15. MATHEMATICAL REASONING AND STATISTICS

1. MATHEMATICAL REASONING

1.1 Introduction

The literal meaning of the word 'Logic' is the use and study of valid reasoning. Reasoning is the process of understanding and conveying, logically, ideas, concepts, facts etc., in a particular language, and sentences are the linguistic unit used to communicate in any language. Therefore, it is imperative that we understand the different types of sentences as all reasoning is based on the way we put sentences together. The different types of sentences are as follows:

(1) Assertive sentence **(2)** Imperative sentence

(3) Exclamatory sentence **(4)** Interrogative sentence

In this chapter, we will be discussing assertive sentences, also known as statements or propositions.

2. STATEMENT/PROPOSITION

A statement or a proposition is an assertion (assertive sentence) that can be either proved true or false, but not both. If the statement is proven to be true, then it is called a valid statement, else it is called an invalid statement.

Open statement: An assertive sentence containing one or more variables is an open statement when the variable(s) is/are replaced by some definite value(s).

Truth set: An open statement becomes a valid statement when the values replacing the variables make the statement true. Therefore, all the values which make an open statement a true statement are collectively referred to as the truth set of the open statement.

Truth value: Truth value refers to the truth or falsity of a statement. The truth value of true statement is 'True' or 'T'. Similarly, the truth value of a false statement is 'False' or 'F'

Logical variables: In mathematical reasoning, statements are represented by lower case letters such as p, q, r, s. These are referred to as logical variables.

For example, the statement 'The earth revolves around the sun' can be denoted by p and is written as p: The earth revolves around the sun. [And its truth value is T]

Quantifiers: Symbols like $\forall$ (stands for "for all") and $\exists$ (stands for "there exists") that are used to denote a group of words or a phrase are known as quantifiers.

The symbols $\forall$ and $\exists$ are known as existential quantifiers. An open sentence containing quantifiers always becomes a statement.

Quantified statements: Statements containing quantifiers are known as quantified statements, e.g., $x^2 \geq 0 \ \forall \ x \in R$. Its truth value is T.

2.1 Using Venn Diagrams for Truth Values

Venn diagrams canbe used to represent the truth value of statements. Let us consider the statement "all teachers are scholars." Let us also assume that this statement is true. The truth of the above statement can be represented by the Venn diagram in Fig. 15.1, where:

U = The set of all human beings

S = The set of all scholars

and T= The set of all teachers

Clearly, $S \subset U$ and $T \subset U$

From the above statement, it can be logically inferred that $T \subset S$.

Thus, the truth value of the above statement can be represented by the Venn diagram shown in Fig. 15.1.

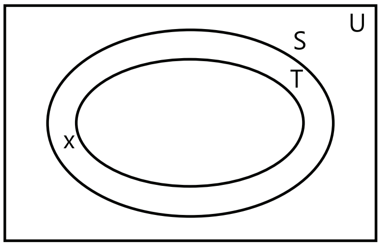

Figure 15.1

Similarly, we can also check the truth value of other statements which are connected to a given statement. For example, consider the statement "There are some scholars who are teachers". It is evident from the Venn diagram (Fig. 15.1) that not all scholars are teachers. Therefore, this statement is also true and its truth value is T.

2.2 Types of Statements

In mathematical reasoning, there are two types of statements, namely, simple statements and compound statements.

(a) **Simple Statements:** If the truth value of a statement is not dependent on another statement is said to be a simple statement. A simple statement does not have any logical connectors like "and", "or", "not", "if . . . then" and " . . .if and only if . . ." and cannot be broken down into simpler statements.

(b) **Compound Statements:** A compound statement is a combination of two or more simple statements. Compound statements use logical connectors to combine two simple statements.

2.3 Truth Tables

A truth table is defined as a table that gives the truth values of a compound statement S (p, q, r,...) for every possible combination of truth values of its sub-statement (simple/component/constituent statements) p, q, r.

Construction of Truth Table: In order to construct the truth table for a compound statement, a table consisting of rows and columns is prepared. At the top of the initial columns, the variables denoting the sub-statements or constituent statements are written and then their truth values are written in the last column. The truth value of the compound statement are written on the basis of the truth values of the constituent statements written in the initial columns. If a compound statement is made up of two simple statements, then the number of rows in the truth table will be $2^2$and if it is made up of three simple statements, then the number of rows will be 2^3. In general, if the compound statement is made up of n sub-statements, then its truth table will contain 2^n rows.

2.4 Logical Connectives or Logical Operators

Logical connectives or operators are the symbols (denoting certain phrases or words) that are used to connect simple statements to create compound statements.

In the following table, we list some possible connectives, their symbols, and the nature of the compound statement formed by them.

Table 15.1: Connectives with their symbols and nature of compound statements

Connective	Symbol	Nature of the compound statement formed by using the connective
And	$\wedge$	Conjunction
Or	$\vee$	Disjunction
If......then	$\Rightarrow$ or $\longrightarrow$	Implication or conditional
If and only if (if)	$\Leftrightarrow$ or $\leftrightarrow$	Equivalence or biconditional
Not	~ or ¬	Negation

(a) Conjunction: Compound statements formed by combining any two simple statements using the connective "and" are called the conjunctions.

For example, if p and q are two simple statements, then $p \wedge q$ denotes the conjunction of p and q and is read as "p and q".

(b) Disjunction or Alternation: Compound statements formed by combining any two simple statements using the connective "or" and are called the disjunctions or alterations.

For example, if p and q are two simple statements, then $p \vee q$ denotes the disjunction of p and q is read as "p or q".

(c) Negation: The denial of a statement is called its negation. For if p is a statement, then its negation is denoted by ~p. A negation, unlike a conjunction or a disjunction does not connect two statements but merely modifies a statement. Still, it is called a connective. Any statement p can be negated by writing "it is not the case that ...." or "It is false that....." before p or, if possible, by inserting the word "not" in p.

(d) Implication or Conditional Statements: Conditional statements formed by combining any two simple statements using the connective "if... then" and are called the conditional statements or implications.For example, if p and q are two statements forming the implication 'if p then q', then we denote this implication by "$p \Rightarrow q$" or "$p \longrightarrow q$", where p is the antecedent and q is the consequent.

Truth table for a conditional statement is as follows:

P	**Q**	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

(e) Biconditional Statement: Statements formed by combining any two conditional statements, one converse to the otherand are called the conditional statements or implications. A statement is biconditional if it is the conjunction of two conditional statements (implications), where one statement is the converse of the other.

For example, if p and q are two conditional statements, then the compound statement "$p \Rightarrow q$ and $q \Rightarrow p$" is called a biconditional statement or an equivalence and is denoted by $p \Leftrightarrow q$. Thus, $p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$

Truth table for a biconditional statement: Since $p \Leftrightarrow q$ is the conjunction of $p \Rightarrow q$ and $q \Rightarrow p$, we have the following truth table for $p \Leftrightarrow q$:

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

2.5 Logical Equivalence Statement

When two compound statements, S_1 (p,q,r,...) and S_2 (p,q,r,...), have the same exact truth values, they are said to be logically equivalent. If statements S_1 (p,q,r,...) and S_2 (p,q,r,...) are logically equivalent, it is written as S_1 (p,q,r,...) = S_2 (p,q,r,...).

2.6 Negation of Compound Statements

A compound sentence can be a conjunction, disjunction, implication, or an equivalence. Therefore, the negation of a compound sentence can be as follows:

(a) Negation of conjunction: If p and q are two statements, then $\sim(p \wedge q) \equiv (\sim p \vee \sim q)$

(b) Negation of disjunction: If p and q are two statements, then $\sim(p \vee q) \equiv (\sim p \wedge \sim q)$

(c) Negation of implication: If p and q are two statements, then $\sim(p \Rightarrow q) \equiv (p \wedge \sim q)$

(d) Negation of biconditional statement or equivalence: If p and q are two statements, then $\sim(p \Leftrightarrow q) \equiv (p \wedge \sim q) \vee (q \wedge \sim p)$

2.7 Tautologies and Contradictions

Statement pattern: Any statement where logical connectives are used to combine constituent statements ($\wedge, \vee, \sim, \Rightarrow, \Leftrightarrow$) is called a statement pattern or Well-Formed Formula (WFF).

For example

(a) $p \vee q$

(b) $p \Rightarrow q$

(c) $(p \wedge q) vr \Rightarrow (s \wedge \sim s)$

(d) $(p \Rightarrow q) \Leftrightarrow (\sim q \Rightarrow \sim p)$ are statement patterns.

A statement is also a statement pattern.

Tautology: A statement pattern that is always true irrespective of the truth values of all combinations of its constitutent statements is called a tautology.

In a tautology the last column of its truth table is always T.

Contradiction: A statement pattern that is always false irrespective of the truth values of all combinations of its constitutent statements.

In a contradiction, the last column of its truth table is always F.

The negation of a tautology is a contradiction and vice versa.

2.8 Algebra of Statements

In the previous section, we have seen that statements satisfy many standard results. These results have been codified as laws of algebra of statements. We will enumerate these laws in this section.

Assuming that p, q, and r to be three statements, the following are some laws of algebra of statements:

(a) Idempotent laws:

(A) $p \vee p \equiv p$

(B) $p \wedge p \equiv p$

(b) Commutative laws:

(A) $p \vee q \equiv q \vee p$

(B) $p \wedge q \equiv q \wedge p$

(c) Associative laws:

(A) $(p \vee q) \vee r \equiv p \vee (q \vee r)$

(B) $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$

(d) Distributive laws:

(A) $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$

(B) $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$

(e) De-Morgan's laws:

(A) $\sim(p \wedge q) \equiv \sim p \vee \sim q$ (B) $\sim(p \vee q) \equiv \sim p \wedge \sim q$

(f) Identity laws: If t and c denote a tautology and a contradiction, respectively, then for any statement p, we have

(A) $p \wedge t \equiv p$ (B) $p \vee c \equiv p$

(C) $p \vee t \equiv t$ (D) $p \wedge c \equiv c$

(g) Complement laws:

(A) $p \vee \sim p = t$ (B) $p \wedge \sim p = c$

(C) $\sim t = c$ (D) $\sim c = t$

where t and c denote a tautology and a contradiction, respectively.

(h) Law of contrapositive: $p \Rightarrow q \equiv \sim q \Rightarrow \sim p$

(i) Involution laws: For any statement p, we have $\sim(\sim p) \equiv p$

MASTERJEE CONCEPTS

- While attempting a question, always convert mathematical statements to assertive statements or propositions.
- Theorems are not entirely reliable. There is scope for error in their application..

Akshat Kharaya (JEE 2009, AIR 235)

Illustration 1: Which of the following is not a negation of the statement p: $\sqrt{5}$ is rational? **(JEE MAIN)**

(A) It is not the case that $\sqrt{5}$ is rational (B) $\sqrt{5}$ is not rational

(C) $\sqrt{5}$ is an irrational number. (D) None of these

Sol: (A), (B), and (C) are negations of p.

Illustration 2: Negation of the statement " $\sqrt{5}$ is irrational or 3 is rational" is **(JEE MAIN)**

(A) $\sqrt{5}$ is rational or 3 is irrational. (B) $\sqrt{5}$ is rational and 3 is rational.

(C) $\sqrt{5}$ is rational and 3 is irrational. (D) None of these

Illustration 3: Contrapositive of the statement "If a number is divisible by 9, then it is divisible by 3" is **(JEE MAIN)**

(A) If a number is not divisible by 3, it is not divisible by 9.

(B) If a number is not divisible by 3, it is divisible by 9.

(C) If a number is not divisible by 9, it is not divisible by 3.

(D) None of these.

Sol: Contrapositive of $p \rightarrow q$ is $\sim q \rightarrow \sim p$.

Illustration 4: Converse of the statement "If a number n is even, then n^2 is even" is **(JEE MAIN)**

(A) If a number n^2 is even, then n is even (B) If n is not even, then n^2 is not even

(C) Neither n or n^2 is even (D) None of these

Sol: Converse of $p \rightarrow q$ is $q \rightarrow p$.

Illustration 5: Which of the following statements is not equivalent to $p \leftrightarrow q$? **(JEE MAIN)**

(A) p if and only if q (B) q if and only if p

(C) p is necessary and sufficient for q (D) None of these

Illustration 6: The statement $p \rightarrow (q \rightarrow p)$ is equivalent to **(JEE MAIN)**

(A) $p \rightarrow (p \rightarrow q)$ (B) $p \rightarrow (p \vee q)$

(C) $p \rightarrow (p \wedge q)$ (D) $p \rightarrow (p \leftrightarrow q)$

Sol: $p \rightarrow (q \rightarrow p) \equiv \sim p \vee (q \rightarrow p)$

$\equiv (\sim p) \vee (\sim q \vee p) \equiv (\sim q) \vee (p \vee \sim p) \equiv (\sim q) \vee T = T$

$\therefore$ $p \rightarrow (q \rightarrow p)$ is a tautology.

Also $p \rightarrow (p \vee q) \equiv \sim p \vee (p \vee q) \equiv (\sim p \vee p) \vee q \equiv T \vee q = T$

$\therefore$ $p \rightarrow (p \vee q)$ is also a tautology. Thus, $p \rightarrow (q \rightarrow p)$ is equivalent to $p \rightarrow (p \vee q)$

Illustration 7: The statement $p \rightarrow (q \vee r)$ is not equivalent to **(JEE MAIN)**

(A) $(p \rightarrow q) \vee (p \rightarrow r)$ (B) $p \wedge (\sim q) \rightarrow r$

(C) $p \wedge (\sim r) \rightarrow q$ (D) $p \wedge q \rightarrow (p \wedge r) \vee (q \wedge r)$

Sol: Use logical connectives to conclude.

$p \rightarrow (q \vee r) \equiv (\sim p) \vee (q \vee r)$

$\equiv (\sim p \vee q) \vee (\sim p \vee r) \equiv (p \rightarrow q) \vee (p \rightarrow r)$

Also, $p \rightarrow (q \vee r) \equiv (\sim p) \vee (q \vee r) \equiv (\sim p \vee q) \vee r \equiv (p \wedge (\sim q)) \vee r \equiv p \wedge (\sim q) \rightarrow r$

Interchanging the roles of q and r, we get $p \rightarrow (q \vee r) \equiv p \wedge (\sim q) \rightarrow r \equiv p \wedge (\sim r) \rightarrow q$

For p = T, q = F, r = F, $p \rightarrow (q \vee r)$ is F, but

$(p \wedge q) \rightarrow (p \vee r) \vee (q \wedge r)$ is T. $\therefore$ $p \rightarrow (q \vee r)$ and $p \wedge q \rightarrow (p \wedge r) \vee (q \wedge r)$ are not equivalent.

Illustration 8: Assuming that S is a nonempty subset of R, which of the following statements is the negation of the statement p: There is a rational number $x \in S$ such that $x > 0$? (A) Every rational number $x \in S$ satisfies $x \le 0$. (B) $x \in S$ and $x \le 0 \Rightarrow x$ is not a rational number. (C) There is a rational number $x \in S$ such that $x \le 0$. (D) There is no rational number $x \in S$ such that $x \le 0$. **(JEE MAIN)**

Sol: The statement p can be written as follows:

$P : \exists x \in Q \cap S$ Such that $x > 0$

Negation of p is $\sim P : x \in Q \cap S$ satisfies $x \le 0$

Illustration 9: Consider the following statements:

p: Suman is brilliant, q: Suman is rich, r: Suman is honest

The negation of the "Suman is brilliant and dishonest if and only if Suman is rich" can be expressed as **(JEE MAIN)**

(A) $\sim\left(P\wedge\sim R\right)\leftrightarrow Q$ (B) $\sim p\wedge\left(Q\leftrightarrow\sim R\right)$ (C) $\sim\left(Q\leftrightarrow\left(P\wedge\sim R\right)\right)$ (D) $\sim Q\leftrightarrow\sim p\wedge R$

Sol: $P\wedge\sim R$ stands for Suman is brilliant and dishonest. Thus, $P\wedge\sim R\leftrightarrow Q$ stands for Suman is brilliant and dishonest if and only if Suman is rich. Its negation is $\sim\left(P\wedge\sim R\leftrightarrow Q\right)$ or $\sim\left(Q\leftrightarrow p\wedge\sim R\right)$

Illustration 10: Which of the following is the conditional statement $p\to q$? **(JEE MAIN)**

(A) p is necessary for q (B) p is sufficient for q

(C) p only if q (D) If q then p

Sol: As $p\to q$, the truth of p is sufficient for truth of q.

Illustration 11: Statement-I: $\sim\left(p\leftrightarrow\sim q\right)$ is equivalent to $p\leftrightarrow q$ **(JEE MAIN)**

Statement-II: $\sim\left(p\leftrightarrow\sim q\right)$ is a tautology

Sol: Use truth table.

Table for basic logical connectives.

p	**q**	**~q**	**$p\leftrightarrow\sim q$**	**$\sim(p\to\sim q)$**	**$(p\leftrightarrow q)$**
T	T	F	F	T	T
T	F	T	T	F	F
F	T	F	T	F	F
F	F	T	F	T	T

Note that $\sim\left(p\to\sim q\right)$ is not tautology. $\therefore$ Statement-II is false.

From table $\sim\left(p\leftrightarrow\sim q\right)$ is equivalent to $\left(p\leftrightarrow q\right)$. Thus, statement-I is true.

Illustration 12: Let p, q, and r be the statements. **(JEE MAIN)**

p: X is a rectangle; q: X is a square; r: $p\to q$

Statement-I: $p\to r$ is a tautology. Statement-II: A tautology is equivalent to T.

Sol: $p\to r\equiv p\to(p\to q)$

$\equiv(\sim p)\vee(p\to q)\equiv(\sim p)\vee[(\sim p)\vee q]\equiv[(\sim p)\vee(\sim p)]\vee q\equiv(\sim p)\vee q\equiv p\to q\equiv r$

Thus, statement-I is not a tautology.

Illustration 13: $(p\wedge\sim q)\wedge(\sim p\vee q)$ is **(JEE MAIN)**

(A) A contradiction (B) A tautology (C) Either (A) or (B) (D) Neither (A) nor (B)

Sol:

p	**q**	**~p**	**~q**	**$p\wedge\sim q$**	**$\sim p\vee q$**	**$(p\wedge\sim q)\wedge(\sim p\vee q)$**
T	T	F	F	F	T	F

T	F	F	T	T	F	F
F	T	T	F	F	T	F
F	F	T	T	F	T	F

Clearly, $(p\wedge \sim q)\wedge(\sim p\vee q)$ is a contradiction.

3. BOOLEAN ALGEBRA

Introduction: Boolean algebra is a tool for studying and applying mathematical logic which was originated by the English mathematician George Boolean. In 1854, he wrote a book "An investigation of the law of thought", which developed a theory of logic using symbols instead of word. This more algebraic treatment of subject is now called Boolean algebra.

Definition: A nonempty set B together with two binary operations $\vee$ and $\wedge$ and a unary operation ʹ is said to be a Boolean algebra if the following axioms hold:

(a) For all x, y, $z\in B$

(i) $x \vee y\in B$ (Closure property for $\vee$)

(ii) $x \wedge y\in B$ (Closure property for $\wedge$)

(b) For all x, $y\in B$

(i) $x \vee y = y \vee x$ (Commutative law for $\vee$)

(ii) $x \wedge y = y \wedge x$ (Commutative law for $\wedge$)

(c) For all x, y, and z in B,

(i) $(x \vee y) \vee z = x \vee (y \vee z)$ (Associative law of $\vee$)

(ii) $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ (Associative law of $\wedge$)

(d) For all x, y, and z in B,

(i) $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ (Distributive law of $\vee$ over $\wedge$)

(ii) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ (Distributive law of $\wedge$ over $\vee$)

(e) There exist elements denoted by 0 and 1 in B such that for all $x\in B$

(i) $x \vee 0 = x$ (0 is identity for $\vee$)

(ii) $x \wedge 1 = x$ (1 is identify for $\wedge$)

(f) For each $x\in B$, there exists an element denoted by xʹ, called the complement or negation of x in B such that

(i) $x \vee x' = 1$

(ii) $x \wedge x' = 0$ (Complement laws)

MASTERJEE CONCEPTS

- One can use the symbols + (addition) and (multiplication) instead of $\vee$ and $\wedge$. These are just binary operators and has nothing to do with ordinary addition and multiplication of numbers.
- We sometimes designate a Boolean algebra by (B, `$\vee$', '$\wedge$', ' '', 0, 1) in order to emphasize its six parts; namely the set B, the two binary operations $\vee$ and $\wedge$, the unary operation ' and the two special elements 0 and 1. These special elements are called the zero element and the unit element. These are two specific elements belonging to set B, satisfying certain properties. These symbols 0 and 1 have nothing to do with numbers zero and one.
- For the set S of all logical statements, the operations + and play the roles of $\vee$ and $\wedge$, respectively. The tautology t and the contradiction c play the roles of 1 and 0, and the operation `~' plays the role of''
- For P(A), the set of all subsets of a set A, the operations $\cup$ and $\cap$ play the roles of `$\vee$' and `$\wedge$', A and ϕ play the role of 1 and 0, and complementation plays the role of ''.

Rohit Kumar (JEE 2012, AIR 79)

3.1 Boolean Functions

Definition: Any expression consisting of combinations by disjunctions ($\vee$) and conjunctions ($\wedge$) of finite number of elements of a Boolean algebra B is called a Boolean function. For example, $x \wedge x'$, $a \wedge b'$, $[a \wedge (b \vee c')] \vee (a' \wedge b' \wedge c)$.

Let B = {a, b, c...} be a Boolean algebra. A constant means any symbol such as 0 and 1 which represents a specified element of B. A variable means a symbol which represents an arbitrary element of B. If in the expression $x' \vee (y \wedge z)$,we replace $\vee$ by + and $\wedge$ by ., we get x'+y.z. Here x' and $y \wedge z$ are called monomials and the whole expression $x' \vee (y \wedge z)$ is called a polynomial.

3.2 Switching Circuits

The most common functional application of the Boolean algebra is the electrical switching system that involve two-state devices. The simplest possible example of such a device is an ordinary ON-OFF switch.

A switch here refers to a device which controls an electric circuit by allowing or inhibiting the flow of current through the circuit. The switch can either be in a closed or an open state (ON or OFF). In the first case, the current flows and in the second, the current does not flow.

The symbols a, b, c, p, q, r, x, y, z ....etc., will denote switches in a circuit. Switches are generally interconnected either serially or in parallel.

(a) Series: If two switches a and b are serially connected, the current can pass only when both are in closed state, and not otherwise Fig. 15.2 shows this circuit.

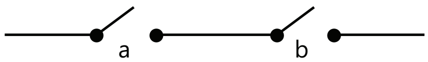

Figure 15.2

(b) Parallel: If two switches a and b are connected `in parallel' the current flows when any one or both are closed, and not otherwise. Fig. 15.3 represents this circuit given by $a \vee b$.

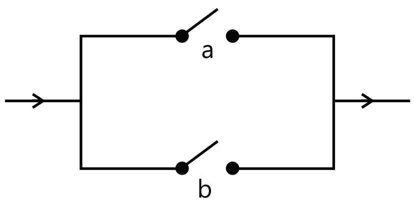

Figure 15.3

If two switches in a circuit are open (closed) simultaneously, we will represent them by the same letter. Again, if one switch is open if the other is closed, we will represent them by a and a'.

The value of a closed switch (on) is equal to 1 and when it is open (off) the value equal to 0.

As open switch r is indicated in Fig. 15.4.

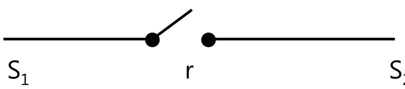

Figure 15.4

A closed switch r is indicated in Fig. 15.5.

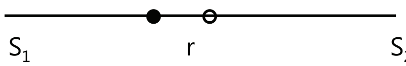

Figure 15.5

3.3 Boolean Operations on Switching Circuits

(a) Boolean Multiplication: In Fig. 15.6, the two serially connected switches, r and s, will perform the operation of Boolean multiplication.

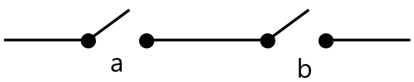

Figure 15.6

Evidently, the current will pass from point S_1 to S_2 only when both r and s are closed.

r	**s**	**r ∧ s**
1	1	1
1	0	0
0	1	0
0	0	0

As seen in the above truth table, the operation is true only in one of the four cases i.e., when both the switches are closed.

(b) Boolean Addition: In Fig. 15.7, the two switches connected in parallel will perform the operation of addition. The circuit shows that the current will pass when either or both the switches are closed.

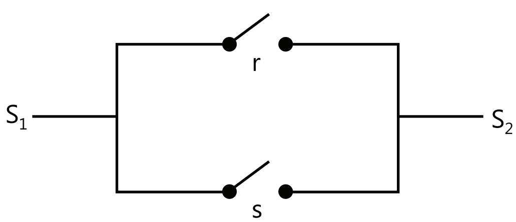

Figure 15.7

R	**S**	**r ∨ s**
1	1	1
1	0	1
0	1	1
0	0	0

As seen in the above truth table, the operation is not true only in one of the four cases i.e., when both r and s are open.

(c) Circuits with Composite Operations:

(i) Circuit showing: $r \wedge (s \vee q)$

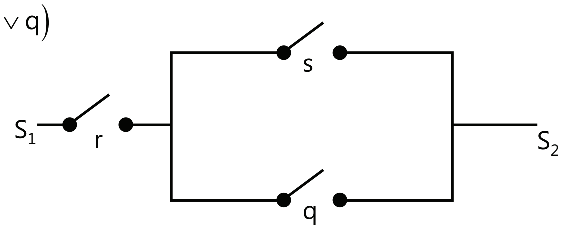

Figure 15.8

(ii) Circuit showing: $r \vee (s \wedge q)$

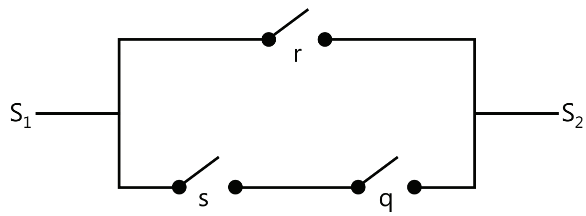

Figure 15.9

(iii) Circuit showing: $(r \vee s) \wedge (r \vee q)$

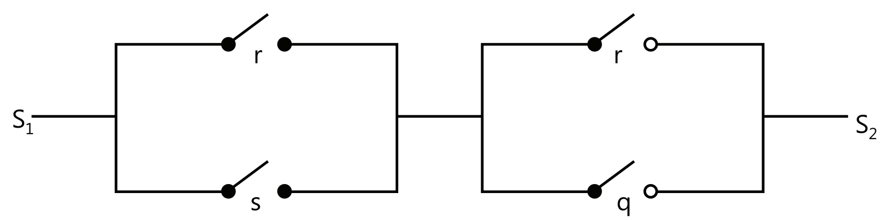

Figure 15.10

(iv) Circuit for: $(r \vee s) \wedge q \wedge (u \vee \nu \vee \omega)$

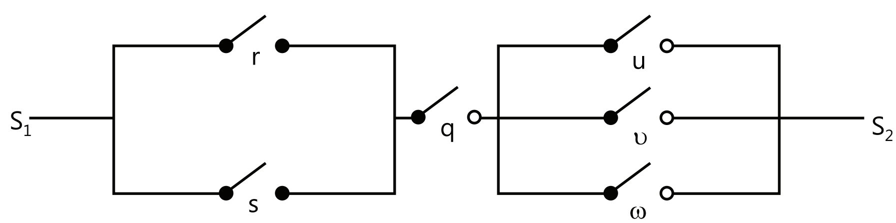

Figure 15.11

3.4 Simplification of Circuits

A simple circuit is one which is least complicated, is cheap, and is efficient. For this, a number of factors like the cost of equipment, positioning and number of switches, types of material used etc., will have to be considered. In the context of Boolean algebra, circuits can be simplified reducing the number of switches.

This can be achieved by using different properties of Boolean algebra. For example, consider the circuits given by $(a \wedge b) \vee (a \wedge c)$. This is represented by the circuit shown in Fig. 15.12.

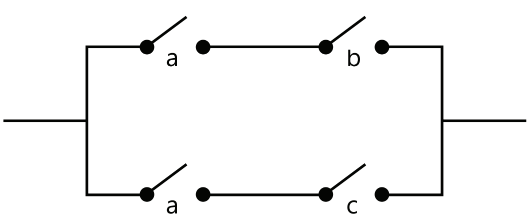

Figure 15.12

Since $(a \wedge b) \vee (a \wedge c) = a \wedge (b \vee c)$.

Therefore the circuit could be simplified to as shown in Fig. 15.13.

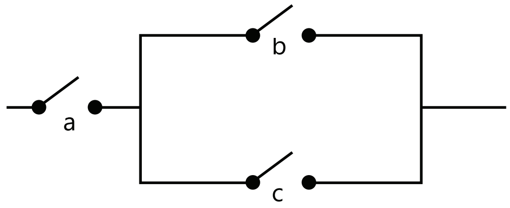

Figure 15.13

MASTERJEE CONCEPTS

- While attempting a question, converting Boolean math into a statement will make it easier to solve the problem.
- However,answers should be cross-checked by solving the question in both statement and mathematical forms.

Ravi Vooda (JEE 2009, AIR 71)

Which of the following is correct?

Illustration 14: (A) $p \vee p' = 0$ (B) $p \wedge p' = 1$ (C) $p \vee p' = 1$ (D) None of these **(JEE MAIN)**

Sol: (C) $p \vee p' = 1$

Illustration 15: (A) $p \wedge p' = 1$ (B) $p \wedge p' = 0$ (C) $p \vee p' = 0$ (D) None of these **(JEE MAIN)**

Sol: (B) $p \wedge p' = 0$

Illustration 16: (A) $a \vee 1 = a$ (B) $a \vee 1 = 1$ (C) $a \vee 1 = 0$ (D) None of these **(JEE MAIN)**

Sol: (B) $a \vee 1 = 1$

Illustration 17: (A) $x \vee x' \wedge x = x'$ (B) $x \vee x' \wedge x = x$ (C) $x \vee x' \wedge x = 1$ (D) None of these **(JEE MAIN)**

Sol: (B) $x \vee x' \wedge x = (x \vee x') \wedge (x \vee x) = 1 \wedge (x \vee x) = 1 \wedge x = x$

Illustration 18: (A) $x \vee x' \wedge y = x$ (B) $x \vee x' \wedge y = y$ (C) $x \vee x' \wedge y = x \vee y$ (D) None of these **(JEE MAIN)**

Sol: (B) $x \vee x' \wedge y = (x \vee x') \wedge (x \vee y) = 1 \wedge (x \vee y) = x \vee y$

Illustration 19: For the circuits shown below, the Boolean polynomial is **(JEE MAIN)**

(A) $(\sim p \vee q) \vee (p \vee \sim q)$ (B) $(\sim p \wedge p) \wedge (q \wedge q)$

(C) $(\sim p \wedge \sim q) \wedge (q \wedge p)$ (D) $(\sim p \wedge q) \vee (p \wedge \sim q)$

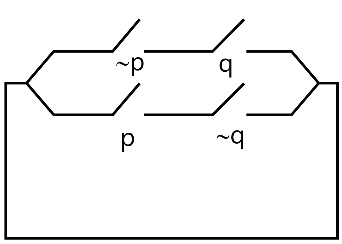

Figure 15.14

Sol: (D) For the given circuit, the Boolean polynomial is

$(\sim p \wedge q) \vee (p \wedge \sim q)$

4. LOGIC GATES

A logic gate performs fundamental logical operations in a Boolean function. Most logic gates take an input of two binary values, and output a single value of a 1 or 0.

The different types of logic gates are outlined below:

(a) AND: It is the Boolean function defined by

$f(x_1, x_2) = x_1 \wedge x_2; x_1, x_2 \in \{0,1\}$. It is shown in the Fig. 15.15.

Input		Output
x_1	x_2	$x_1 \wedge x_2$
1	1	1
1	0	0
0	1	0
0	0	0

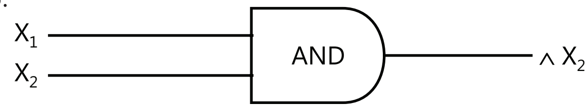

Figure 15.15

(b) OR: It is the Boolean function defined by

$f(x_1, x_2) = x_1 \vee x_2; x_1, x_2 \in \{0,1\}$. It is shown in Fig. 15.16.

Input		Output
x_1	x_2	$x_1 \vee x_2$
1	1	1
1	0	1
0	1	1
0	0	0

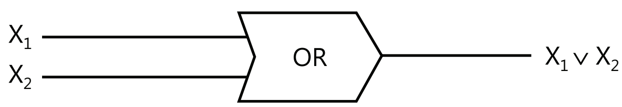

Figure 15.16

(c) NOT: It is the Boolean function defined by $f(x) = x', x \in \{0,1\}$

It is shown in Fig. 15.17.

Input	Output
X	**x′**
1	0
0	1

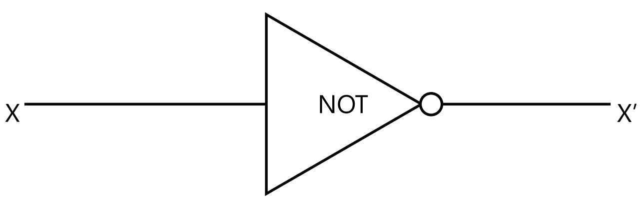

Figure 15.17

Combinational Circuit: In a combinational circuit there is a combination of different logic gates in a circuit, yet the output (s) is uniquely defined for each combination of inputs (x_1, x_2, and x_3). See Fig. 15.18.

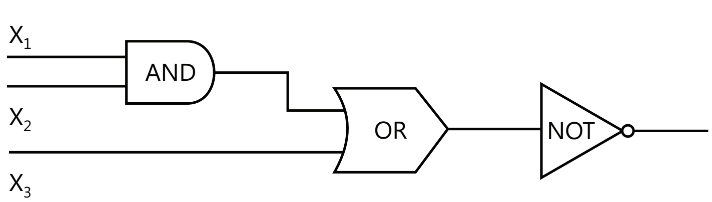

Figure 15.18

If the output is not uniquely defined for the inputs then the circuit is not combinational. In Fig. 15.19, if $x_1 = 1$, $x_2 = 0$, then the inputs to the AND gate are 1 and 0 and so the output of the AND gate is '0' (Minimum of 1 and 0). This is the input of NOT gate which gives the output s=1. But the diagram states that $x_2 = s$ i.e., 0=1, a contradiction.

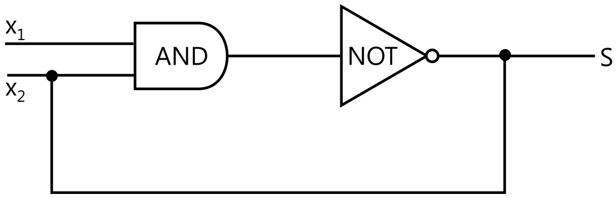

Figure 15.19

Two Combinational Circuits: Circuit having the input $x_1, x_2, ...x_n$ and a single output are said to be combinational circuit if the circuits receive the same input and they produce the same output i.e., if the input/output tables are identical.

Illustration 20: An AND gate is the Boolean function defined by **(JEE MAIN)**

(A) $f(x_1, x_2) = x_1.x_2$, $x_1, x_2 \in \{0,1\}$

(B) $f(x_1, x_2) = x_1 + x_2$ $x_1, x_2 \in \{0,1\}$

(C) $f(x_1, x_2) = x_1$, $x_1, x_2 \in \{0,1\}$

(D) $f(x_1, x_2) = x_2$ $x_1, x_2 \in \{0,1\}$

Sol: (A) It is the definition.

Illustration 21: The output of the circuit is **(JEE MAIN)**

(A) $(x_2 + x_3).[(x_1.x_2).x'_3]$

(B) $(x_2 + x'_3).[(x_1.x_2).x'_3]$

(C) $(x_2 + x_3) + [(x_1.x_2).x'_3]$

(D) $(x_2.x_3) + [(x_1.x_2).x'_3]$

(E) $(x_1 + x_3).[(x_1.x_2).x'_3]$

Sol: (A) Here, x_1, x_2 and x'_3 are connected by 'AND'

$\therefore$ Combinatorial circuit is $(x_1.x_2).x'_3$. Also, x_2 and x_3 are connected by 'OR'

$\therefore$ Combinatorial circuit is $(x_2 + x_3)$. These two combinatorial circuit are connected by 'AND'

$\therefore$ Output is $(x_2 + x_3).[(x_2.x_2).x'_3]$

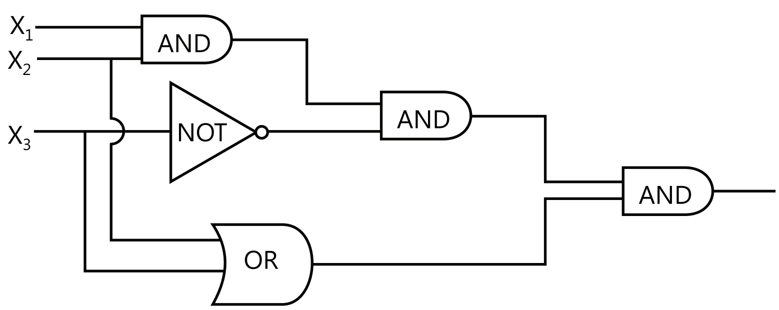

Figure 15.20

5 LOGICAL REASONING

Questions that test logical reasoning require innovative and broad thinking. There are no tried and tested methods to solve these questions. Given below are a few illustrations of these kind of questions.

Illustration 22: The missing number in the given figure is: **(JEE MAIN)**

Sol: Answer is 44

In the First diagram, the difference is 6, 8, and 14.

In the second diagram, the difference of is 3, 4, and 7.

So suppose unknown is x. difference of three should be 12, 16, and 28.

60-x=16; 72-x=28, hence x=44

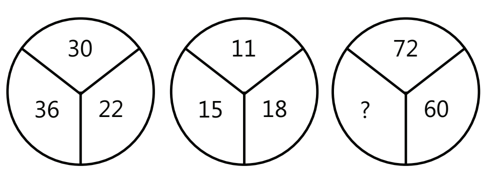

Figure 15.21

Illustration 23: Find the next number in the series. **(JEE MAIN)**

1, 2, 9, 28, 65,_____.

Sol: $1 = 0^3 + 1$; $2 = 1^3 + 1$; $9 = 2^3 + 1$; $65 = 4^3 + 1$;

Hence next term would be $5^3 + 1 = 126$

Illustration 24: It was Wednesday on July 15, 1964. What was the day on July 15, 1965? **(JEE MAIN)**

Sol: This question can be a little tricky as some years (leap years) have 366 days. Since 1964 is divisible by 4, 1964 is a leap year. However, the first date is July 15, 1964 and July comes after February. Hence, there are 365 days between the given dates. That translates to 52 weeks and 1 day. So July 15, 1965 will be 1 day after Wednesday i.e., Thursday.

Note: If February 29th was included in the given range then you have to consider 366 days.

Illustration 25: Arrange the number 1 through 9 on a tic-tac-toe board such that the numbers in each row, column, and diagonal add up to 15 **(JEE MAIN)**

Sol: It can be observed that '5' is the most recurrent number. So, this should be in the middle. Now, try to solve by filling some places with random numbers. Remember, there is no unique solution for this. One of the solutions is given below.

4	3	8
9	5	1
2	7	6

6. STATISTICS

6.1 Introduction

Statistics plays a significant role in engineering. However, for JEE Main we will only study mean, mode, median, variance, standard deviation, and mean deviation. It is quite easy as it has simple, straight formulae to remember but it involves a fair amount of calculations. The key to score well here is to remember the exact formula and not commit mistakes while doing the calculations.

6.2 Mean

Mean, as a stand-alone word, generally refers to arithmetic mean, unless specified.

Arithmetic Mean

(a) **Individual Series:** For a set of values $x_1, x_2, \ldots\ldots x_n$ the arithmetic mean usually denoted by $\bar{x}$ and is given by

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

(b) **Discrete Series:** For a set of values $x_1, x_2, \ldots\ldots x_n$ with corresponding frequencies $f_1, f_2 \ldots f_n$, the arithmetic mean

$\bar{x}$ is given by $\bar{x} = \dfrac{f_1 x_1 + f_2 x_2 \ldots + f_n x_n}{f_1 + f_2 \ldots + f_n} = \dfrac{1}{N}\sum_{i=1}^{n} f_i x_i$. Where $N = \sum_{i=1}^{n} f_i$

(c) Continuous Series: In case the data is provided in the form of class intervals, we have to try to obtain an approximate value of the mean. The method of approximation is by determining mid-value of each class interval. If $x_1, x_2, ... x_n$ are the mid values of the class intervals having corresponding frequencies $f_1, f_2 ... f_n$, then we apply the same formula as in discrete series

i.e., $\bar{x} = \frac{1}{N}\sum_{i=1}^{n} f_i x_i$, $\quad N = \sum_{i=1}^{n} f_i$

MASTERJEE CONCEPTS

If $\bar{x}$ is the mean of $x_1, x_2, ... x_n$, then the mean of $ax_1, ax_2, ... ax_n$ is $a\bar{x}$, where a is any number other than zero.

Anvit Tawar (JEE 2009, AIR 9)

Illustration 26: If the mean of a set of observations $x_1, x_2, ...x_{10}$ is 20, the mean of $x_1 + 4$, $x_2 + 8$, $x_3 + 12$ .... $x_{10} + 40$ is **(JEE MAIN)**

(A) 34 (B) 42 (C) 38 (D) 40

Sol: (B) Use $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Required mean $= \frac{1}{10}[(x_1, x_2, ... x_{10}) + (4 + 8 + ...40)] = \frac{1}{10}(x_1 + ... + x_{10}) + \frac{4}{10}[(1 + 2 + ...10)] = 20 + \frac{4 \times 10 \times 11}{10 \times 2} = 42$

Illustration 27: The arithmetic mean (A.M.) of the observations 1.3.5, 3.5.7, 5.7.9 ... $(2n-1)(2n+1)(2n+3)$ is **(JEE MAIN)**

(A) $2n^3 + 6n^2 + 7n - 2$ (B) $n^3 + 8n^2 + 7n - 2$ (C) $2n^3 + 5n^2 + 6n - 1$ (D) $2n^3 + 8n^2 + 7n - 2$

Sol: Use $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$$A.M. = \frac{1}{n}[1.3.5 + 3.5.7 + ... + (2n-1)(2n+1)(2n+3)] = \frac{1}{n}\left[\sum_{r=1}^{n}(2r-1)(2r+1)(2r+3)\right] = \frac{1}{n}\sum_{r=1}^{n}(8r^3 + 12r^2 - 2r - 3)$$

$$= \frac{1}{n}\left[8\left(\frac{n(n+1)}{2}\right)^2 + 12\frac{n(n+1)(2n+1)}{6} - 2\frac{n(n+1)}{2} - 3n\right] = 2n(n+1)^2 + 2(n+1)(2n+1) - (n+1) - 3 = 2n^3 + 8n^2 + 7n - 2.$$

Illustration 28: If the arithmetic mean of 7 consecutive integers starting with a is m, the arithmetic mean of 11 consecutive integers starting with a+2. **(JEE MAIN)**

(A) 2a (B) 2m (C) a +4 (D) m+4 (e) a +m+2

Sol: 7 consecutive integers starting from a are a ,a + 1,a + 2,..........,a + 6

A.M. of these numbers is

(Given) $m = \frac{(a) + (a+1) + (a+2) + ... + (a+6)}{7} \Rightarrow m = \frac{7a + 21}{7} = a + 3$ ... (i)

Now, 11 consecutive integers starting from $a+2, a+3, ... a+12$

A.M. of these numbers is Mathematical Reasoning and Statistics = $a+\frac{2+3+...+12}{11}=a+7=m+4$

6.3 Median

Median refers to the middle value in a distribution.

For individual series, the median can be found by arranging the data in ascending or descending order of magnitude. In case the total number of values is odd, median = size of $\frac{n+1}{2}$th item. In case the total number of values is even median = average of $\frac{n}{2}$th and $\frac{n+2}{2}$th observation. In case of discrete frequency distribution, find $\frac{N}{2}$, where N $=\sum_{i=1}^{n} f_i$. Find the cumulative frequency (c.f.) just more than $\frac{N}{2}$. The corresponding value of x is median. In case of continuous distribution, the class corresponding to c.f. just more than $\frac{N}{2}$ is called the median class and the median $= l+\frac{h}{f}\left(\frac{N}{2}-C\right)$

Where l = lower limit of the median class; f = frequency of the median class; h = width of the median class; C = c.f. of the class preceding to the median class and $N=\sum_{i=1}^{n} f_i$

Illustration 29: The median from the following data is **(JEE MAIN)**

Wages/week (Rs.)	No. of workers
50-59	15
60-69	40
70-79	50
80-89	60
90-99	45
100-109	40
110-119	15

(A) 83.17 (B) 84.08 (C) 82.17 (D) 85.67

Sol: (B) Use median formula.

Wages/week (Rs.)	No. of workers	c.f.
50-59	15	15
60-69	40	55
70-79	50	105
Median class 80-89	60	165
90-99	45	210
100-109	40	250
110-119	15	265

Median = size of $\left(\frac{N}{2}\right)$ item$=\frac{265}{2}=132.5$th item

Median lies in the class 80-89. But the true limit of the class is 79.5 - 89.5

Here, N = 265, C = 105, f = 60, h = 10 and l = 79.5

Median= $l+\frac{h}{f}\left(\frac{N}{2}\right)=79.5+\frac{10}{60}(132.5-105)=84.08$

Illustration 30: The median of 21 observations is 40 and if the observations greater than the median are increased by 6 then the median of the new data will be **(JEE MAIN)**

(A) 40 (B) 46 (C) 46+40/21 (D) 46-40/21

Sol: (A) Median is not altered if we increase the value of terms that come after the median.

Illustration 31: Given the following frequency distribution with some missing frequencies, **(JEE MAIN)**

Class	Frequency
10-20	180
20-30	----
30-40	34
40-50	180
50-60	136
60-70	----
70-80	50

if the total frequency is 685 and median is 42.6, then the missing frequencies are

(A) 81, 24 (B) 80, 25 (C) 82, 23 (D) 832, 22

Sol: (C) Use median formula

Let the missing frequencies be f_1 and f_2, respectively.

Class	Frequency (f)	Cum. Freq. (c.f.)
10-20	180	180
20-30	f_1	$180 + f_1$
30-40	34	$214 + f_1$
40-50	180	$394 + f_1$
50-60	136	$530 + f_1$
60-70	f_2	$530 + f_1 + f_2$
70-80	50	$580 + f_1 + f_2$
Total	685	

We have $508+f_1+f_2=685$ $\Rightarrow f_1+f_2=105$

As median 42.6 lies in the class 40-50, median class is 40-50. We have that $l=40$, f=180, c= 214 + f_1 and h=10

Now, Median= $l+\frac{h}{f}\left(\frac{N}{2}-C\right)$

$$\Rightarrow 42.6 = 40 + \frac{10}{180}\left(\frac{685}{2} - 214 - f_1\right) \Rightarrow 2.6 \times 18 = 342.5 - 214 - f_1 \Rightarrow f_1 = 81.7$$

As frequency of an item is always a whole number, we take $f_1 = 82$. $\therefore f_2 = 105 - 82 = 23$.

6.4 Mode

Mode is that value which repeats itself the maximum number of times in a set of data. For individual series, the value that occurs with greatest frequency is the mode. In case of discrete series, mode can be determined just by inspection. In case of continuous series,

$$\text{Mode} = l + \frac{f_1 - f_0}{2f_1 - f_2 - f_0} \times h$$

where l = lower limit of the modal class i.e., the class having maximum frequency; f_1 = frequency of the modal class; f_0 = frequency of the class preceding the modal class, f_2 = frequency of the class succeeding the modal class and h = width of the modal class.

When there are two or more values with greatest frequency, the mode can computed with the help of the following equation: Mode = 3 median – 2 mean.

MASTERJEE CONCEPTS

There may be multiple modes if there are two or more values with greatest frequency. We do not need to take average of both modes.

Rohit Kumar (JEE 2012, AIR 79)

Illustration 32: A market with 3900 operating firms has the following distribution for firms arranged according to various income groups of workers **(JEE MAIN)**

Income Group	No. of firms
150-300	300
300-500	500
500-800	900
800-1200	1000
1200-1800	1200

If a histogram for the above distribution is constructed, the highest bar in the histogram would correspond to the class

(A) 500-800 (B) 1200-1800 (C) 800-1200 (D) 150-300

Sol: (B) Maximum number of firms is 1200 in the group of 1200-1800. Therefore, highest bar would correspond to 1200-1800 class.

Illustration 33: Mode of 7, 6, 10, 7, 5, 9, 3, 7, 5 is **(JEE MAIN)**

(A) 6 (B) 3 (C) 5 (D) 7

Sol: (D) Mode is the value that occurs most frequently. So, mode =7 (with 3 appearances).

6.5 Mean Deviation

In a set of values, $x_1, x_2 ...... x_n$, the mean deviation (M.D.) about a point A is given by

M.D. $= \frac{1}{n}\sum |x_i - A|$. In the case of discrete or continuous series M.D. $= \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - A|, N = \sum_{i=1}^{n} f_i$

M.D. is least when taken from the median.

MASTERJEE CONCEPTS

- Mean deviation is deviation of all values from the mean. Since mean is the average so it will always be zero. Hence, absolute mean deviation is considered without sign.
- Generally, mean deviation refers to absolute mean deviation from mean.

Vaibhav Gupta (JEE 2009, AIR 54)

Illustration 34: The mean deviation from mean of the observations $a, a+d, a+2d, ..., a+2nd$ is **(JEE MAIN)**

(A) $\frac{n(n+1)d^2}{3}$ (B) $\frac{n(n+1)}{2}d^2$ (C) $a + \frac{n(n+1)d^2}{2}$ (D) None of these

Sol: $\bar{X} = \frac{1}{2n+1}\left[a + (a+d) + ... + (a+2nd)\right] = \frac{1}{2n+1}\left[(2n+1)a + d(1+2+...+2n)\right] = a + d\frac{2n}{2}\frac{(1+2n)}{2n+1} = a + nd$

M.D. from mean $= \frac{1}{2n+1}2|d|(1+2+...+n) = \frac{n(n+1)|d|}{(2n+1)}$

Illustration 35: The mean deviation from the mean for the set of observations -1, 0, 4 is **(JEE MAIN)**

(A) $\sqrt{\frac{14}{3}}$ (B) 2 (C) $\frac{2}{3}$ (D) None of these

Sol: Mean $= \frac{-1+0+4}{3} = 1$. Hence M.D. (about mean)$= \frac{|-1-1| + |0-1| + |4-1|}{3} = 2$.

6.6 Standard Deviation and Variance

In statistics and probability theory, the standard deviation (represented by the Greek letter sigma, σ) shows the extent of variation or dispersion of data from the average. Variance is represented by σ^2 and it is just square of standard deviation.

In the case of individual series, variance σ^2 is given by $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

If $x_1, x_2, ..., x_n$ occur with frequency $f_1, f_2, ... f_n$, respectively, then

$$\sigma^2(\text{variance}) = \frac{1}{N}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2 = \frac{1}{N}\sum f_i x_i^2 - \left(\frac{1}{N}\sum f_i x_i\right)^2$$

And standard deviation $= +\sqrt{\text{variance}}$.

MASTERJEE CONCEPTS

- Standard deviation as well as variance are always positive.
- A low standard deviation indicates that the data points in a statistical distribution are close to the average or mean value (also called expected value); a high standard deviation indicates that the data points diverge greatly average value.

Ravi Vooda (JEE 2009, AIR 71)

Std. dev. (X+Y) ≠ Std. dev. (X) + std. dev. (Y). There is no effect of change of origin on standard deviation i.e., if

$d_i = x_i - A$ then $\sigma_x = \sigma_d$. If $d_i = \frac{x_i - A}{h}$ then; $\sigma_x = |h|\sigma_d$ equivalently

$$\sigma_x^2 = h^2\left[\frac{1}{N}\sum f_i d_i^2 - \left(\frac{1}{N}\sum f_i d_i\right)^2\right];\ \text{var}(X) \equiv \sigma_x^2$$

$$\text{var}(X_1 + X_2) \equiv \text{var}(X_1) + \text{var}(X_2);\ \text{var}(cX_1) \equiv c^2\,\text{var}(X_1)$$

$$\text{std. dev.}(X+Y) = \sqrt{\text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X,Y)}.$$

Illustration 36: If $\sum_{i=1}^{18}(x_i - 8) = 9$ and $\sum_{i=1}^{18}(x_i - 8)^2 = 45$ then the standard deviation of $x_1, x_2 \ldots x_{18}$ is **(JEE MAIN)**

(A) 4/9 (B) 9/4 (C) 3/2 (D) None of these

Sol: Let $d_i = x_i - 8$, but $\sigma_x^2 = \sigma_d^2 = \frac{1}{18}\sum d_i^2 - \left(\frac{1}{18}\sum d_i\right)^2 = \frac{1}{18}\times 45 - \left(\frac{9}{18}\right)^2 = \frac{5}{2} - \frac{1}{4} = \frac{9}{4}$; Therefore $\sigma_x = 3/2$.

Illustration 37: If the standard deviation of $x_1, x_2, \ldots, x_n$ is 3.5, then the standard deviation of $-2x_1 - 3, -2x_2 - 3, \ldots, -2x_n - 3$ is **(JEE MAIN)**

(A) -7 (B) -4 (C) 7 (D) 1.75

Sol: We know that if $d_i = \frac{x_i - A}{h}$ then $\sigma_x = |h|\sigma_d$. In this case, $-2x_i - 3 = \frac{x_i + 3/2}{-1/2}$ so h=-1/2.

Thus $\sigma_d = \frac{1}{|h|}\sigma_x = 2 \times 3.5 = 7$.

Illustration 38: The variance of first 20 natural numbers is **(JEE MAIN)**

(A) 133/4 (B) 379/12 (C) 133/2 (D) 399/4

Sol: Use the formula $\sigma^2 = \frac{1}{N}\sum d_i^2 - \left(\frac{1}{N}\sum d_i\right)^2$

$$\sigma^2 = \frac{1}{20}\left[1^2 + 2^2 + \ldots + 20^2\right] - \left[\frac{1}{20}(1 + 2 + \ldots + n)\right]^2 = \frac{1}{20}\frac{20\times 21\times(2\times 20+1)}{6} - \left[\frac{1}{20}\frac{20\times 21}{2}\right]^2 = \frac{7\times 41}{2} - \frac{441}{4} = \frac{133}{4}.$$

In fact, the variance of first n-natural numbers is $\frac{n^2-1}{12}$.

Illustration 39: Suppose a population A has 100 observations 101, 102, ... 200 and another population B has 100 observations 151, 152,...,250. If V_A and V_B represent the variances of the two populations, respectively, then $\frac{V_A}{V_B}$ is **(JEE MAIN)**

(A) 1 (B) $\frac{9}{4}$ (C) $\frac{4}{9}$ (D) $\frac{2}{3}$

Sol: Find the variance for both the series individually.

$\bar{X}$ for population A = $\frac{101+102+....+200}{100} = \frac{(100/2)\left[101+200\right]}{100} = 150.5$

$\bar{X}$ for population B = $\frac{151+152+....+250}{100}$

$= \frac{(100/2)\left[151+250\right]}{100} = 200.5; V_A = \frac{(101-150.5)^2+(102-150.5)^2+.....+(200-150.5)^2}{100}$

$= \frac{(49.5)^2+(48.5)^2+...+(0.5)^2+(0.5)^2+(1.5)^2+...+(49.5)^2}{100}$

$V_B = \frac{(151-200.5)^2+...+(250-200.5)^2}{100} = \frac{(49.5)^2+.....+(0.5)^2+(0.5)^2+....+(49.5)^2}{100} \Rightarrow \frac{V_A}{V_B} = 1.$

Illustration 40: What is the standard deviation of the following series? **(JEE MAIN)**

Measurements	0-10	10-20	20-30	30-40
Frequency	1	3	4	2

(A) 81 (B) 7.6 (C) 9 (D) 2.26

Sol: (C)

Class	f_1	y_1	**d = y_1 – A, A = 25**	$f_i d_i$	$f_i d_i^2$
0-10	1	5	-20	-20	400
10-20	3	15	-10	-30	300
20-30	4	25	0	0	0
30-40	2	35	10	20	200
Total	10			-30	900

$\sigma^2 = \frac{\sum f_i d_i^2}{\sum f_i} - \left(\frac{\sum f_i d_i}{\sum f_i}\right)^2 = \frac{900}{10} - \left(\frac{-30}{10}\right)^2; \sigma^2 = 90-9=81 \Rightarrow \sigma = 9.$

6.7 Quantile

Quantile is a class of values of a variate which divides an ordered dataset into a certain number of equal proportions. n quantiles of a dataset means that the data has been divided into n equal-sized data subsets. Following are the commonly used quantiles for statistical analysis

100-quantiles are called percentiles (denoted by p)

10-quantiles are called deciles (denoted by D), 4-quantiles are called quartiles (denoted by Q)

2-quartile is called median. For data divided into quartiles, the median is the middle value of the dataset.

$\therefore \text{Median} = \frac{1}{2}(n+1)\text{th}$ value, where n is the number of data values in the dataset.

The lower quartile (Q_1) is the median of the lower half of the dataset.

$\therefore Q_1 = \frac{1}{4}(n+1)\text{th}$ value, where n is the number of data values in the dataset.

The upper quartile (Q_3) is the median of the upper half of the dataset.

$\therefore Q_3 = \frac{3}{4}(n+1)\text{th}$ value, where n is the number of data values in the dataset.

Illustration 41: Find the median, lower quartile, upper quartile, and interquartile range of the following dataset of scores. **(JEE MAIN)**

18 20 23 20 23 27 24 23 29

Sol: Arrange the values in ascending order of magnitude:

18 20 20 23 23 23 24 27 29

There are 9 values in the dataset. $\therefore n = 9$

Now, median $= \left(\frac{n+1}{2}\right)\text{th value} = \left(\frac{9+1}{2}\right)\text{th value} = \frac{10}{2}\text{th value} = 5\text{th value} = 23$

$$\text{Lower quartile} = \frac{1}{4}(n+1)\text{th value} = \frac{1}{4}(9+1)\text{th value} = \frac{1}{4}(10)\text{th value}$$

$$= 2.5\text{th value} = \frac{20+20}{2} \qquad (\text{Average of the 2nd and 3rd values}) = \frac{40}{2} = 20$$

$$\text{Upper quartile} = \frac{3}{4}(n+1)\text{th value} = \frac{3}{4}(9+1)\text{th value} = \frac{3}{4}(10)\text{th value} = \frac{30}{4}\text{th value}$$

$$= 7.5\text{th value} = \frac{24+27}{2} \quad (\text{Average of the 7th and 8th values}) = \frac{51}{2} = 25.5$$

Illustration 42: Given the series: 3, 5, 2, 7, 6, 4, 9, 1, Calculate the mode, median, mean, average deviation, variance, standard deviation, the quartiles 1 and 3, the deciles 2 and 7, and the percentiles 32 and 85. **(JEE MAIN)**

Sol: 3, 5, 2, 7, 6, 4, 9, 1

Mode. Does not exist because all the scores have the same frequency.

$\text{Median} = \frac{4+5}{2} = 4.5$

$\text{Mean } \bar{X} = \frac{2+3+4+5+6+7+9+1}{8} = 4.625$

Variance $\sigma^2 = \dfrac{3^2+5^2+2^2+7^2+6^2+4^2+9^2+1^2}{8} - 4.625^2 = 6.234$

Standard Deviation $\sigma = \sqrt{6.234} = 2.497$

Average Deviation

$$D_{\overline{X}} = \frac{|3-4.625|+|5-4.625|+|2-4.625|+|7-4.625|+|6-4.625|+|4-4.625|+|9-4.625|+|1-4.625|}{8} = 2.123$$

Range r = 9 – 1 = 8

Quartiles

1, $\underline{2, 3}$, $\underline{4, 5}$, $\underline{6, 7}$, 9

2.5 4.5 6.5

↓ ↓ ↓

Q_1 Me Q_3

Deciles $8.(2/10) = 1.6 D_2 = 2$; $8.(7/10) = 5.6 D_7 = 6$

Percentiles $8.(32/100) = 2.56 P_{32} = 3$; $8.(85/100) = 6.8 P_{85} = 7$

PROBLEM-SOLVING TACTICS

- Formulates and solves a variety of meaningful problems.
- Extracts pertinent information from situations and figures out what additional information is needed.
- Formulates conjectures and argues why they must be or seem true.
- Makes sensible, reasonable estimates.
- Makes justified, logical statements.
- Employs forms of mathematical reasoning and proof appropriate to the solution of the problem at hand, including deductive and inductive reasoning, making and testing conjectures, and using counterexamples and indirect proof.
- Differentiates clearly between giving examples that support a conjecture and giving a proof of the conjecture.

FORMULAE SHEET

Arithmetic mean: (a) For ungrouped data (individual series) $\bar{x} = \dfrac{X_1 + X_2 + \ldots\ldots + X_n}{n(\text{no.of terms})} = \dfrac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$

(b) For grouped data (continuous series)

(i) Direct method $\overline{X} = \dfrac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$, where $x_i, i = 1.....n$, (n = observations, fi = corresponding frequencies)

(ii) Shortcut method : $\overline{X} = A + \dfrac{\sum f_i d_i}{\sum f_i}$, Where A=assumed mean, $d_i = x_i - A$ = deviation for each term

Median: (a) Individual series (ungrouped data): If data is raw, arrange in ascending or descending order. n denotes the number of observations. If n is odd, median = value of $\left(\frac{n+1}{2}\right)^{th}$ observations

If n is even, median= $\frac{1}{2}$[value of $\left(\frac{n}{2}\right)^{th}$ + value of $\left(\frac{n}{2}+1\right)^{th}$] observation

(b) Discrete series: First find cumulative frequencies of the variables arranged in ascending of descending order and Median = $\left(\frac{n+1}{2}\right)^{th}$ observation, where n is the cumulative frequency.

(c) Continuous distribution (grouped data)

(i) For a series in the ascending order, median = $\ell + \frac{((N/2)-C)}{f} \times i$

where ℓ = Lower limit of the median class.

F = Frequency of the median class.

N = Sum of all frequencies.

i = The width of the median class.

C = Cumulative frequency of the class preceding to median class.

(ii) For a series in descending order, median = $u - \frac{((N/2)-C)}{f} \times i$

Where u = upper limit of median class.

Mode: (a) For individual series: In the case of individual series, the value which is repeated maximum number of times is the mode of the series.

(b) For discrete frequency distribution series: In the case of discrete frequency distribution, mode is the value of the variate corresponding to the maximum frequency.

(c) For continuous frequency distribution: First find the modal class i.e., the class which has maximum frequency.

For continuous series, mode= $\ell_1 + \left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right] \times i$

Where, ℓ_1 = Lower limit of the modal class; f_1 = Frequency of the modal class

f_0 = Frequency of the class preceding modal class; f_2 = Frequency of the class succeeding modal class ; i = Size of the modal class

Relation between mean, mode and median:

(i) In symmetrical distribution: Mean=mode=median

(ii) In moderately symmetrical distribution: Mode=3median-2mean

Measure of Dispersion: The degree to which data points diverge from the average or mean value is called variation or dispersion. Popular methods of measure of dispersion:

(a) Mean Deviation: The arithmetic average of deviations from the mean, median, or mode is known as mean deviation.

(i) Individual series (ungrouped data). Mean deviation = $\frac{\sum|x-S|}{n}$

Where n = number of terms, S = deviation of variate from mean, mode, and median.

(ii) Continuous series (grouped data). Mean deviation = $\frac{\sum f|x-S|}{\sum f} = \frac{\sum f|x-s|}{N}$

Note: Mean deviation is the least when measured from the median.

(b) Standard Deviation: S.D. (σ) is the square root of the arithmetic mean of the squares of the deviations of the terms from their arithmetic mean (A.M.).

(i) For individual series (ungrouped data)

$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}}$ where $\bar{x}$ = arithmetic mean of the series. N=Total frequency

(ii) For continuous series (grouped data)

- Direct method $\sigma = \sqrt{\frac{\sum f_i(x_i-\bar{x})^2}{N}}$

  where $\bar{x}$ =Arithmetic mean of series

  x_i =Mid-value of the class

  f_i =Frequency of the corresponding x_i. N= $\sum f$ = Total frequency

- Shortcut method $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$ or $\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2}$

  where d = x - A = Derivation from assumed mean A

  f = Frequency of item (term), N= $\sum f$ =Total frequency.

  Variance – Square of standard direction i.e., variance = $(S.D)^2 = (\sigma)^2$

  Coefficient of variance = coefficient of S.D. X 100 $= \frac{\sigma}{x} \times 100$