CHAPTER ONE

INTRODUCTION TO ENGINEERING STATISTICS

1.1 Introduction

Engineering is about bridging the gaps between problems and solutions, and that process requires an approach called the scientific method. Many aspects of engineering practice involve collecting, working with, and using data in the solution of a problem, so knowledge of **statistics** is just as important to the engineer as knowledge of any of the other engineering sciences. **Statistical methods** are a powerful aid in designing new products and systems, improving existing designs, and designing, developing, and improving production operations. Statistical methods are used to help us describe and understand variability.

All natural processes, as well as those devised by humans, are subject to variability. Civil engineers are aware, for example, the compressive strength of concrete, soil pressures, traffic flow, floods, and pollution loads in streams have wide variations. To cope with uncertainty, the engineer must first obtain and investigate a sample of data, such as a set of flow data or soil test results. The sample is used in applying statistics and probability at the descriptive stage. For inferential purposes, however, one needs to make decisions regarding the population from which the sample is drawn. A data set comprises a number of measurements of a phenomenon such as the failure load of a structural component. The quantities measured are termed variables, each of which may take any one of a specified set of values.

1.2 Inferential statistics and probability models

After the preceding experiment is completed and the data are described and summarized, we hope to be able to draw a conclusion. This part of statistics, concerned with the drawing of conclusions, is called inferential statistics. To be able to draw a conclusion from the data, we must take into account the possibility of chance. To be able to draw logical conclusions from data, we usually make some assumptions about the chances (or probabilities) of obtaining the different data values. The totality of these assumptions is referred to as a probability model for the data. Sometimes the nature of the data suggests the form of the probability model that is assumed.

In other situations, the appropriate probability model for a given data set will not be readily apparent. However, careful description and presentation of the data sometimes enable us to infer a reasonable model, which we can then try to verify with the use of additional data. Because the basis of statistical inference is the formulation of a probability model to describe the data, an understanding of statistical inference requires some knowledge of the theory of probability. In other words, statistical inference starts with the assumption that important aspects of the phenomenon under study can be described in terms of probabilities; it then draws conclusions by using data to make inferences about these probabilities.

1.3 Populations and samples

In everyday language, the word "population" refers to all the people or organisms contained within a specific country, area, region, etc. When we talk about the population of IRAQ, we usually mean something like "the total number of people who currently reside in IRAQ" In the field of statistics, however, the term population is defined operationally by the question we ask: it is the entire collection of measurements about which we want to make a statement.

In statistics, we are interested in obtaining information about a total collection of elements, which we will refer to as the population. The population is often too large for us to examine each of its members. In such cases, we try to learn about the population by choosing and then examining a subgroup of its elements. This subgroup of a population is called a sample.

A population consists of all possible observations available from a particular probability distribution. A sample is a particular subset of the population that an experimenter measures and uses to investigate the unknown probability distribution. A random sample is one in which the elements of the sample are chosen at random from the population, and this procedure is often used to ensure that the sample is representative of the population.

1.4 A brief definition in statistics

Statistics is a discipline of study dealing with the collection, analysis, interpretation, and presentation of data. **Descriptive statistics** is the use of graphs, charts, and tables and the calculation of various statistical measures to organize and summarize information. Descriptive statistics help to reduce our information to a manageable size and put it into focus. The **population** is the complete collection of individuals, items, or data under consideration in a statistical study. The portion of the population selected for analysis is called the **sample**. **Inferential statistics** consist of techniques for reaching conclusions about a population based upon information contained in a sample. A **variable** is a characteristic of interest concerning the individual elements of a population or a sample. A variable is often represented by a letter such as x, y, or z. The value of a variable for one particular element from the sample or population is called an **observation**. A **data set** consists of the observations of a variable for the elements of a sample. A **quantitative variable** is determined when the description of the characteristic of interest results in a numerical value. When a measurement is required to describe the characteristic of interest or it is necessary to perform a count to describe the characteristic, a quantitative variable is defined. A **discrete variable** is a quantitative variable whose values are countable. Discrete variables usually result from counting. A **continuous variable** is a quantitative variable that can assume any numerical value over an interval or over several intervals. A continuous variable usually results from making a measurement of some type. A **qualitative variable** is determined when the description of the characteristic of interest results in a nonnumeric value. A qualitative variable may be classified into two or more categories.

CHAPTER TWO

PRESENTATION OF STATISTICS DATA

2.1 Introduction

Once a data set has been collected, the experimenter's next task is to find an informative way of presenting it. In general, a table of numbers is not very informative, whereas a picture or graphical representation of the data set can be quite informative. If "a picture is worth a thousand words," then it is worth at least a million numbers.

2.2 Frequency distributions

One way of describing a distribution of sample values, which is particularly useful in large samples, is to construct a frequency distribution of the sample values. We distinguish between two types of frequency distributions, namely, frequency distributions of: (i) discrete variables; and (ii) continuous variables.

A random variable, X, is called discrete if it can assume only a finite number of different values. For example, the number of defective computer cards in a production lot is a discrete random variable. A random variable is called continuous if, theoretically, it can assume all possible values in a given interval. For example, the output voltage of a power supply is a continuous random variable.

2.2.1 Frequency distributions of discrete random variables

Consider a random variable, X, that can assume only the value $x_1, x_2, x_3, \ldots, x_k$, where $x_1 < x_2 < \ldots.. < x_k$. Suppose that we have made n different observations on X. The frequency of x_i $(i=1,\ldots k)$ is defined as the number of observations having the value x_i. We denote the frequency of x_i by f_i. Notice that

$$\sum_{i=1}^{k} f_i = f_1 + f_2 + \ldots\ldots + f_k = n$$

We can present a frequency distribution in a tabular form as:

Value	Frequency
x_1	f_1
x_2	f_2
.	.
.	.
.	.
x_k	f_k
Total	n

It is sometimes useful to present a frequency distribution in terms of the proportional or relative frequencies p_i, which are defined by

$$p_i = \frac{f_i}{n} \qquad (i = 1, 2, \ldots\ldots, k)$$

In addition to the frequency distribution, it is often useful to present the cumulative frequency distribution of a given variable. The cumulative frequency of x_i is defined as the sum of frequencies of values less than or equal to x_i. We denote it by F_i , and the proportional cumulative frequencies or cumulative relative frequency by

$$P_i = \frac{F_i}{n} \qquad (i = 1, 2, \ldots\ldots, k)$$

A table of proportional cumulative frequency distribution could be represented as follows:

Value	Relative frequency (p)	Cumulative relative frequency (P)
x_1	p_1	$P_1=p_1$
x_2	p_2	$P_2=p_1+p_2$
.	.	.
.	.	.
.	.	.
x_k	p_k	$P_k=p_1+p_2+\ldots+p_k=1$
Total	*1*	

2.2.2 Frequency distributions of continuous random variables

For the case of a continuous random variable, we partition the possible range of variation of the observed variable into k subintervals. Generally speaking, if the possible range of X is between L and H, we specify numbers $b_0, b_1, b_2, \cdots, b_k$ such that $L = b_0 < b_1 < b_2 < \cdots < b_{k-1} < b_k = H$. The values $b_0, b_1, \cdots, b_k$ are called the limits of the k subintervals. We then classify the X values into the interval (b_{i-1}, b_i) if $b_{i-1} < X \leq b_i$ $(i = 1, \cdots, k)$. (If $X = b_0$, we assign it to the first subinterval.) Subintervals are also called bins, classes or class-intervals.

In order to construct a frequency distribution we must consider the following two questions:

(i) How many sub-intervals should we choose? and

(ii) How large should the width of the subintervals be?

In general, it is difficult to give to these important questions exact answers which apply in all cases. However, the general recommendation is to use between 10 and 15 subintervals in large samples, and apply equal width subintervals. The frequency distribution is given then for the subintervals, where the mid-point of each subinterval provides a numerical representation for that interval. A typical frequency distribution table might look like the following:

Subinterval	Mid-Point	Frequency	Cumulative frequency
$b_0 - b_1$	$\bar{b}_1$	f_1	$F_1=f_1$
$b_1 - b_2$	$\bar{b}_2$	f_2	$F_2=f_1+f_2$
.	.	.	.
.	.	.	.
.	.	.	.
$b_{k-1} - bk$	$\bar{b}_k$	f_k	$F_k=f_1+f_2+\ldots+f_k=n$

2.3 Frequency distributions table

The easy method for construction of the frequency distribution table, it must conduct the following steps below:-

1. Find the Largest and Smallest values in the data.
2. Calculate the Range of data. (Range = Largest Value – Smallest Value)
3. Select the suitable Number of Classes according to the data under study (generally 5 – 15).
4. Find the Class Interval (C.I.):-

$$Class\ Interval = \frac{Range}{No.of\ Classes}$$

Note: (Class Interval coming near to the largest number depending on the accuracy of the data)

5. Determine the Class Limits (C.L.), that start with the lowest value of the sample data which is the lowest limit of the first class..
6. Calculate the class boundaries based on required accuracy.
7. Calulate the Class Mark (C.M.) of each Class Interval :-

$$Class\ Mark = \frac{(Upper\ Limit + Lower\ Limit)}{2}$$

8. Find the frequency of each class, the number of observations (data) corresponding to that class.

To separate one class from another, we use class boundaries (C.B.). If the data are given to the nearest integer, the class boundaries should be given to the nearest half. The class boundaries equal to the half way between the classes limits. These are called the class boundaries.

Example 2.1 :If we have the following frequency table. Complete the other elements of that table .

Score	*Frequency*
80-94	*8*
95-109	*14*
110-124	*24*
125-139	*16*
140-154	*13*

Solution :-

Largest value = 154
Smallest value = 80
Range = 154 – 80 = 74
Take No. of classes = 6
Class Interval = (74)/5 =14.8 → take C.I.= 15
Lower boundary of the first class = 80-0.5 = 79.5
Upper boundary of the first class = 79.5 + 15 = 94.5

No.	Class Limits	Class Boundaries	Class Mark	Frequency
1	***80-94***	79.5 – 94.5	87	***8***
2	***95-109***	94.5 – 109.5	102	***14***
3	***110-124***	109.5 – 124.5	117	***24***
4	***125-139***	124.5 – 139.5	132	***16***
5	***140-154***	139.5 – 154.5	147	***13***

Note: When forming a frequency distribution, the following general guidelines should be followed:

1. The number of classes should be between 5 and 15
2. Each data value must belong to one, and only one, class.
3. When possible, all classes should be of equal width (Class Interval) .

Example 2.2 :-Group the following data into classes and show its frequency table.

111	***120***	***127***	***129***	***130***	***145***	***145***	***150***	***153***	***155***	***160***	***161***
165	***167***	***170***	***171***	***174***	***175***	***177***	***179***	***180***	***180***	***185***	***185***
190	***195***	***195***	***201***	***210***	***220***	***224***	***225***	***230***	***245***	***248***	

Solution :-

Largest value = 248
Smallest value = 111
Range = 248 – 111 = 137
Take No. of classes = 6
Class Interval = (137)/6 =22.8333 → take C.I.= 23
Lower boundary of the first class = 111-0.5 = 110.5
Upper boundary of the first class = 110.5 + 23 = 133.5

No.	Class Limits	Class Boundaries	Class Mark	Tally	Frequency
1	111 – 133	110.5 – 133.5	122	~~llll~~	5
2	134 – 156	133.5 – 156.5	145	~~llll~~	5
3	157 – 179	156.5 – 179.5	168	~~llll~~ ~~llll~~	10
4	180 – 202	179.5 – 202.5	191	~~llll~~ lll	8
5	203 – 225	202.5 – 225.5	214	llll	4
6	226 – 248	225.5 – 248.5	237	lll	3

2.4 Relative frequency

The relative frequency of a class is obtained by dividing the frequency for a class by the sum of all the frequencies. The relative frequencies for the five classes in Example 2.1 are shown below. The sum of the relative frequencies will always equal one.

No.	Frequency (F)	Relative frequency (RF)
1	8	8/75 = 0.11
2	14	14/75 = 0.19
3	24	24/75 = 0.32
4	16	16/75 = 0.21
5	13	13/75 = 0.17
Total	**75**	**1**

2.5 Percentage

The percentage for a class is obtained by multiplying the relative frequency for that class by 100. The percentages for the six classes in Example 2.2 are shown in below. The sum of the percentages for all the categories will always equal 100 percent.

No.	Frequency (F)	Relative frequency (RF)	Percentage
1	5	5/35 = 0.14	0.14 × 100 = 14 %
2	5	5/35 = 0.14	0.14 × 100 = 14 %
3	10	10/35 = 0.29	0.29 × 100 = 29 %
4	8	8/35 = 0.23	0.23 × 100 = 23 %
5	4	4/35 = 0.11	0.11 × 100 = 11 %
6	3	3/35 = 0.09	0.09 × 100 = 9 %
Total	**35**	**1**	**100 %**

2.6 Cumulative frequency distribution

A cumulative frequency distribution gives the total number of values that fall below various class boundaries of a frequency distribution. The cumulative frequency distribution for the five classes in Example 2.1 are shown below

No.	Frequency (F)	Cumulative frequency (CF)
1	8	8
2	14	22
3	24	46
4	16	62
5	13	75

Total	75

There are different types of Cumulative frequency; which are shown below:

1. Ascending Cumulative Frequency :-

For any class, the ascending cumulative frequency is equal to the number of observations that are less than the upper boundary of that class.

2. Descending Cumulative Frequency :-

For any class, the descending cumulative frequency is equal to the number of observations that are greater than the lower boundary of that class.

3. Relative Cumulative Frequency:-

Relative Cumulative Frequency is obtained by dividing a cumulative frequency by the total number of observations in the data set. The cumulative frequencies for the frequency distribution given in Example 2.1 are shown below. Cumulative percentages are obtained by multiplying relative cumulative frequencies by 100.

No.	Frequency (F)	Relative frequency (RF)	Ascending cumulative frequency (ACF)	Relative Ascending cumulative frequency (RACF)	Descending cumulative frequency (DCF)	Relative Descending cumulative frequency (RDCF)
1	8	8/75 = 0.11	8	8/75 = 0.11	75	75/75 = 1
2	14	14/75 = 0.19	22	22/75 = 0.29	67	67/75 = 0.89
3	24	24/75 = 0.32	46	46/75 = 0.61	53	53/75 = 0.71
4	16	16/75 = 0.21	62	62/75 = 0.82	29	29/75 = 0.39
5	13	13/75 = 0.17	75	75/75 = 1	13	13/75 = 0.17
Total	**75**					

2.7 Graphical representation of data

If "a picture is worth a thousand words," then graphical techniques provide an excellent method to visualize the variability and other properties of a set of data. We proceed by assembling the data into graphs, scanning the details, and noting the important characteristics. There are numerous types of graphs. Histograms, frequency polygons, and cumulative frequency curves are given in this section.

2.7.1 Histogram

The data are divided into groups according to their magnitudes. The horizontal axis of the graph gives the magnitudes. Blocks are drawn to represent the groups, each of which has a different upper and lower limit. The area of a block is proportional to the number of occurrences in the group. The variability of the data is shown by the horizontal spread of the blocks, and the most common values are found in blocks with the largest areas. Other features such as the symmetry of the data or lack of it are also shown. The first step is to take into account the range *r* of the observations, that is, the difference between the largest and smallest values.

2.7.2 Frequency polygon

A frequency polygon is a useful characteristic tool to determine the distribution of a variable. It can be drawn by joining the midpoints of the tops of the rectangles of a histogram after extending the diagram by one class on both sides. We assume that equal class widths are used. If the ordinates of a histogram are divided by the total number of observations, then a relative frequency histogram is obtained. Thus, the ordinates for each class denote the probabilities bounded by 0 and 1, by which we simply mean the chances of occurrence. The resulting diagram is called the relative frequency polygon. The polygon ends on the horizontal axis at a distance equal to one half class interval after the upper boundary.

2.7.3 Cumulative frequency curves

If a cumulative sum is taken of the relative frequencies step by step from the smallest class to the largest, then the line joining the ordinates (cumulative relative frequencies) at the ends of the class boundaries forms a cumulative relative frequency or probability diagram. On the vertical axis of the graph, this line gives the probabilities of non exceedance of values shown on the horizontal axis. There are two types of cumulative frequency curves:

1. Ascending Cumulative Frequency Curve :-

It is drawn on a pair of perpendicular axes, just as the histogram and the frequency polygon, with the horizontal axis representing the values of the upper boundaries of the classes and the vertical axis representing the corresponding ascending cumulative frequencies of these classes.

2. Descending Cumulative Frequency Curve :-

It is drawn, just as the ascending curve, but the horizontal axis representing the values of the lower boundaries of the classes and the vertical axis representing the corresponding descending cumulative frequencies of these classes.

Example 2.3:-

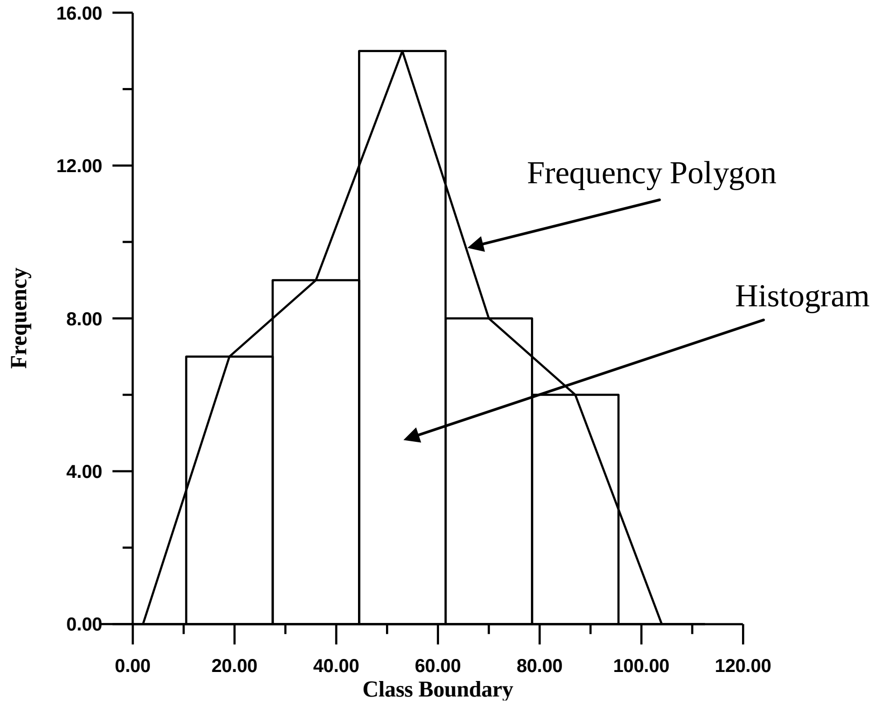
For the following classified data, draw the histogram, frequency polygon, and the cumulative frequency curves.

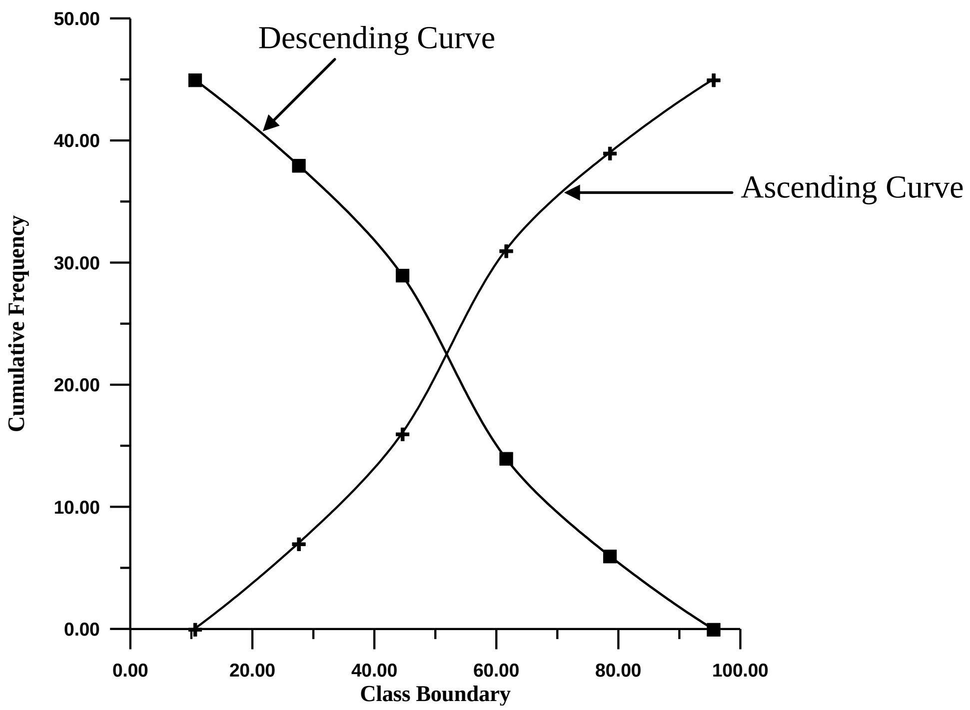
Class Limits	11 – 27	28 – 44	45 – 61	62 – 78	79 – 95
Frequency	7	9	15	8	6

Solution :-

No.	C.L	F	C.M.	C.B.	ACF	DCF
1	11 – 27	7	19	10.5 – 27.5	7	45
2	28 – 44	9	36	27.5 – 44.5	16	38
3	45 – 61	15	53	44.5 – 61.5	31	29

4	62 – 78	8	70	61.5 – 78.5	39	14
5	79 – 95	6	87	78.5– 95.5	45	6





CHAPTER THREE

MEASURES OF CENTRAL LOCATION AND DISPERSION

3.1 Introduction

When describing a numerical data set, it is common to report both a value that describes where the data distribution is centered along the number line and a value that describes how spread out the data distribution is.

Measures of center describe where the data distribution is located along the number line. A measure of center provides information about what is "typical."

Measures of Spread describe how much variability there is in a data distribution. A measure of spread provides information about how much individual values tend to differ from one another.

There is more than one way to measure center and spread in a data distribution.

3.2 Measures of center

A data set consisting of the observations for some variable is referred to as raw data or **ungrouped data**. Data presented in the form of a frequency distribution are called **grouped data**. The measures of central tendency discussed in this chapter will be described for both grouped and ungrouped data since both forms of data occur frequently. There are many different measures of central tendency. The five most widely used measures of central tendency are the mean, the median, the mode, geometric mean, and the harmonic mean. These measures are defined for both samples and populations.

3.2.1 The mean

The mean of a numerical data set is just the familiar arithmetic average—the sum of all of the observations in the data set divided by the total number of observations. It is denoted by the symbol $\bar{x}$. Notice that the value of the subscript on x doesn't tell us anything about how small or large the data value is.

Mean of ungrouped Data

The sum of $x_1, x_2, \dots, x_n$ can be written $x_1+x_2+\dots+x_n$, but this can be shortened by using the Greek letter Σ , which is used to denote summation. In particular Σx is used to denote the sum of all of the x values in a data set.

Mean is the most commonly used measure of central tendency. The mean of a set of N-numbers $x_1, x_2, x_3, \dots\dots, x_n$ is denoted by ($\bar{x}$) and is defined as :-

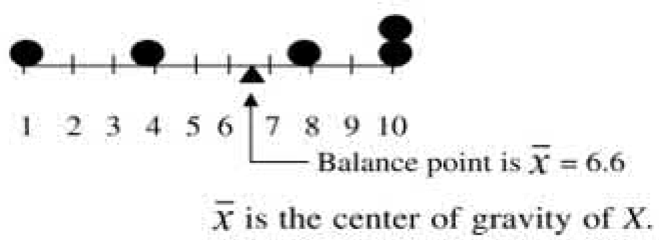
$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Example 3.1:-

For the following raw data on a variable *X*: 1, 4, 10, 8, 10. What is the mean of *X* or the value of $\bar{x}$?

Solution :-

$$\bar{x} = \frac{1+4+10+8+10}{5} = \frac{33}{5} = 6.6$$



Notes: The properties of the mean are as follows:

1. The mean always exists.
2. The mean is unique.

Mean of grouped (Classified) Data

For the samples that have frequency distribution, the mean of it can be obtained as:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i^{/} f_i}{\sum_{i=1}^{n} f_i}$$

where :-

$x_i^{/}$: is the class mark of class (i).

f_i : is the frequency of class (i).

n : the number of classes.

Example 3.2 :-

For the following classified data, find the mean.

Class Limits	84-86	87-89	90-92	93-95	96-98
Frequency	4	12	13	17	3

Solution :-

No.	Class Limits	Class Mark (x'_i)	Frequency (f_i)	$x'_i \cdot f_i$
1	84-86	85	4	340
2	87-89	88	12	1056
3	90-92	91	13	1183
4	93-95	94	17	1598
5	96-98	97	3	291
			Σ 49	Σ 4468

$\bar{x}$ = 4468/49 = 91.18

3.2.2 The median

When the values of a data set of size n were ordered from smallest to largest. If n is odd, the median is the value in position $(n+1)/2$; if n is even, it is the average of the values in positions $(n/2)$ and $(n/2)+1$. Thus the sample median of a set of three values is the second smallest; of a set of four values, it is the average of the second and third smallest.

Median of ungrouped Data

The sample median of n measurements $x_1, x_2, \ldots, x_n$ is the middle value when measurements are arranged from the smallest to largest. The median is the value that divides the data into two equal halves. In other words, 50% of the data lie below the median and 50% lie above it. If n is an odd number, there is a unique middle value and it is the median. If n is an even number, there are two middle values and the median is defined as their average value.

Example 3.3:-

For the following raw data on a variable X: 1, 8, 5, 10, 15, 2 and variable Y: 8, 7, 12, 8, 6, 2, 4, 3, 5, 11, 10. What is the median of X and Y?

Solution :-

For variable X:-

Arranging the data in an increasing sequence gives, 1, 2, 5, 8, 10, 15

Median = (5+8)/2=6.5

For variable Y:-

Arranging the data in an increasing sequence yields, 2, 3, 4, 5, 6, 7, 8, 8, 10, 11, 12

Median = 7

Notes: The properties of the median are as follows:

1. The median may or may not equal the mean.
2. The median always exists.
3. The median is unique.

Median of Classified Data

The median of classified data can be founded by determined first the smallest class that has an ascending cumulative frequency (*ACF*) greater than the half of summation of all frequencies. This class called the median class. Thus, the median can be determined as follows :-

$$Median = B_L + \left[\frac{\frac{n}{2} - (\sum f)_L}{f_{median\,class}} \right] \times C$$

where :-

B_L : lower boundary of the median class.

n : number of observations (data).

$(\sum f)_L$: sum of all class frequencies lower than the median class.

$f_{median\ class}$: frequency of median class.

C : the median class interval.

Example 3.4 :-

For the following classified data, find the median.

Class Limits	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79
Frequency	6	12	15	13	8

Solution :-

No.	Class Limits	Frequency	Class Boundaries	ACF
1	30 – 39	6	29.5 – 39.5	6
2	40 – 49	12	39.5 – 49.5	18
3	50 – 59	15	49.5 – 59.5	33
4	60 – 69	13	59.5– 69.5	46
5	70 – 79	8	69.5 – 79.5	54

n/2 = 54/2=27

Then the median class is (49.5 – 59.5) which is the third class.

B_L= 49.5

n = 54

$(\sum f)_L$= 18

$f_{median\ class}$= 15

C = 10

$$Median = 49.5 + \left[\frac{\frac{54}{2} - 18}{15} \right] \times 10 = 55.5$$

3.2.3 The mode

The mode is the value in a data set that occurs the most often. If no such value exists, we say that the data set has no mode. If two or more such values exist, we say the data set is multimodal. There is no symbol that is used to represent the mode.

Mode of ungrouped Data

The mode of X is the value that occurs with the highest frequency; it is the most common or most probable value of X.

Example 3.5:- For the following raw data on a variable X: 1, 2, 4, 4, 5 , variable Y: 1, 2, 4, 4, 5, 6, 6, 9 and variable Z: 1, 2, 3, 7, 9, 11. What is the mode of X ,Y,and Z ?

Solution :-
For variable X:-
Mode = 4 (its frequency is two)
There is only one mode.

For variable Y:-
Modes = 4 and 6 (Y is thus a multimodal variable)
There are two modes.

For variable Z:-
Mode = no mode (no one value appears more often than any other).
There is no mode.

Notes: The properties of the mode are as follows:

1. The mode may or may not equal the mean and median.
2. The mode may not exist.
3. If the mode exists, it may not be unique.

Mode of Classified Data

In classified data, the mode can be calculated from the following formula :-

$$Mode = L_w + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2} \right] \times C$$

where :-
L_w : lower boundary of the modal class.
Δ_1 : excess of modal class frequency over frequency of the next lower class.
Δ_2 : excess of modal class frequency over frequency of the next higher class.
C : the modal class interval.

Modal Class: *The class which has the maximum number of observations.*

Example 3.6:-

Find the mode for the following classified data?

Class Limits	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79
Frequency	6	12	15	13	8

Solution :-

Modal class is (50 – 59) which is the third class.

$L_w = 49.5$

$\Delta_1 = 15-12 = 3$

$\Delta_2 = 15-13 = 2$

$C = 10$

:. Mode = 49.5 + {3 / (3+2)} x 10 = 55.5

3.2.4 The geometric mean

Geometric mean is a kind of average of a set of numbers that is different from the arithmetic average (mean). The geometric mean is well defined only for sets of positive real numbers. This is calculated by multiplying all the numbers (call the number of numbers *n*), and taking the *nth* root of the total.

$$\text{Geometric Mean } (G) = [x_1 \times x_2 \times x_3 \times \ldots\ldots \times x_n]^{1/n} = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots\ldots\ldots \times x_n}$$

G can be computed by logarithms such as:

$$\log G = \frac{\sum_{i=1}^{n} \log x_i}{n} \qquad i = 1, 2, 3, \ldots\ldots, n$$

Example 3.7:- Find the geometric mean for the following data 1,2,3,4,5.

Solution :-

n=5, total number of values.

Then G = $[x_1 \times x_2 \times x_3 \times \ldots\ldots \times x_n]^{1/n} = [1\times 2\times 3\times 4\times 5]^{1/5} = 2.60517$

3.2.5 The harmonic mean

The harmonic mean of a set of *n*-numbers $x_1, x_2, x_3, \ldots\ldots, x_n$, can be evaluated as follows :-

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \ldots\ldots + \frac{1}{x_n}}$$

Example 3.8:- Find the harmonic mean for the following data 1,2,3,4,5.

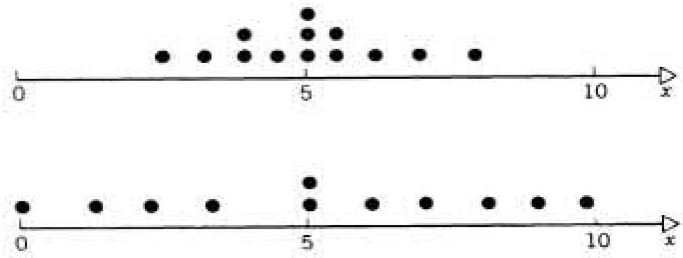
Solution :-

n=5, total number of values

Then, H= $\frac{5}{\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}} = 2.189781$

3.3 Measures of dispersion

In addition to locating the center of the data, another important aspect of a descriptive study of data is numerically measuring the extent of variation around the center. Two data sets may show similar positions of center but may be remarkably different with respect to variability. For example, the following figure that show dot diagrams with similar center values but different variations.



There are several forms that are used as measures of variation such as range, mean absolute deviation, variance, and standard deviation.

3.3.1 The range (R)

We previously defined the range of a variable X as *range (R)= max* x_i - *min* x_i. Clearly, the range is a measure of the total spread of the data.

3.3.2 Mean absolute deviation (M.A.D.)

Mean absolute deviation is used to find how consistent a set of data. It describes the average distance from the mean for the numbers in the data set. M.A.D. for ungrouped data, can be calculated as;

$$M.A.D. = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

For grouped (classified) data, the M.A.D. can be calculated by the following formula:-

$$M.A.D. = \frac{\sum_{i=1}^{c} f_i |x_i^{/} - \bar{x}|}{n}$$

where :-

c : number of classes.

n : number of observation.

$x_i^{/}$ the class mark in grouped data.

f_i the frequency of class i in grouped data.

$x_i^{/}, f_i$: the mark and frequency of class i.

$\bar{x}$: the mean of (grouped or ungrouped) data.

Example 3.9:- Find the mean absolute deviation for the following data 1,2,4,5.

Solution :- *Mean = (1+2+4+5)/4=12/4=3*

M.A.D. =[(2+1+1+2)/4]=6/4=1.5

3.3.3 Variance (S^2) and standard deviation (S)

Variance is constructed by adding the squared deviations and dividing the total by the number of observations minus one. If the dataset contains *n* measurements labeled, $x_1, x_2, x_3, \ldots, x_n$, then the variance is defined as :

$$S^2 = \frac{\text{sum of squared deviations}}{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

Whereas, the variance for (grouped) classified data can be calculated as :-

$$S^2 = \frac{\sum_{i=1}^{c}(x_i^{/} - \bar{x})^2}{n-1}$$

The standard deviation is the square root of the variance.

$$S = \sqrt{\text{Variance}} = \sqrt{S^2}$$

Example 3.10:- Find the Variance and standard deviation for the following data 3,4,6,7,10

Solution :- Mean ($\bar{x}$) = (3+4+6+7+10)/5=6

Variance (S^2) = $[(-3)^2+(-2)^2+0^2+1^2+4^2)/4=7.5$

Standard deviation (S) = $\sqrt{7.5}$ =*2.738*

3.3.4 Coefficient of Variance (C.V.)

The standard formulation of the Coefficient of Variance, is the ratio of the standard deviation to the mean. It can be calculated as follows:

$$C.V. = \frac{S}{\bar{x}} \times 100$$

where :-

S : standard deviation for (grouped or ungrouped) data.

$\bar{x}$: the mean of (grouped or ungrouped) data.

Example 3.11 :- Find the coefficient of variance for data in example 3.10 above?

Solution :- Mean ($\bar{x}$) = (3+4+6+7+10)/5=6

Variance (S^2) = $[(-3)^2+(-2)^2+0^2+1^2+4^2)/4=7.5$

Standard deviation (S) = $\sqrt{7.5}$ =*2.738*

Coefficient of variance (*C.V.*) $= \frac{S}{\bar{x}} \times 100 = \frac{2.738}{6} \times 100 = 45.6333\,\%$

Example 3.12 :- Find the mean absolute deviation (M.A.D.), standard deviation (S), and coefficient of variance (C.V.) for the grouped (classified) data below?

Class Limits	60 – 62	63 – 65	66 – 68	69 – 71	72 – 74
Frequency	5	18	42	27	8

Solution :-

Classes	x'_i	f_i	$x'_i f_i$	$\lvert x'_i - \bar{x}\rvert$	$f_i \lvert x'_i - \bar{x}\rvert$	$(x'_i - \bar{x})^2$
60 – 62	61	5	305	6.45	32.225	41.6025
63 – 65	64	18	1152	3.45	62.1	11.9025
66 – 68	67	42	2814	0.45	18.9	0.2025
69 – 71	70	27	1890	2.55	68.85	6.5025
72 – 74	73	8	584	5.55	44.4	30.8025
		Σ 100	Σ 6745		Σ 226.5	Σ 91.0125

$$\bar{x} = \frac{\sum_{i=1}^{n} x'_i f_i}{\sum_{i=1}^{n} f_i} = 6745/100 = 67.45$$

$$M.A.D. = \frac{\sum_{i=1}^{c} f_i \lvert x'_i - \bar{x}\rvert}{n} = 226.5/100 = 2.265$$

$$S = \sqrt{\frac{\sum_{i=1}^{c} (x'_i - \bar{x})^2}{n-1}} = \sqrt{\frac{91.0125}{99}} = 0.9193$$

$$C.V. = \frac{S}{\bar{x}} \times 100 = (0.9193/67.45) * 100 = 1.3629\ \%$$

Homework A: For the following table that represent the statistical data for ages and its frequencies. Find the following: the mean, median, mode, mean absolute deviation, standard deviation, and Coefficient of Variance.

Age	5 – 14	15 – 24	25 – 34	35 – 44	45 – 54
Frequency	750	2005	1950	195	100

Homework B: As shown in the table below that gives the age distribution of individuals starting new companies. Find the mean, median, mean absolute deviation, standard deviation, and Coefficient of Variance for this distribution.

Age	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69
Frequency	11	25	14	7	3

CHAPTER FOUR

THE PROBABILITY

4.1 Introduction

Probability theory is a mathematical theory to describe and analyze situations where randomness or uncertainty is present. An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a ***random experiment***. After the experiment is over, we call the result an ***outcome***. For any given experiment, there is a set of possible outcomes, and can be state as the set of all possible outcomes in a random experiment is called the ***sample space***, denoted ***S***. The subsets of the sample space of random experiment called ***event***.

4.2 Probability theory

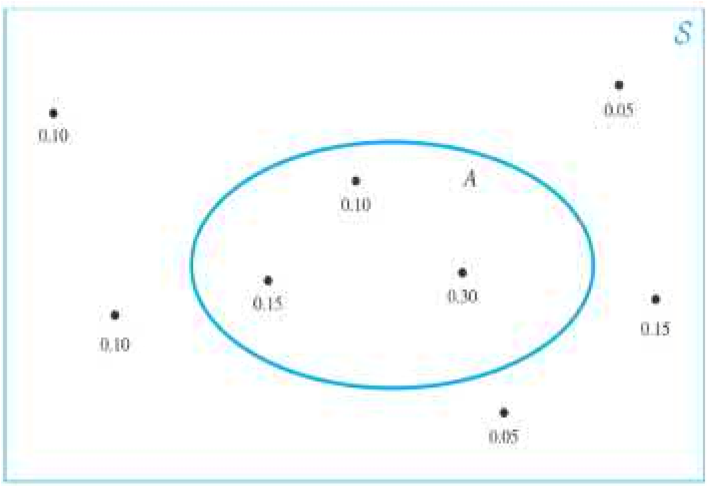
There is a finite number **n** of possibilities (often called outcomes) and each of them has the same probability **1/n**. A collection **A** (event) of **k** outcomes with **k ≤ n** is called an event and its probability **P(A)** is calculated as **k/n**:

$$P(A) = \frac{k}{n} = \frac{\text{the number of outcomes in A}}{\text{the total number of outcomes}}$$

An empty collection has probability zero and the whole collection one. Also, the probability of non-occurrence of event (A) is referred :

$$q(A) = P(not\ A) = \frac{n-k}{n} = 1 - \frac{k}{n} = 1 - P(A)$$

Example 4.1:- as shown in the figure below, a sample space S consisting of eight outcomes, each of which is labeled with a probability value. Find P(A) and q(A)?



Solution :-

P(A)= 0.10+0.15+0.30=0.55

q(A)=0.10+0.10+0.05+0.15+0.05=0.45

or q(A)=1-P(A)=1-0.55=0.45 because P(A)+q(A)=1, which is general rule.

Example 4.2:- ***As shown in the table below that gives the age distribution of individuals starting new companies. Find the following probabilities: 1-Ages (30-39). 2-Age=45. 3-Age>50. 4-Age<40.***

Age	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69
Frequency	11	25	14	7	3

Solution :-

1- The probability of ages between (30-39) = 25/60= 0.41667
2- The probability of age (45) = probability of the class/class interval
   = (14/60)/10=0.02334
3- The probability of age > 50 = probability of the classes (50-59) and (60-69)
   = (7/60) + (3/60) = (10/60) = 0.16667
4- The probability of age < 40 = probability of the classes (20-29) and (30-39)
   = (11/60) + (25/60) = (36/60) = 0.6

4.3 Combinations of Events

For two events **A** and **B**, in addition to the consideration of the probability of event **A** occurring and the probability of event **B** occurring, it is often important to consider other probabilities such as the probability of both events occurring simultaneously. Other quantities of interest may be the probability that neither event **A** nor event **B** occurs, the probability that at least one of the two events occurs, or the probability that event **A** occurs, but event **B** does not.

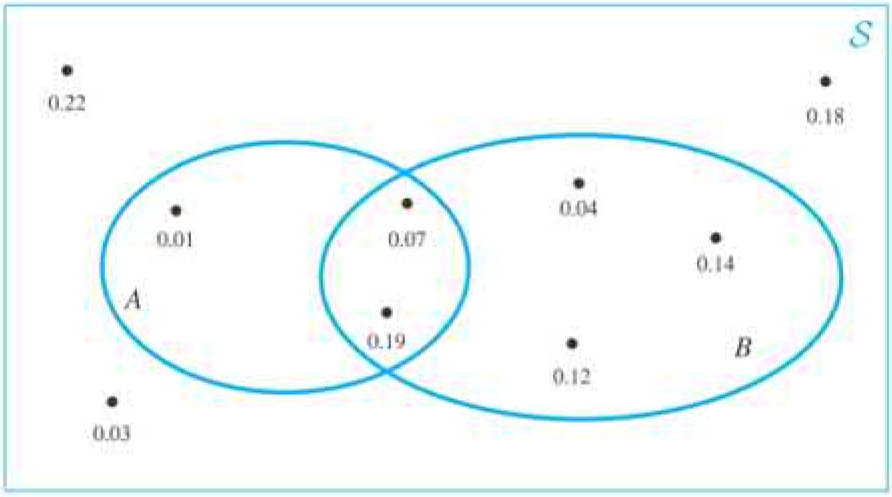
4.3.1 Intersections of Events

Consider first the calculation of the probability that both events occur simultaneously. This can be done by defining a new event to consist of the outcomes that are in both event **A** and event **B**.

The event **A∩B** is the intersection of the events **A** and **B** and consists of the outcomes that are contained within both events **A** and **B**. The probability of this event, **P(A∩B)**, is the probability that both events **A** and **B** occur simultaneously.

Example 4.3:- ***Figure below shows a sample space S that consists of nine outcomes. Event A consists of three outcomes, and event B consists of five outcomes. Find* P(A∩B), P(A∩A′), P(A′∩B), and P(A∩B′).**

Solution :-



P(A) = 0.01 + 0.07 + 0.19 = 0.27

P(B) = 0.07 + 0.19 + 0.04 + 0.14 + 0.12 = 0.56

P(A ∩ B) = 0.07 + 0.19 = 0.26 (there are two outcomes that contained within both events A and B), Which is the probability that both events A and B occur simultaneously.

Event **A´**, the ***complement*** of the event **A**, is the event consisting of the six outcomes that are not in event **A**. Notice that there are obviously no outcomes in **A ∩ A´**, and this is written as **A ∩ A´ = ∅.** Where **∅** is referred to as the "empty set," a set that does not contain anything.
P(A ∩ A´) = P(∅) = 0

And it is impossible for the event **A** to occur at the same time as its complement (**A´**). A more interesting event is the event **A´ ∩ B.** This event consists of the three outcomes that are contained within event **B** but that are not contained within event **A**. It has a probability of

P(A´ ∩ B) = 0.04 + 0.14 + 0.12 = 0.30 (which is the probability that event **B** occurs but event **A** does not occur.)

Similarly, event **A ∩ B´**, which has a probability of

P(A ∩ B´) = 0.01 (This is the probability that event **A** occurs but event **B** does not.)

Notice that
P(A ∩ B) + P(A ∩ B´) = 0.26 + 0.01 = 0.27 = P(A)
and similarly that

$\mathbf{P(A \cap B) + P(A' \cap B) = 0.26 + 0.30 = 0.56 = P(B)}$

The following two equalities hold in general for all events **A** and **B**:

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

Two events **A** and **B** that have no outcomes in common are said to be ***mutually exclusive events***. In this case $\mathbf{A \cap B = \emptyset}$ and $\mathbf{P(A \cap B) = 0}$.

Some other simple results concerning the intersections of events are as follows:

$\mathbf{A \cap B = B \cap A}$	$\mathbf{A \cap A = A}$	$\mathbf{A \cap S = A}$
$\mathbf{A \cap \emptyset = \emptyset}$	$\mathbf{A \cap A' = \emptyset}$	$\mathbf{A \cap (B \cap C) = (A \cap B) \cap C}$

4.3.2 Unions of Events

The event $\mathbf{A \cup B}$ is the union of events **A** and **B** and consists of the outcomes that are contained within at least one of the events **A** and **B**. The probability of this event, $\mathbf{P(A \cup B)}$, is the probability that at least one of the events **A** and **B** occurs.

Notice that the outcomes in the event $\mathbf{A \cup B}$ can be classified into three kinds. They are

1- in event **A**, but not in event **B**
2- in event **B**, but not in event **A**
3- in both events **A** and **B**

Since the probability of $\mathbf{A \cup B}$ is obtained as the sum of the probability values of the outcomes within (mutually exclusive) events, the following result is obtained:

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

This equality can be presented in another form using the relationships

$$\mathbf{P(A \cap B') = P(A) - P(A \cap B)} \quad \text{and} \quad \mathbf{P(A' \cap B) = P(B) - P(A \cap B)}$$

Substituting in these expressions for $\mathbf{P(A \cap B')}$ and $\mathbf{P(A' \cap B)}$ gives following result:

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

This equality has the intuitive interpretation that the probability of at least one of the events **A** and **B** occurring can be obtained by adding the probabilities of the two events **A** and **B** and then subtracting the probability that both the events occur simultaneously.

If the events **A** and **B** are mutually exclusive so that $\mathbf{P(A \cap B)=0}$,
Then $\mathbf{P(A \cup B)=P(A) + P(B)}$

Example 4.4:- The sample space of nine outcomes illustrated in Example 4.3 can be used to demonstrate some general relationships between unions and intersections of events.

Solution :-

The event $\mathbf{(A \cup B)}$ consists of the six outcomes, and it has a probability of

$$\mathbf{P(A \cup B) = 0.01 + 0.07 + 0.19 + 0.04 + 0.14 + 0.12 = 0.57}$$

The event $\mathbf{(A \cup B)'}$, which is the complement of the union of the events **A** and **B**, consists of the three outcomes that are neither in event **A** nor in event **B**. It has a probability of

$$\mathbf{P((A \cup B)') = 0.03 + 0.22 + 0.18 = 0.43 = 1 - P(A \cup B)}$$

Notice that the event $\mathbf{(A \cup B)'}$ can also be written as $\mathbf{A' \cap B'}$ since it consists of those outcomes that are simultaneously neither in event **A** nor in event **B**. This is a general result:

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

Furthermore, the event $\mathbf{A' \cup B'}$ consists of the seven outcomes, and it has a probability of

$$\mathbf{P(A' \cup B') = 0.01 + 0.03 + 0.22 + 0.18 + 0.12 + 0.14 + 0.04 = 0.74}$$

However, this event can also be written as $\mathbf{(A \cap B)'}$ since it consists of the outcomes that are in the complement of the intersection of sets A and B. Hence, its probability could have been calculated by

$$\mathbf{P(A' \cup B') = P((A \cap B)') = 1 - P(A \cap B) = 1 - 0.26 = 0.74}$$

Again, this is a general result:

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

Finally, if event **A** is contained within event **B**, $\mathbf{A \subset B}$, then clearly $\mathbf{A \cup B = B}$.
Some other simple results concerning the unions of events are as follows:

$\mathbf{A \cup B = B \cup A}$ $\quad$ $\mathbf{A \cup A = A}$ $\quad$ $\mathbf{A \cup S = S}$

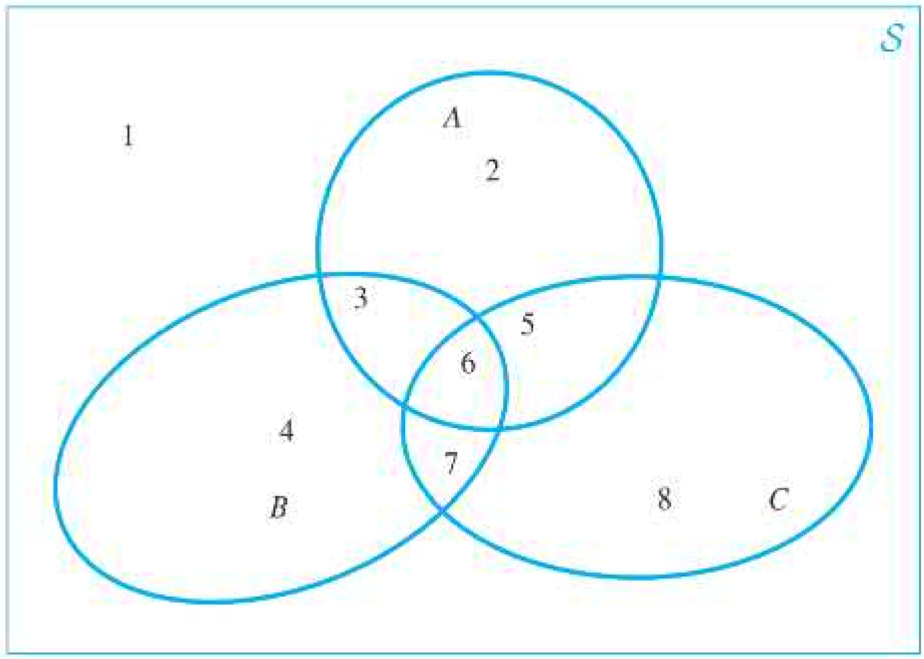
$\mathbf{A \cup \emptyset = A}$ $\quad$ $\mathbf{A \cup A' = S}$ $\quad$ $\mathbf{A \cup (B \cup C) = (A \cup B) \cup C}$

4.3.3 Combinations of Three or More Events

The probability of the union of three events **A**, **B**, and **C** is the sum of the probability values of the simple outcomes that are contained within at least one of the three events. It can also be calculated from the expression

$$P(A\cup B\cup C) = [P(A)+P(B)+P(C)]-[P(A\cap B)+P(A\cap C)+P(B\cap C)]+P(A\cap B\cap C)$$

Example 4.5:- Three events decompose the sample space into eight regions as shown in the figure below. Find $P(A\cup B\cup C)$ **?**



Solution :-

The event **A**, is composed of the regions 2, 3, 5, and 6, and the event $A\cap B$ is composed of the regions 3 and 6. The event $A\cap B\cap C$, the intersection of the events **A**, **B**, and **C**, consists of the outcomes that are simultaneously contained within all three events **A**, **B**, and **C** , it corresponds to region 6. The event $A\cup B\cup C$, the union of the events **A**, **B**, and **C**, consists of the outcomes that are in at least one of the three events **A**, **B**, and **C** , it corresponds to all of the regions except for region 1. Hence region 1 can be referred to as $(A\cup B\cup C)'$ since it is the complement of the event $A\cup B\cup C$.

In general, care must be taken to avoid ambiguities when specifying combinations of three or more events. For example, the expression

$A\cup B\cap C$

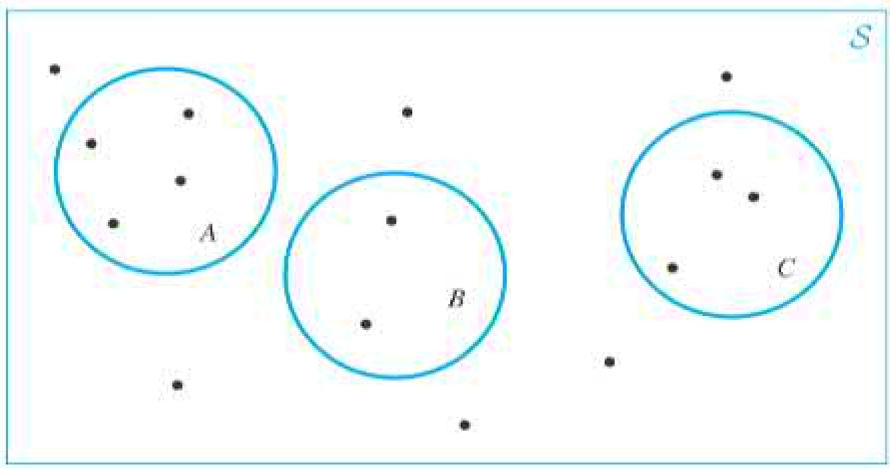
since the two events are different

$A\cup (B\cap C)$ and $(A\cup B)\cap C$

The event $\mathbf{B \cap C}$ is composed of regions 6 and 7, so $\mathbf{A \cup (B \cap C)}$ is composed of regions 2, 3, 5, 6, and 7. In contrast, the event $\mathbf{A \cup B}$ is composed of regions 2, 3, 4, 5, 6, and 7, so $\mathbf{(A \cup B) \cap C}$ is composed of just regions 5, 6, and 7.

The required probability, $\mathbf{P(A \cup B \cup C)}$, is the sum of the probability values of the outcomes in regions 2, 3, 4, 5, 6, 7, and 8. However, the sum of the probabilities **P(A)**, **P(B)**, and **P(C)** counts regions 3, 5, and 7 twice, and region 6 three times. Subtracting the probabilities $\mathbf{P(A \cap B)}$, $\mathbf{P(A \cap C)}$, and $\mathbf{P(B \cap C)}$ removes the double counting of regions 3, 5, and 7 but also subtracts the probability of region 6 three times. The expression is then completed by adding back on $\mathbf{P(A \cap B \cap C)}$, the probability of region 6.

The figure below illustrates three events **A**, **B**, and **C** that are mutually exclusive because no two events have any outcomes in common. In this case,
$\mathbf{P(A \cup B \cup C) = P(A) + P(B) + P(C)}$



In general

The union and intersection operations satisfy the following laws: For any subsets **A**, **B**, **C** of **S**, we have:

Commutative Law: $\mathbf{A \cup B = B \cup A}$,
$\mathbf{A \cap B = B \cap A}$

Associative Law: $\mathbf{(A \cup B) \cup C = A \cup (B \cup C)}$,
$\mathbf{(A \cap B) \cap C = A \cap (B \cap C)}$

Distributive Law: $\mathbf{(A \cup B) \cap C = (A \cap C) \cup (B \cap C)}$,
$\mathbf{A \cup (B \cap C) = (A \cup B) \cap (A \cup C)}$

4.4 Combinations Rule

A second method of determining the number of sample points for an experiment is to use combinatorial mathematics. This branch of mathematics is concerned with developing counting rules for given situations. For example, there is a simple rule for finding the number of different samples of 5 student selected from 1,000. This rule, called the combinations rule, is given below.

Suppose a sample of n elements is to be drawn without replacement from a set of N elements. Then the number of different samples possible is denoted by the following:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

Where: -
n! = n (n-1) (n-2) .... (3)(2)(1)
and similarly for N*!* and *(N – n)!*

for example, *5! =5.4.3.2.1*. [Note: the quantity *0!* is defined to be equal to *1*]

Example 4.6: - Consider the task of choosing 2 students from 4 students. Use the combinations counting rule to determine how many different selections can be made.

Solution: -
For this example, N = *4*, *n* = *2*, and

$$\binom{4}{2} = \frac{4!}{2!\,(4-2)!} = \frac{4!}{2!\,2!} = \frac{4.3.2.1}{(2.1)(2.1)} = 6$$

Classwork: - Compute the number of ways you can select n elements from N elements for each of the following:

a. n = 2, N = 5
b. n = 3, N = 6
c. n = 5, N = 20

4.5 Multiplicative Rule

If there are k sets of elements, n_1 in the first set, n_2 in the second set, ….. and n_k in the k^{th} set. Suppose it wish to form a sample of k elements by **taking one element from each** of the k sets. Then the number of different samples that can be formed is the product $(n_1\, n_2\, n_3 \ldots n_k)$

Example 4.7: - There are 20 engineers for three different positions: E_1, E_2, and E_3. How many different ways that the positions can be fill?

Solution: -

There are $k = 3$ sets of elements, corresponding to

Set 1: engineers available to fill position E_1

Set 2: engineers remaining (after filling E_1) that are available to fill E_2

Set 3: engineers remaining (after filling E_1 and E_2) that are available to fill E_3 .

The numbers of elements in the sets are $n_1 = 20$, $n_2 = 19$, and $n_3 = 18$.

Therefore, the number of different ways of filling the three positions is given by the multiplicative rule as $n_1\, n_2\, n_3 = (20)\,(19)\,(18) = 6840$.

4.6 Permutations Rule

When there are a set of N different elements, it wish to select n elements from the N and arrange them within n positions. The number of different permutations of the N elements taken n at a time is denoted by P_n^N and is equal to

$$P_n^N = N\,(N-1)(N-2)\ldots\ldots(N-n-1) = \frac{N!}{(N-n)!}$$

Example 4.8: - Consider the task of choosing 2 students from 4 students. Use the Permutations rule to determine how many different ways.

Solution: -

For this example, N = 4, $n = 2$, and

$$P_n^N = P_2^4 = \frac{4!}{(4-2)!} = \frac{4.3.2.1}{2.1} = 12$$

4.7 Partitions Rule

Suppose we wish to partition a single set of N different elements into k sets, with the first set containing n_1 elements, the second containing n_2 elements, …,and the k^{th} set containing n_k elements. Then the number of different partitions is

$$\frac{N!}{n_1!\; n_2!\; \ldots.. n_k!}$$

Where $n_1 + n_2 + \ldots.. + n_k = N$

Example 4.9: - If you have 12 construction workers and you wish to put 3 workers to site 1, 4 workers to site 2, and 5 workers to site 3. In how many different ways can you make this assignment?

Solution: -

For this example, $k = 3$ (corresponding to the $k = 3$ different sites),

N = 12 , n_1 = 3, n_2 = 4, and n_3 = 5. Then the number of different ways to assign the workers to the sites is

$$\frac{N\,!}{n_1!\;n_2!\;n_3!}=\frac{12!}{3!\;4!\;5!}=\frac{12.11.10.......3.2.1}{(3.2.1)(4.3.2.1)(5.4.3.2.1)}=27720$$

Summary of Counting Rules

1. **Multiplicative rule**. If you are drawing one element from each of k sets of elements , where the sizes of the sets are n_1 , n_2 , ... , n_k , then the number of different results is

$$(n_1\, n_2\, n_3 \,....n_k)$$

2. **Permutations rule**. If you are drawing n elements from a set of N elements and arranging the n elements in a distinct order , then the number of different results is

$$P_n^N=\frac{N\,!}{(N-n)!}$$

3. **Partitions rule**. If you are partitioning the elements of a set of N elements into k groups consisting of n_1 ,n_2 , ... ,n_k elements ($n_1+\ldots+n_k= N$), then the number of different results is

$$\frac{N\,!}{n_1!\;n_2!\;\;.....n_k\,!}$$

4.8 Conditional Probability

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**”. It can be defined by **P(B\A) = P(A∩B)/ P(A)** provided **P(A)>0**.

Example 4.10: - For the table shown below which represent the number of persons for a small city. Find the conditional probability for chosen a man and the one that chosen is educated?

	Educated	*Uneducated*	*Total*
Male	*460*	*40*	*500*
Female	*140*	*260*	*400*
Total	*600*	*300*	*900*

Solution: -

For this example, **M**: a man is chosen, **E**: the one chosen is educated

P(M\E) = P(E∩M)/ P(E)=(460/900)/(600/900)= 0.766666

Example 4.11: - The following table summarizes the analysis of samples of galvanized steel for coating weight and surface roughness. The results from 100 samples are shown.

		Coating Weight	
		High	**Low**
Surface Roughness	**High**	*70*	*9*
	Low	*16*	*5*

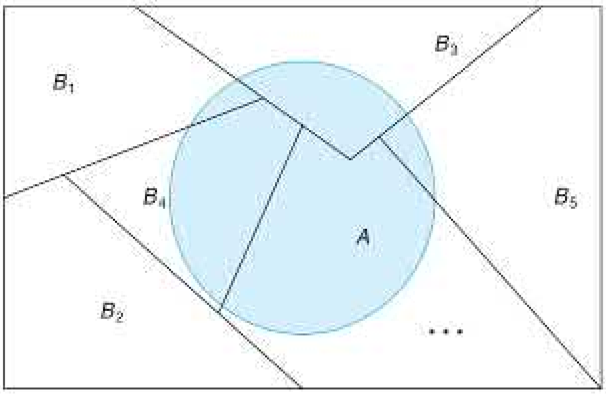
Let A denote the event that a sample has high coating weight, and let B denote the event that a sample has high surface roughness. Determine the following probabilities:
(a) P(A) (b) P(B) (c) P (A \B) (d) P(B \A)

Solution: -

(a) P(A)=(70+16)/100=86/100
(b) P(B)=(70+9)/100=79/100
(c) $P(A \backslash B)=P(B \cap A)/P(B)=(70/100)/(79/100)=70/79$
(d) $P(B \backslash A)=P(A \cap B)/P(A)=(70/100)/(86/100)=70/86$

4.9 Total probability and Bayes' theorems

Sometimes the probability of an event **A** cannot be determined directly. However, its occurrence is accompanied by the occurrence of other events $\mathbf{B_i}$, $i = 1, 2, \dots, n$, such that the probability of **A** will depend on which of the events $\mathbf{B_i}$ has occurred. In such a case, the probability of **A** will be an expected probability, that is, the average probability weighted by those of $\mathbf{B_i}$. This problem can be approached by using the theorem of total probability, which can be derived by the definition of conditional probability. When the sample space is partitioned into **k** subsets is covered by the following theorem, sometimes called the theorem of total probability or the rule of elimination (see figure).



If the events $\mathbf{B_1 , B_2 , \dots , B_k}$ constitute a partition of the sample space **S** such that $\mathbf{P(B_i) \neq 0}$ for i = 1, 2, . . . , k, then for any event **A** of **S**,

$$\mathbf{P(A)=P(B_1)P(A\backslash B_1)+P(B_2)P(A\backslash B_2)+\dots\dots+P(B_k)P(A\backslash B_k)} \text{ , Or}$$

$$P(A)=\sum_{i=1}^{k} P(B_i \cap A)=\sum_{i=1}^{k} P(B_i)P(A \backslash B_i)$$

Example 4.12: - Plant have 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 BAD. Now, suppose that a finished product is randomly selected. What is the probability that it is BAD?

Solution: -

Consider the following events:

A : the product is BAD,

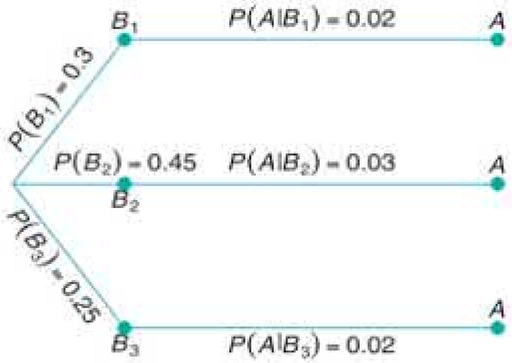
$\mathbf{B_1}$: the product is made by machine $\mathbf{B_1}$,

$\mathbf{B_2}$: the product is made by machine $\mathbf{B_2}$,

$\mathbf{B_3}$: the product is made by machine $\mathbf{B_3}$.

Applying the rule of elimination, and make a tree diagram,

$P(A)= P(B_1)P(A\backslash B_1) + P(B_2)P(A\backslash B_2) + P(B_3)P(A\backslash B_3)$



Referring to the tree diagram, the three branches give the probabilities

$P(B_1)P(A\backslash B_1) = (0.3)(0.02) = 0.006,$

$P(B_2)P(A\backslash B_2) = (0.45)(0.03) = 0.0135,$

$P(B_3)P(A\backslash B_3) = (0.25)(0.02) = 0.005,$

and hence

$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245.$

Bayesian statistics is a collection of tools that is used in a special form of statistical inference which applies in the analysis of experimental data in many practical situations in science and engineering. Bayes' rule is one of the most important rules in probability theory.

Bayes's rule, can be applied when an observed event **A** occurs with any one of several mutually exclusive and exhaustive events, $\mathbf{B_1}$,$\mathbf{B_2}$, … , $\mathbf{B_k}$. The formula for finding the appropriate conditional probabilities is given below

Bayes's Rule

Given **k** mutually exclusive and exhaustive events, $\mathbf{B_1}$, $\mathbf{B_2}$, … ,$\mathbf{B_k}$ such that $\mathbf{P(B_1)+P(B_2)+\ldots+P(B_k)=1}$, and given an observed event **A** , it follows that

$\mathbf{P(B_i\backslash A)=P(B_i \cap A)/P(A)}$

Where: $\mathbf{P(A)=P(B_1)P(A\backslash B_1)+P(B_2)P(A\backslash B_2)+\ldots\ldots+P(B_k)P(A\backslash B_k)}$

Example 4.13: - For example 4.12, if a product was chosen randomly and found to be BAD, what is the probability that it was made by machine B_3 ?

Solution: -

Using Bayes' rule,

$P(B_3 \backslash A)=P(B_3 \cap A)/P(A)$

$P(B_3 \cap A)=P(B_3)P(A \backslash B_3)= (0.25)(0.02) = 0.005$,

$P(A)=P(B_1)P(A \backslash B_1)+P(B_2)P(A \backslash B_2)+P(B_3)P(A \backslash B_3)= 0.006+0.0135+0.005= 0.0245$.

Then,

$P(B_3 \backslash A)= 0.005/0.0245=0.2041$

Example 4.14: - The table shows the probabilities for product failure subjected to level of contamination in manufacturing:

Probability of Failure	Level of Contamination
0.10	High
0.01	Medium
0.001	Low

In a particular production run, 20% of the chips are subjected to high levels of contamination, 30% to medium levels of contamination, and 50% to low levels of contamination. What is the probability that a product using one of these chips fails?

Solution: -

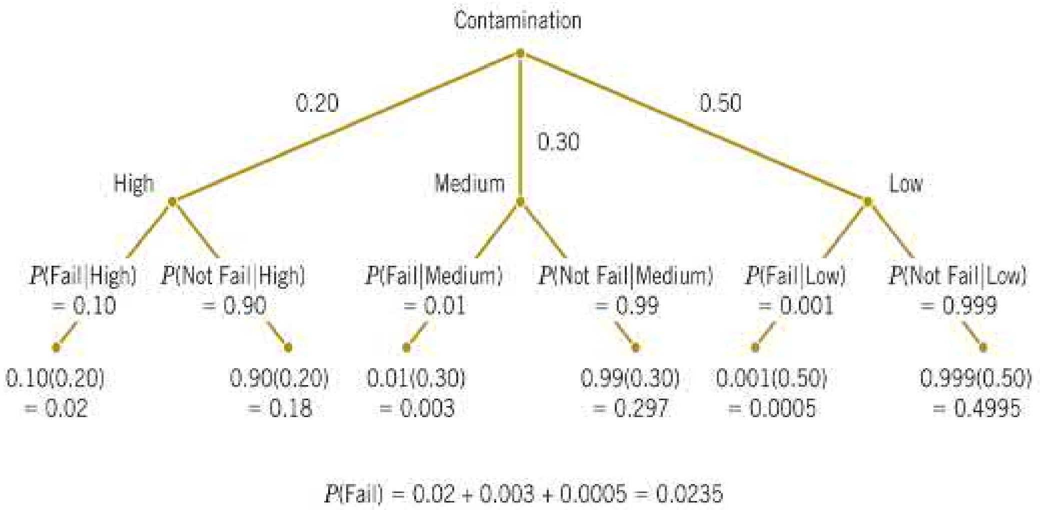
Let:

H denote the event that a chip is exposed to **high** levels of contamination

M denote the event that a chip is exposed to **medium** levels of contamination

L denote the event that a chip is exposed to **low** levels of contamination

Then,



$P(F)=P(F \backslash H)P(H)+P(F \backslash M)P(M)+P(F \backslash L)P(L)=0.10(0.2)+0.01(0.30)+0.001(0.50)=0.0235$

The calculations are conveniently organized with the tree diagram.

4.10 Geometric Probability

Probability is always expressed as a ratio between 0 and 1 that gives a value to how likely an event is to happen. A probability of 0 means there is no chance of that event happening. A probability of 1 means the particular event will always happen. To calculate geometric probability, it will need to find the areas of the shapes involved in the problem. It will need to know the total area, which means the biggest area in the diagram. It will also need to know the desired area. The formula is simply:

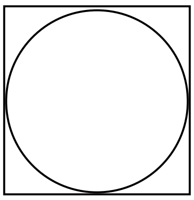
$$P = \frac{desired}{total}$$

Where, P represents the geometric probability. **Desired** stands for the area. **Total** stands for the area of the whole figure

Example 4.14: Find the probability for a circle inscribed in a square of 5 cm?
Solution:-

1- Draw both of circle and square.

2- When circle inscribed in a square, it must have four points in its boundaries in tangent with the square.



3- Find the area of circle and the area of square.

Area of circle=πr^2
$= 3.14 * (2.5)^2 = 19.63\ cm^2$
Area of square=side ^{2}
$= (5)^2 = 25\ cm^2$

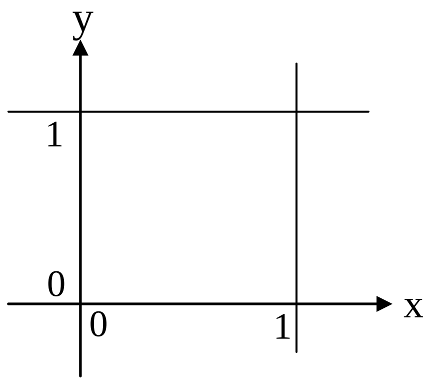
4- Find the geometric probability.

$$P = \frac{Area\ of\ Circle}{Total\ Area(Square\ Area)} = \frac{19.63}{25} = 0.785$$

Example 4.15: Two numbers are chosen which their values between 0 and 1 randomly. What is the probability that the sum of their squares is greater than 1?

Solution:-

1. Represent each number by a coordinate axis and then determine the sample space.



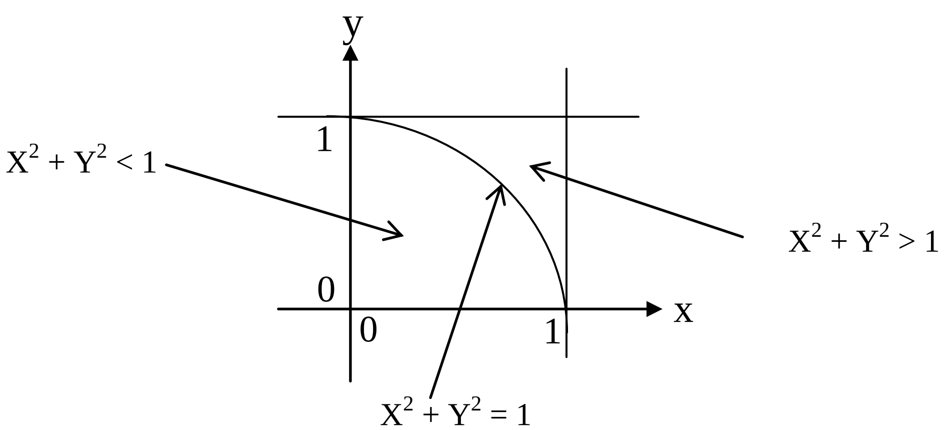
2. Derive a mathematical expression for the required event. So ,

   $X^2 + Y^2 > 1$ (this is the required event)

3. Change the above expression to a mathematical equation.

   $X^2 + Y^2 = 1$ (a circle equation with r = 1)

4. Now, draw the above equation in the sample space.



Then,

P = Desired / Total

Total = 1 x 1 = 1

Desired= $1 - (\pi\, 1^2 / 4) = 1 - \pi/4$

$P = (1 - \pi/4) / 1 = 1 - \pi/4 = 0.21$

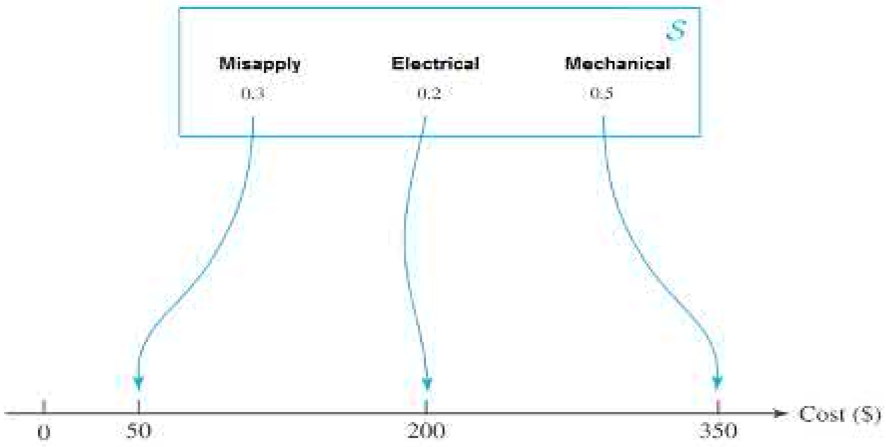
CHAPTER FIVE

PROBABILITY DISTRIBUTION

5.1 Introduction

In an experiment, a measurement is usually denoted by a variable such as X. In a random experiment, a variable whose measured value can change (from one replicate of the experiment to another) is referred to as a **random variable**. There are two types of random variables. A **discrete** random variable is a random variable with a finite (or countable infinite) set of real numbers for its range. A **continuous** random variable is a random variable with an interval (either finite or infinite) of real numbers for its range.

The sample space for the machine breakdown problem is S = {electrical, mechanical, misapply} and each of these failures may be associated with a repair cost. For example, suppose that electrical failures generally cost an average of $200 to repair, mechanical failures have an average repair cost of $350, and operator misapply failures have an average repair cost of only $50. These repair costs generate a random variable cost, as illustrated in Figure below, which as a state space of { 50 , 200 , 350 } .



Notice that cost is a random variable because its values 50, 200, and 350 are numbers. The breakdown cause, defined to be electrical, mechanical, or operator misapply, is not considered to be a random variable because its values are not numerical.

The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X. The two kinds of probability distribution of a random variable are Discrete and Continuous.

5.2 Discrete probability distribution

For a discrete random variable X, its distribution can be described by a function that specifies the probability at each of the possible discrete values for X.

For a discrete random variable X with possible values x_1 , x_2 ,…., x_n a **probability mass function** (***pmf***) is a function such that

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $f(x_i) = P(X = x_i)$

The cumulative distribution function of a discrete random variable X, denoted as F(x), is $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$ and F(x) satisfies the following properties,

(1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

(2) $0 \leq F(x) \leq 1$

(3) $if\ x \leq y,\ then\ F(x) \leq F(y)$

Example 5.1: **-** Consider randomly selecting a student at university by X=1 if the selected student does not qualify and X=0 if the student does qualify. If 20% of all students do not qualify, find the probability mass function (pmf) for X ?

Solution.

The pmf for X is

p(0)=P(X=0)=P(the selected student does qualify)= 0.8

p(1)=P(X=1)=P(the selected student does not qualify)= 0.2

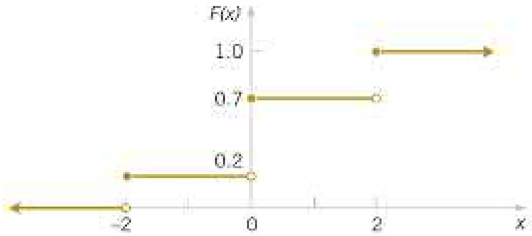
p(x)=P(X=x)=0 for x≠0 or 1.

$$p(x) = \begin{cases} 0.8 & if \quad x = 0 \\ 0.2 & if \quad x = 1 \\ 0 & if \quad x \neq 0 \ or \ 1 \end{cases}$$

Example 5.2: **-** Determine the probability mass function (pmf) of X from the following cumulative distribution function:

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

Solution.



The Figure above displays a plot of F(x) . From the plot, the only points that receive nonzero probability are –2, 0, and 2. The probability mass function (pmf) at each point is the jump in the cumulative distribution function at the point.
Therefore,

$f(-2) = 0.2 - 0 = 0.2$

$f(0) = 0.7 - 0.2 = 0.5$

$f(2) = 1.0 - 0.7 = 0.3$

5.2.1 Discrete Uniform Distribution

The simplest discrete random variable is one that assumes only a finite number of possible values, each with equal probability. A random variable X that assumes each of the values $x_1, x_2, \ldots, x_n$ with equal probability 1/n. Then

$$f(x_i) = \frac{1}{n}$$

Suppose that X is a discrete uniform random variable on the consecutive integers a, a+1 ,a+2,…….. b, for $a \le b$. Then

The mean of X is

$$\mu = E(X) = \frac{b+a}{2}$$

The variance of X is

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12}$$

Example 5.3: - The first digit of serial number is any one of the digits 0 through 9. If one part is selected and X is the first digit, find f(x)?

Solution.

X has a discrete uniform distribution with probability 0.1 for each value. That is, $f(x) = 0.1$

Example 5.4: - A voice communication system contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used. Find the mean and the variance?

Solution

Let the random variable X denote the number of 48 (lines) in use.

Then X can assume any of the integer values with a range of 0 to 48.

$$\mu = E(X) = \frac{b+a}{2} = \frac{48+0}{2} = 24$$

And,

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12} = \frac{(48-0+1)^2 - 1}{12}$$

$$\sigma = \sqrt{\frac{(48-0+1)^2 - 1}{12}} = 14.14$$

5.2.2 Binomial Distribution

A trial with only two possible outcomes is used so frequently as a building block of a random experiment that it is called a Bernoulli trial.

A random experiment consists of n Bernoulli trials such that

(1) The trials are independent.

(2) Each trial results in only two possible outcomes, labeled as "success" and "failure."

(3) The probability of a success in each trial, denoted as p, remains constant.

The random variable X that equals the number of trials that result in a success is a binomial random variable with parameters $0 < p < 1$ and $n = 1,2$ …. The probability mass function of X is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad , \quad x = 0,1,2,......,n$$

Where, $\binom{n}{x}$ represent the following $\frac{n!}{x!\ (n-x)!}$ value.

If X is a binomial random variable with parameters p and n,

The mean of X is

$$\mu = E(X) = n\,p$$

The variance of X is

$$\sigma^2 = n\,p\,(1-p)$$

Example 5.5:- Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the following :-

(a) The probability that in the next 18 samples, exactly 2 contain the pollutant.

(b) The probability that at least four samples contain the pollutant in 18 samples.

(c) The probability that $3 \le X < 7$ where X are samples contain the pollutant .

Solution

Let X = the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with p = 0.1 and n = 18.

(a) $P(X=2)=\binom{18}{2}0.1^2\ (1-0.1)^{18-2}=0.284$

(b) $P(X\geq 4)=\sum_{x=4}^{18}\binom{18}{x}0.1^x\ (1-0.1)^{18-x}=1-\sum_{x=0}^{3}\binom{18}{x}0.1^x\ (1-0.1)^{18-x}=0.098$

(c) $P(3\leq X<7)=\sum_{x=3}^{6}\binom{18}{x}0.1^x\ (1-0.1)^{18-x}=0.265$

Example 5.6:- Let X be a binomial random variable with p=0.1, and n=10. Calculate the following probabilities from the binomial probability mass function:

$(a).\ P(X\leq 2)\quad (b).\ P(X=4)\quad (c).\ P(5\leq X\leq 7)$

Solution

(a) $P(X=2)=\sum_{x=0}^{2}\binom{10}{x}0.1^x\ (1-0.1)^{10-x}=0.9298$

(b) $P(X=4)=\binom{10}{4}0.1^4\ (1-0.1)^6=0.0112$

(c) $P(5\leq X\leq 7)=\sum_{x=5}^{7}\binom{10}{x}0.1^x\ (1-0.1)^{10-x}=0.0016$

NOTE: Using Binomial Tables

Calculating binomial probabilities becomes tedious when n is large. For some values of n and p , the binomial probabilities have been tabulated in Tables below. The entries in the table represent cumulative binomial probabilities.

For example, N=10, the entry in the column corresponding to p = 0.10 and the row corresponding to x = 2 is 0.930 , and its interpretation is

$$P(x\leq 2)=\sum_{x=0}^{2}\binom{10}{x}0.1^x\ (1-0.1)^{10-x}=P(x=0)+P(x=1)+P(x=2)=0.930$$

$$For\ x=2,\quad P(x=2)=P(x\leq 2)-P(x\leq 1)=0.930-0.736=0.194$$

$$For\ x>2,\quad P(x>2)=1-P(x\leq 2)=1-0.930=0.070$$

Classwork: - For N=20, p=0.6. Use **Binomial Tables** to calculate the following:

1- The probability that $x\leq 10$
2- The probability that $x>12$
3- The probability that $x=11$

Ans.

1- 0.245 2- 0.416 3- 0.159

N=5

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000
1	.999	.977	.919	.737	.528	.337	.188	.087	.031	.007	.000	.000	.000
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000
3	1.000	1.000	1.000	.993	.969	.913	.812	.663	.472	.263	.081	.023	.001
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049

N=6

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.941	.735	.531	.262	.118	.047	.016	.004	.001	.000	.000	.000	.000
1	.999	.967	.886	.655	.420	.233	.109	.041	.011	.002	.000	.000	.000
2	1.000	.998	.984	.901	.744	.544	.344	.179	.070	.017	.001	.000	.000
3	1.000	1.000	.999	.983	.930	.821	.656	.456	.256	.099	.016	.002	.000
4	1.000	1.000	1.000	.998	.989	.959	.891	.767	.580	.345	.114	.033	.001
5	1.000	1.000	1.000	1.000	.999	.996	.984	.953	.882	.738	.469	.265	.059

N=7

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.932	.698	.478	.210	.082	.028	.008	.002	.000	.000	.000	.000	.000
1	.998	.956	.850	.577	.329	.159	.063	.019	.004	.000	.000	.000	.000
2	1.000	.996	.974	.852	.647	.420	.227	.096	.029	.005	.000	.000	.000
3	1.000	1.000	.997	.967	.874	.710	.500	.290	.126	.033	.003	.000	.000
4	1.000	1.000	1.000	.995	.971	.904	.773	.580	.353	.148	.026	.004	.000
5	1.000	1.000	1.000	1.000	.996	.981	.937	.841	.671	.423	.150	.044	.002
6	1.000	1.000	1.000	1.000	1.000	.998	.992	.972	.918	.790	.522	.302	.068

N=8

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.923	.663	.430	.168	.058	.017	.004	.001	.000	.000	.000	.000	.000
1	.997	.943	.813	.503	.255	.106	.035	.009	.001	.000	.000	.000	.000
2	1.000	.994	.962	.797	.552	.315	.145	.050	.011	.001	.000	.000	.000
3	1.000	1.000	.995	.944	.806	.594	.363	.174	.058	.010	.000	.000	.000
4	1.000	1.000	1.000	.990	.942	.826	.637	.406	.194	.056	.005	.000	.000
5	1.000	1.000	1.000	.999	.989	.950	.855	.685	.448	.203	.038	.006	.000
6	1.000	1.000	1.000	1.000	.999	.991	.965	.894	.745	.497	.187	.057	.003
7	1.000	1.000	1.000	1.000	1.000	.999	.996	.983	.942	.832	.570	.337	.077

N=9

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.914	.630	.387	.134	.040	.010	.002	.000	.000	.000	.000	.000	.000
1	.997	.929	.775	.436	.196	.071	.020	.004	.000	.000	.000	.000	.000
2	1.000	.992	.947	.738	.463	.232	.090	.025	.004	.000	.000	.000	.000
3	1.000	.999	.992	.914	.730	.483	.254	.099	.025	.003	.000	.000	.000
4	1.000	1.000	.999	.980	.901	.733	.500	.267	.099	.020	.001	.000	.000
5	1.000	1.000	1.000	.997	.975	.901	.746	.517	.270	.086	.008	.001	.000
6	1.000	1.000	1.000	1.000	.996	.975	.910	.768	.537	.262	.053	.008	.000
7	1.000	1.000	1.000	1.000	1.000	.996	.980	.929	.804	.564	.225	.071	.003
8	1.000	1.000	1.000	1.000	1.000	1.000	.998	.990	.960	.866	.613	.370	.086

N=10

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.904	.599	.349	.107	.028	.006	.001	.000	.000	.000	.000	.000	.000
1	.996	.914	.736	.376	.149	.046	.011	.002	.000	.000	.000	.000	.000
2	1.000	.988	.930	.678	.383	.167	.055	.012	.002	.000	.000	.000	.000
3	1.000	.999	.987	.879	.650	.382	.172	.055	.011	.001	.000	.000	.000
4	1.000	1.000	.998	.967	.850	.633	.377	.166	.047	.006	.000	.000	.000
5	1.000	1.000	1.000	.994	.953	.834	.623	.367	.150	.033	.002	.000	.000
6	1.000	1.000	1.000	.999	.989	.945	.828	.618	.350	.121	.013	.001	.000
7	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.322	.070	.012	.000
8	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.624	.264	.086	.004
9	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.893	.651	.401	.096

N=15

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.860	.463	.206	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
1	.990	.829	.549	.167	.035	.005	.000	.000	.000	.000	.000	.000	.000
2	1.000	.964	.816	.398	.127	.027	.004	.000	.000	.000	.000	.000	.000
3	1.000	.995	.944	.648	.297	.091	.018	.002	.000	.000	.000	.000	.000
4	1.000	.999	.987	.838	.515	.217	.059	.009	.001	.000	.000	.000	.000
5	1.000	1.000	.998	.939	.722	.403	.151	.034	.004	.000	.000	.000	.000
6	1.000	1.000	1.000	.982	.869	.610	.304	.095	.015	.001	.000	.000	.000
7	1.000	1.000	1.000	.996	.950	.787	.500	.213	.050	.004	.000	.000	.000
8	1.000	1.000	1.000	.999	.985	.905	.696	.390	.131	.018	.000	.000	.000
9	1.000	1.000	1.000	1.000	.996	.966	.849	.597	.278	.061	.002	.000	.000
10	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.164	.013	.001	.000
11	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.352	.056	.005	.000
12	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.602	.184	.036	.000
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.833	.451	.171	.010
14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.794	.537	.140

N=20

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.818	.358	.122	.012	.001	.000	.000	.000	.000	.000	.000	.000	.000
1	.983	.736	.392	.069	.008	.001	.000	.000	.000	.000	.000	.000	.000
2	.999	.925	.677	.206	.035	.004	.000	.000	.000	.000	.000	.000	.000
3	1.000	.984	.867	.411	.107	.016	.001	.000	.000	.000	.000	.000	.000
4	1.000	.997	.957	.630	.238	.051	.006	.000	.000	.000	.000	.000	.000
5	1.000	1.000	.989	.804	.416	.126	.021	.002	.000	.000	.000	.000	.000
6	1.000	1.000	.998	.913	.608	.250	.058	.006	.000	.000	.000	.000	.000
7	1.000	1.000	1.000	.968	.772	.416	.132	.021	.001	.000	.000	.000	.000
8	1.000	1.000	1.000	.990	.887	.596	.252	.057	.005	.000	.000	.000	.000
9	1.000	1.000	1.000	.997	.952	.755	.412	.128	.017	.001	.000	.000	.000
10	1.000	1.000	1.000	.999	.983	.872	.588	.245	.048	.003	.000	.000	.000
11	1.000	1.000	1.000	1.000	.995	.943	.748	.404	.113	.010	.000	.000	.000
12	1.000	1.000	1.000	1.000	.999	.979	.868	.584	.228	.032	.000	.000	.000
13	1.000	1.000	1.000	1.000	1.000	.994	.942	.750	.392	.087	.002	.000	.000
14	1.000	1.000	1.000	1.000	1.000	.998	.979	.874	.584	.196	.011	.000	.000
15	1.000	1.000	1.000	1.000	1.000	1.000	.994	.949	.762	.370	.043	.003	.000
16	1.000	1.000	1.000	1.000	1.000	1.000	.999	.984	.893	.589	.133	.016	.000
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.965	.794	.323	.075	.001
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.992	.931	.608	.264	.017
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.988	.878	.642	.182

N=25

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.778	.277	.072	.004	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.974	.642	.271	.027	.002	.000	.000	.000	.000	.000	.000	.000	.000
2	.998	.873	.537	.098	.009	.000	.000	.000	.000	.000	.000	.000	.000
3	1.000	.966	.764	.234	.033	.002	.000	.000	.000	.000	.000	.000	.000
4	1.000	.993	.902	.421	.090	.009	.000	.000	.000	.000	.000	.000	.000
5	1.000	.999	.967	.617	.193	.029	.002	.000	.000	.000	.000	.000	.000
6	1.000	1.000	.991	.780	.341	.074	.007	.000	.000	.000	.000	.000	.000
7	1.000	1.000	.998	.891	.512	.154	.022	.001	.000	.000	.000	.000	.000
8	1.000	1.000	1.000	.953	.677	.274	.054	.004	.000	.000	.000	.000	.000
9	1.000	1.000	1.000	.983	.811	.425	.115	.013	.000	.000	.000	.000	.000
10	1.000	1.000	1.000	.994	.902	.586	.212	.034	.002	.000	.000	.000	.000
11	1.000	1.000	1.000	.998	.956	.732	.345	.078	.006	.000	.000	.000	.000
12	1.000	1.000	1.000	1.000	.983	.846	.500	.154	.017	.000	.000	.000	.000
13	1.000	1.000	1.000	1.000	.994	.922	.655	.268	.044	.002	.000	.000	.000
14	1.000	1.000	1.000	1.000	.998	.966	.788	.414	.098	.006	.000	.000	.000
15	1.000	1.000	1.000	1.000	1.000	.987	.885	.575	.189	.017	.000	.000	.000
16	1.000	1.000	1.000	1.000	1.000	.996	.946	.726	.323	.047	.000	.000	.000
17	1.000	1.000	1.000	1.000	1.000	.999	.978	.846	.488	.109	.002	.000	.000
18	1.000	1.000	1.000	1.000	1.000	1.000	.993	.926	.659	.220	.009	.000	.000
19	1.000	1.000	1.000	1.000	1.000	1.000	.998	.971	.807	.383	.033	.001	.000
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.910	.579	.098	.007	.000
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.967	.766	.236	.034	.000
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.902	.463	.127	.002
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.973	.729	.358	.026
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.928	.723	.222

5.2.3 Geometric Distribution

In a series of Bernoulli trials (independent trials with constant probability p of a success), the random variable X that equals the number of trials until the first success is a **geometric random variable** with parameter $0 < p < 1$ and

$$f(x) = (1-p)^{x-1} p \quad , \quad x = 1, 2, ...$$

The mean of X is

$$\mu = E(X) = \frac{1}{p}$$

The variance of X is

$$\sigma^2 = V(X) = \frac{(1-p)}{p^2}$$

Example 5.7:- Let X denotes a random variable having a geometric distribution, with probability of success on any trial "p" (p=0.4) . Find $(a). P(X \le 2)$ $(b). P(4 < X < 7)$

Solution

$$a - f(x) = P(X \le 2) = (1-0.4)^{1-1} 0.4 + (1-0.4)^{2-1} 0.4 = 0.64$$

$$b - f(x) = P(4 < X < 7) = (1-0.4)^{5-1} 0.4 + (1-0.4)^{6-1} 0.4 = 0.0829$$

5.2.4 Negative binomial Distribution

In Section 5.2.3, the geometric distribution models show the probabilistic behavior of the number of the trial on which the first success occurs in a sequence of independent Bernoulli trials. But what if the interest is in the number of the trial for the second success, or the third success, or in general, the r^{th} success? The distribution governing the probabilistic behavior in these cases is called the **negative binomial distribution**.

Let Y denote the number on the trial on which the r^{th} success occurs in a sequence of independent Bernoulli trials with p denoting the common probability of "success." The negative binomial distribution is defined by two parameters, r and p.

$$P(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r} \quad , \quad y = r, r+1, ... \textit{ and } \; r = 1, 2, 3, ... \; \textit{for } 0 < p < 1$$

The mean of Y is

$$\mu = E(Y) = \frac{r}{p}$$

The variance of Y is

$$\sigma^2 = V(Y) = \frac{r(1-p)}{p^2}$$

Example 5.8:- Let Y denotes a random variable having a negative binomial distribution, with p=0.4 . Find $P(Y \geq 4)$ if **a-** r=2 and **b-** r=4

Solution

$$a - P(Y \geq 4) = \binom{y-1}{2-1} 0.4^2 (1-0.4)^{y-2} \ , \quad y = 2,3,4,5,.......$$

$$P(Y \geq 4) = 1 - \left[\binom{1}{1} 0.4^2 (0.6)^{2-2} + \binom{2}{1} 0.4^2 (0.6)^{3-2} \right] = 1 - [0.16 + 0.192] = 0.648$$

$$b - P(Y \geq 4) = \binom{y-1}{4-1} 0.4^4 (1-0.4)^{y-4} \ , \quad y = 4,5,6,7,.......$$

$$P(Y \geq 4) = 1$$

Example 5.9:- It was found that 30% of the applicants for a certain job have advanced training in computer programming. Applicants are selected at random and are interviewed sequentially.

a- Find the probability that the first applicant having advanced training is found on the fifth interview.

b- Suppose three jobs requiring advanced programming training are open. Find the probability that the third qualified applicant is found on the fifth interview, if the applicants are interviewed sequentially and at random.

Solution

$$a - P(Y = 5) = (1-0.3)^{5-1} 0.3 = 0.072$$

b- It was assumed independent trials, with the probability of finding a qualified applicant on any one trials being 0.3. Let Y denote the number of the trial on which the third qualified applicant is found. Then Y can reasonably be assumed to have a negative binomial distribution with r=3 and p=0.3 , so

$$P(Y = 5) = \binom{5-1}{3-1} 0.3^3 (1-0.3)^{5-3} = 0.0794$$

Example 5.10:- Camera Flashes consider the time to recharge the flash. The probability that a camera passes the test is 0.8, and the cameras perform independently. What is the probability that the third failure is obtained in five or fewer tests?

Solution

Let Y denote the number of cameras tested until three failures have been obtained. The requested probability is P (Y ≤ 5).

Here Y has a negative binomial distribution with p = (1-0.8) = 0.2 and r = 3. Therefore,

$$P(Y \leq 5) = \sum_{y=3}^{5} \binom{y-1}{3-1} 0.2^3 (1-0.2)^{y-3} = 0.056$$

5.3 Continuous Probability Distributions

Density functions are commonly used in engineering to describe physical systems. A **probability density function** f (x) can be used to describe the probability distribution of a continuous random variable X. If an interval is likely to contain a value for X, its probability is large and it corresponds to large values for f (x) . The probability that X is between a and b is determined as the integral of f (x) from a to b. Then,

For a continuous random variable X, a **probability density function (pdf)** is a function such that

(1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x)dx = 1$

(3) $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx = \text{area under } f(x) \text{ from } a \text{ to } b \text{ for any } a \text{ and } b$

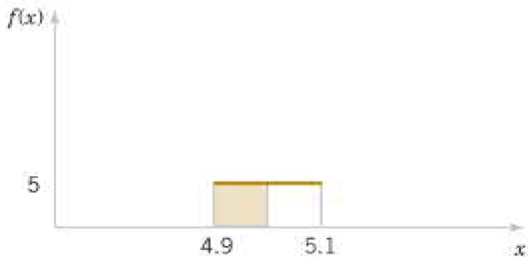
If X is a continuous random variable, for any x_1 and x_2 ,

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2)$$

Example 5.11:- Let X denote the continuous random variable of the current measured in a thin copper wire in milliamperes. Assume that the range of X is [4.9, 5.1] mA, and assume that the probability density function of X is f (x) = 5 for 4.9 ≤ x≤5.1. What is the probability that a current measurement is less than 5 milliamperes?

Solution

The probability density function is shown in the figure. It is assumed that f (x) =0 wherever it is not specifically defined. The shaded area in the figure indicates the probability.

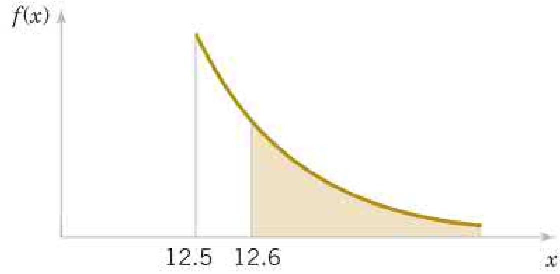


$$P(X < 5) = \int_{4.9}^{5} f(x)dx = \int_{4.9}^{5} 5dx = 0.5$$

Example 5.12:- Let X denote the continuous random variable of the diameter of a hole drilled in a metal sheet. The target diameter is 12.5 millimeters. From past data show that the distribution of X can be modeled by a probability density function of $f(x) = 20e^{-20(x-12.5)}$ for $x \geq 12.5$. If a part with a diameter greater than 12.6 mm is required, what proportion of those parts?

Solution

From the question, A part is required if X >12.6. Now, the density function is shown in the figure.



$$P(X > 12.6) = \int_{12.6}^{\infty} f(x)dx = \int_{12.6}^{\infty} 20e^{-20(x-12.5)}dx$$

$$P(X > 12.6) = -e^{-20(x-12.5)} = 0.135$$

Now, it can be calculate the proportion of parts between 12.5 and 12.6 mm,

$$P(12.5 < X < 12.6) = \int_{12.5}^{12.6} f(x)dx = -e^{-20(x-12.5)} = 0.865 \quad \text{Or}$$

Total area under f (x) = 1 , then P(X>12.6)=1- P(12.5<X<12.6)= 1-0.865=0.135

The **cumulative distribution function (cdf)** of a continuous random variable X is

$$F(x) = P(X \leq x) = \int_{-\infty}^{\infty} f(u)du \quad for -\infty < x < \infty$$

For example, 5.12, the cumulative distribution function can be shown below, consists of two expressions. F (x) =0 for x<12.5

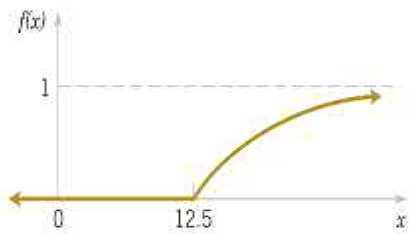
And the other expression can be determine when $x \geq 12.5$ as shown below

$$F(x) = \int_{12.5}^{x} 20e^{-20(x-12.5)}du = 1 - e^{-20(x-12.5)}$$

Therefore,

$$F(x) = \begin{cases} 0 & x < 12.5 \\ 1 - e^{-20(x-12.5)} & 12.5 \leq x \end{cases}$$

The figure represents the graph of F(x).



The probability density function of a continuous random variable can be determined from the cumulative distribution function by differentiating. The fundamental theorem of calculus states that $\frac{d}{dx}\int_{-\infty}^{x} f(u)du = f(x)$

Then given F(x), $f(x) = \frac{dF(x)}{dx}$

Example 5.13:- The time until a chemical reaction is complete (in milliseconds) is approximated by the cumulative distribution function F(x), Find the probability density function of X.

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

Solution

Using the result that the probability density function is the derivative of F (x) ,

$$f(x) = \begin{cases} 0 & x < 0 \\ 0.01e^{-0.01x} & 0 \le x \end{cases}$$

The probability that a reaction completes within 200 milliseconds is

$$P(X < 200) = F(200) = 1 - e^{-2} = 0.8647$$

The mean and variance can also be defined for a continuous random variable. Suppose that X is a continuous random variable with probability density function f (x) . The mean or expected value of X, denoted as μ or E (X) , is

$$\mu = E(X) = \int_{-\infty}^{\infty} x\,f(x)dx$$

The variance of X, denoted as V (X) or σ^2 , is

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

The standard deviation of X is

$$\sigma = \sqrt{\sigma^2}$$

Example 5.14:- Find the mean and variance of X in example 5.11?

Solution

The mean will be,

$$\mu = E(X) = \int_{4.9}^{5.1} x\,f(x)dx = 5x^2/2 = 5$$

The variance will be,

$$\sigma^2 = V(X) = \int_{4.9}^{5.1} (x-10)^2 f(x)dx = 5(x-10)^3/3 = 0.0033$$

5.3.1 Continuous Uniform Distributions

A continuous random variable X with probability density function f(x)=1(b-a), $a \le x \le b$ is a continuous uniform random variable.

The mean of continuous uniform random variable, is

$$\mu = E(X) = \frac{(a+b)}{2}$$

The variance of continuous uniform random variable, is

$$\sigma^2 = V(X) = \frac{(b-a)^2}{12}$$

Example 5.15:- Find the mean and variance of X in example 5.11?

Solution

The mean and variance formulas can be applied with a=4.9 and b=5.1. Therefore,

$$\mu = E(X) = \frac{(4.9+5.1)}{2} = 5\ mA \quad \text{and} \quad \sigma^2 = V(X) = \frac{(5.1-4.9)^2}{12} = 0.0033\ mA$$

Consequently, the standard deviation of X is $0.0577\ mA$.

The cumulative distribution function of a continuous uniform random variable is obtained by integration. If $a < x < b$,

$$F(x) = \int_a^x \frac{1}{b-a} du = \frac{x-a}{b-a}$$

Therefore, the complete description of the cumulative distribution function of a continuous uniform random variable is,

$$f(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x < b \\ 1 & b \le x \end{cases}$$

5.3.2 Normal Distributions

A random variable X with probability density function

$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$ is a normal random variable with parameters μ where $-\infty < \mu < \infty$, and σ > 0. Also $E(X) = \mu$ and $V(X) = \sigma^2$ and the notation $N(\mu, \sigma^2)$ is used to denote the distribution.

A normal random variable with μ=0 and $\sigma^2=1$ is called a standard normal random variable and is denoted as Z . The cumulative distribution function of a standard normal random variable is denoted as $\Phi(z) = P(Z \le z)$

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$

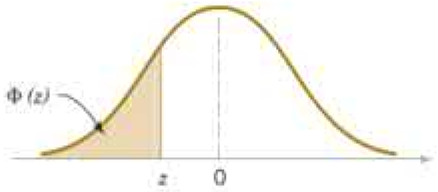


TABLE • Cumulative Standard Normal Distribution

z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.03	−0.01	−0.00
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687
−3.1	0.000711	0.000736	0.000762	0.000789	0.000816	0.000845	0.000874	0.000904	0.000935	0.000968
−3.0	0.001001	0.001035	0.001070	0.001107	0.001144	0.001183	0.001223	0.001264	0.001306	0.001350
−2.9	0.001395	0.001441	0.001489	0.001538	0.001589	0.001641	0.001695	0.001750	0.001807	0.001866
−2.8	0.001926	0.001988	0.002052	0.002118	0.002186	0.002256	0.002327	0.002401	0.002477	0.002555
−2.7	0.002635	0.002718	0.002803	0.002890	0.002980	0.003072	0.003167	0.003264	0.003364	0.003467
−2.6	0.003573	0.003681	0.003793	0.003907	0.004025	0.004145	0.004269	0.004396	0.004527	0.004661
−2.5	0.004799	0.004940	0.005085	0.005234	0.005386	0.005543	0.005703	0.005868	0.006037	0.006210
−2.4	0.006387	0.006569	0.006756	0.006947	0.007143	0.007344	0.007549	0.007760	0.007976	0.008198
−2.3	0.008424	0.008656	0.008894	0.009137	0.009387	0.009642	0.009903	0.010170	0.010444	0.010724
−2.2	0.011011	0.011304	0.011604	0.011911	0.012224	0.012545	0.012874	0.013209	0.013553	0.013903
−2.1	0.014262	0.014629	0.015003	0.015386	0.015778	0.016177	0.016586	0.017003	0.017429	0.017864
−2.0	0.018309	0.018763	0.019226	0.019699	0.020182	0.020675	0.021178	0.021692	0.022216	0.022750
−1.9	0.023295	0.023852	0.024419	0.024998	0.025588	0.026190	0.026803	0.027429	0.028067	0.028717
−1.8	0.029379	0.030054	0.030742	0.031443	0.032157	0.032884	0.033625	0.034379	0.035148	0.035930
−1.7	0.036727	0.037538	0.038364	0.039204	0.040059	0.040929	0.041815	0.042716	0.043633	0.044565
−1.6	0.045514	0.046479	0.047460	0.048457	0.049471	0.050503	0.051551	0.052616	0.053699	0.054799
−1.5	0.055917	0.057053	0.058208	0.059380	0.060571	0.061780	0.063008	0.064256	0.065522	0.066807
−1.4	0.068112	0.069437	0.070781	0.072145	0.073529	0.074934	0.076359	0.077804	0.079270	0.080757
−1.3	0.082264	0.083793	0.085343	0.086915	0.088508	0.090123	0.091759	0.093418	0.095098	0.096801
−1.2	0.098525	0.100273	0.102042	0.103835	0.105650	0.107488	0.109349	0.111233	0.113140	0.115070
−1.1	0.117023	0.119000	0.121001	0.123024	0.125072	0.127143	0.129238	0.131357	0.133500	0.135666
−1.0	0.137857	0.140071	0.142310	0.144572	0.146859	0.149170	0.151505	0.153864	0.156248	0.158655
−0.9	0.161087	0.163543	0.166023	0.168528	0.171056	0.173609	0.176185	0.178786	0.181411	0.184060
−0.8	0.186733	0.189430	0.192150	0.194894	0.197662	0.200454	0.203269	0.206108	0.208970	0.211855
−0.7	0.214764	0.217695	0.220650	0.223627	0.226627	0.229650	0.232695	0.235762	0.238852	0.241964
−0.6	0.245097	0.248252	0.251429	0.254627	0.257846	0.261086	0.264347	0.267629	0.270931	0.274253
−0.5	0.277595	0.280957	0.284339	0.287740	0.291160	0.294599	0.298056	0.301532	0.305026	0.308538
−0.4	0.312067	0.315614	0.319178	0.322758	0.326355	0.329969	0.333598	0.337243	0.340903	0.344578
−0.3	0.348268	0.351973	0.355691	0.359424	0.363169	0.366928	0.370700	0.374484	0.378281	0.382089
−0.2	0.385908	0.389739	0.393580	0.397432	0.401294	0.405165	0.409046	0.412936	0.416834	0.420740
−0.1	0.424655	0.428576	0.432505	0.436441	0.440382	0.444330	0.448283	0.452242	0.456205	0.460172
0.0	0.464144	0.468119	0.472097	0.476078	0.480061	0.484047	0.488033	0.492022	0.496011	0.500000

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$

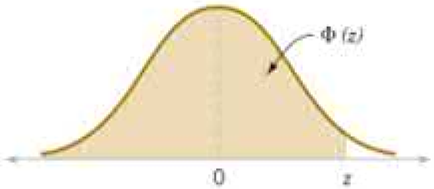


TABLE • Cumulative Standard Normal Distribution (*Continued*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959071	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965621	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967