;
- the chance that a person born within 25 miles of Manchester actually supports United is 7/10;
- The probability that a person not born within 25 miles of Manchester supports United with probability 1/10.

Soln Define

- B - event that the person is born within 25 miles of Manchester
- U - event that the person supports United.

We want $P(B|U)$: Using Bayes' Theorem;

$$(P|U) = \frac{P(U|B)P(B)}{P(U)} = \frac{P(U|B)P(B)}{P(U|B)P(B) + P(U|B')P(B')}$$

$$= \frac{\frac{7}{10} * \frac{1}{20}}{\frac{7}{10} * \frac{1}{20} + \frac{1}{10} * \frac{19}{20}} = \frac{7}{26} = 0.269$$