STUDY MATERIAL FOR B.COM.,

BUSINESS MATHEMATICS AND STATISTICS

SEMESTER – III

ACADEMIC YEAR 2024-25

PREPARED BY

COMMERCE DEPARTMENT

INDEX

Business Mathematics & Statistics

Syllabus

Unit – I – Ratio

Ratio, Proportion and Variations, Indices and Logarithms.

Unit – II - Interest and Annuity

Banker's Discount - Simple and Compound Interest -Arithmetic, Geometric and Harmonic Progressions-Annuity - Meaning - Types of Annuity Applications.

Unit – III - Business Statistics Measures of Central Tendency

Arithmetic Mean, Geometric Mean -Harmonic Mean - Mode and Median - Quartiles - Deciles - Percentiles. Measures of Variation - Range - Quartile Deviation and Mean Deviation - Variance and Standard Deviation & Co-efficient.

Unit – IV - Correlation and Regression

Correlation - Karl Pearson's Coefficient of Correlation -Spearman's Rank Correlation -Regression Lines and Coefficients.

Unit – V - Time Series Analysis and Index Numbers

Time Series Analysis: Secular Trend - Seasonal Variation -Cyclical variations - Index Numbers - Aggregative and Relative - Index - Chain and Fixed Index -Wholesale Index - Cost of Living Index.

Unit – I

Ratio

Meaning

A ratio is a way to compare two quantities or values by showing the relative size of one to the other. It's expressed as a fraction, with a colon, or with the word "to." For example, if you have a ratio of 3:4, it means that for every 3 units of the first quantity, there are 4 units of the second quantity. Ratios are often used to compare things like ingredients in a recipe, aspects of a financial statement, or dimensions in design.

1. Simple Ratio:

Example: If you have 4 apples and 3 oranges, the ratio of apples to oranges is 4:3. This means for every 4 apples, there are 3 oranges.

2. Ratio in Recipes:

Example: A pancake recipe calls for 2 cups of flour to 1 cup of milk. The ratio of flour to milk is 2:1. This indicates that for every 2 cups of flour, you use 1 cup of milk.

3. Aspect Ratio:

Example: A television screen has an aspect ratio of 16:9. This means the width of the screen is 16 units for every 9 units of height.

4. Financial Ratios:

Example: A company's current ratio (a measure of liquidity) is 1.5:1. This indicates that the company has Rs.1.50 in current assets for every Rs.1.00 in current liabilities.

5. Scaling and Proportions:

Example: If a model car is built at a scale of 1:20, this means that 1 unit of measurement on the model corresponds to 20 units of the same measurement on the real car.

6. Speed Ratios:

Example: A car travels 60 miles in 1 hour and 30 minutes. The ratio of distance to time is 60 miles to 1.5 hours, which simplifies to 40 miles per hour.

7. Ratios in Surveys:

Example: In a survey of 200 people, 120 said they preferred chocolate ice cream while 80 preferred vanilla. The ratio of people who prefer chocolate to those who prefer vanilla is 120:80, which simplifies to 3:2.

8. Chemical Ratios:

Example: A chemical solution is made by mixing two substances in a ratio of 5:2. This means for every 5 parts of one substance, you mix in 2 parts of the other substance.

Ratios can be used to compare quantities, analyze proportions, and understand relationships in various contexts.

Compound Ratio

A compound ratio is a ratio that combines two or more ratios into a single, simplified ratio. It is used when you need to compare multiple quantities or combine different proportions.

Definition:

A compound ratio is obtained by multiplying the terms of two or more ratios. For example, if you have two ratios, a:b and c:d, the compound ratio is found by multiplying a with c and b with d, resulting in the ratio (a.b): (b.d)

Examples:

1. **Basic Example:**
   - Ratios:2:3 and 4:5
   - Compound Ratio Calculation:
     - Multiply the first terms: $2 \times 4 = 8$
     - Multiply the second terms: $3 \times 5 = 15$
     - Compound Ratio:8:15
2. **Application in Recipes:**
   - **Ratios for Ingredients:** A recipe calls for a ratio of flour to sugar as 3:2and the ratio of sugar to butter as 4:1. To find the compound ratio of flour to butter:
   - Combine 3:23:23:2 (flour to sugar) with 4:14:14:1 (sugar to butter).
   - **Calculation:**
     - Flour to sugar ratio: 3:2
     - Sugar to butter ratio: 4:1
     - Multiply the terms: $3 \times 4 = 12$ (flour to butter) and $2 \times 1 = 2$ (butter)
     - Compound Ratio:12:2 (which simplifies to 6:1)

3. **Ratios in Finance:**

- Ratios: Suppose a company has a profit-to-revenue ratio of 5:8 and a revenue-to-expense ratio of 3:2. To find the compound ratio of profit to expense:
- Combine 5:85:85:8 (profit to revenue) with 3:23:23:2 (revenue to expense).
- Calculation:
  - Profit to revenue ratio: 5:8
  - Revenue to expense ratio: 3:2
  - Multiply the terms: 5×3=15 (profit to expense) and 8×2=16(expense)
  - Compound Ratio:15:16

4. **Ratios in Mixes:**

- **Ratios:** A solution requires mixing two solutions in the ratio of 7:3 and then mixing this with another solution in the ratio of 2:5. To find the overall ratio of the first solution to the third solution:
- Combine 7:3 (first solution to second solution) with 2:5 (second solution to third solution).
- **Calculation:**
  - First to second solution ratio: 7:3
  - Second to third solution ratio: 2:5
  - Multiply the terms: 7×2=14 (first to third solution) and 3×5=15(third solution)
  - Combound Ratio:14:15

By combining ratios in this way, you can get a more comprehensive understanding of the relationships between multiple quantities.

Types of proportion with examples

Proportions are statements that two ratios or fractions are equal. There are several types of proportions, each useful in different contexts. Here's a rundown of the main types with examples:

Types of Proportions

Based on the type of relationship two or more quantities share, the proportion can be classified into different types. There are two types of proportions.

- Direct Proportion
- Inverse Proportion

Direct Proportion

Direct proportion (or direct proportionality) is a relationship between two quantities in which an increase or decrease in one quantity causes a proportional increase or decrease in the other quantity. In other words, the two quantities change in the same direction, and their ratio remains constant.

Inverse Proportion

Inverse proportion (or inverse proportionality) describes a relationship between two quantities where an increase in one quantity results in a corresponding decrease in the other, and vice versa. In other words, the two quantities are inversely related if their product remains constant.

Example No. 1

$$\frac{x}{3} = \frac{8}{12}$$

12x = 24

x = 2

Example No. 2

Solve the proportion $\frac{n}{18} = \frac{20}{45}$

45n = 18x20

45n = 360

n = 360/45

n = 8

Example No. 3

In a computer lab, there are 3 computers for every 6 students. How many computers will be needed for 24 students?

For 6 students = 3 computers

Therefore for 24 students = $\frac{24 \times 3}{6}$ = 12 computers

Example No. 4

A pet store has 8 cats, 12 dogs, and 3 rabbits. What does the ratio 8 : 23 compares?

Ans : cats and all animals

Example No. 5

If 15 workers can build a wall in 48 hours, how many workers will be required to do the same work in 30 hours?

In 48 hrs, the wall is built by 15 workers In 48 hrs,

Thus, in 30 hrs, the wall will be built by

15×48/30 workers

=24 workers

Example No. 6

There are 15 girls and 12 boys in a class. What is the ratio of girls to boys?

Ans : 5:4

Law of Indices

The law of indices (or laws of exponents) are rules that describe how to handle operations involving powers (exponents) of numbers. These laws are crucial for simplifying and calculating expressions with exponents. Here's an overview of the main laws of indices:

Rule 1: If a constant or variable has index as '0', then the result will be equal to one, regardless of any base value.

a0 = 1

Example: 50 = 1, 120 = 1, y0= 1

Rule 2: If the index is a negative value, then it can be shown as the reciprocal of the positive index raised to the same variable.

a-p = 1/ap

Example: 5-1 = ⅕, 8-3=1/83

Rule 3: To multiply two variables with the same base, we need to add its powers and raise them to that base.

ap.aq = ap+q

Example: 52.53 = 52+3 = 55

Rule 4: To divide two variables with the same base, we need to subtract the power of denominator from the power of numerator and raise it to that base.

ap/aq = ap-q

Example: 104/102 = 104-2 = 102

Rule 5: When a variable with some index is again raised with different index, then both the indices are multiplied together raised to the power of the same base.

(ap)q = apq

Example: (82)3 = 82.3 = 86

Rule 6: When two variables with different bases, but same indices are multiplied together, we have to multiply its base and raise the same index to multiplied variables.

ap.bp = (ab)p

Example: 32.52 = (3 x 5)2 = 152

Rule 7: When two variables with different bases, but same indices are divided, we are required to divide the bases and raise the same index to it.

ap/bp = (a/b)p

Example: 32/52 = (⅗)2

Rule 8: An index in the form of a fraction can be represented as the radical form.

ap/q = q√ap

Example: 61/2 = √6

Q.1: Multiply x4y3z2 and xy5z-1

Solution: x4y3z2 and xy5z-1

= x4.x .y3.y5.z2.z-1

= x4+1.y3+5.z2-1

= x5.y8.z

Q.2: Solve a3b2/a2b4

Solution: a3b2/a2b4

= a3-2b2-4

= a1b-2

= a b-2
= a/b2

Q.3: Find the value of 272/3.

Solution: 272/3

= 3√272

= 32

= 9

Logarithms are the other way of writing the exponents. A logarithm of a number with a base is equal to another number. A logarithm is just the opposite function of exponentiation. For example, if 102 = 100 then log10 100 = 2.

Hence, we can conclude that,

Logb x = n or bn = x

Where b is the base of the logarithmic function.

This can be read as “Logarithm of x to the base b is equal to n”.

In this article, we are going to learn the definition of logarithms, two types of logarithms such as common logarithm and natural logarithm

Logarithm Types

In most cases, we always deal with two different types of logarithms, namely

- Common Logarithm
- Natural Logarithm

Common Logarithm

The common logarithm is also called the base 10 logarithms. It is represented as log10 or simply log. For example, the common logarithm of 1000 is written as a log (1000). The common logarithm defines how many times we have to multiply the number 10, to get the required output.

For example, log (100) = 2

If we multiply the number 10 twice, we get the result 100.

Natural Logarithm

The natural logarithm is called the base e logarithm. The natural logarithm is represented as ln or loge. Here, “e” represents the Euler’s constant which is approximately equal to 2.71828. For example, the natural logarithm of 78 is written as ln 78. The natural logarithm defines how many we have to multiply “e” to get the required output.

For example, ln (78) = 4.357.

Thus, the base e logarithm of 78 is equal to 4.357

Logarithm Rules and Properties

There are certain rules based on which logarithmic operations can be performed. The names of these rules are:

- Product rule
- Division rule
- Power rule/Exponential Rule
- Change of base rule
- Base switch rule
- Derivative of log
- Integral of log

Product Rule

In this rule, the multiplication of two logarithmic values is equal to the addition of their individual logarithms.

Logb (mn)= logb m + logb n

For example: log3 (2y) = log3 (2) + log3 (y)

Division Rule

The division of two logarithmic values is equal to the difference of each logarithm.

Logb (m/n)= logb m – logb n

For example, log3 (2/ y) = log3 (2) -log3 (y)

Exponential Rule

In the exponential rule, the logarithm of m with a rational exponent is equal to the exponent times its logarithm.

Logb (mn) = n logb m

Example: logb(23) = 3 logb 2

Change of Base Rule

Logb m = loga m/ loga b

Example: logb 2 = loga 2/loga b

Base Switch Rule

logb (a) = 1 / loga (b)

Example: logb 8 = 1/log8 b

Derivative of log

If f (x) = logb (x), then the derivative of f(x) is given by;

f'(x) = 1/(x ln(b))

Example: Given, f (x) = log10 (x)

Then, f'(x) = 1/(x ln(10))

Integral of Log

∫logb(x)dx = x(logb(x) – 1/ln(b)) + C

Example: ∫ log10(x) dx = x · (log10(x) – 1 / ln(10)) + C

Other Properties

Some other properties of logarithmic functions are:

- Logb b = 1
- Logb 1 = 0
- Logb 0 = undefined

Logarithmic Formulas

logb(mn) = logb(m) + logb(n)

logb(m/n) = logb (m) – logb (n)

Logb (xy) = y logb(x)

Logbm√n = logb n/m

m logb(x) + n logb(y) = logb(xmyn)

logb(m+n) = logb m + logb(1+nm)

logb(m – n) = logb m + logb (1-n/m)

Example 1:

Solve log 2 (64) =?

Solution:

since 26= 2 × 2 × 2 × 2 × 2 × 2 = 64, 6 is the exponent value and log 2 (64)= 6.

Example 2:

What is the value of log10(100)?

Solution:

In this case, 102 yields you 100. So, 2 is the exponent value, and the value of log10(100)= 2

Example 3:

Use of the property of logarithms, solve for the value of x for log3 x= log3 4+ log3 7

Solution:

By the addition rule, log3 4+ log3 7= log 3 (4 * 7)

Log 3 (28). Thus, x= 28.

Example 4:

Solve for x in log2 x = 5

Solution:

This logarithmic function can be written In the exponential form as 25 = x

Therefore, 25= 2 × 2 × 2 × 2 × 2 = 32, x= 32.

Example 5:

Find the value of log5 (1/25).

Solution:

Given: log5 (1/25)

By using the property,

Logb (m/n)= logb m – logb n

log5 (1/25) = log5 1 – log5 25

log5 (1/25) = 0 – log5 52

log5 (1/25) = -2log55

log5 (1/25) = -2 (1) [By using the property loga a = 1)

log5 (1/25) = -2.

Hence, the value of log5 (1/25) = -2

UNIT –II

Banker's Discount-Banker's Gain

In finance, "banker's discount" and "banker's gain" are terms related to the discounting of bills of exchange and promissory notes. They both reflect different aspects of the discounting process, which is essentially the sale of a financial instrument at a price lower than its face value.

Banker's Discount

Banker's Discount refers to the amount of interest deducted by a bank when it discounts a bill of exchange or a promissory note before its maturity date. This discount is calculated based on the face value of the instrument and the time remaining until its maturity. The formula used is:

Banker's Discount=Face Value × Discount Rate × Time ⁄ 360

Where:

- **Face Value** is the amount payable at maturity.
- **Discount Rate** is the annual interest rate.
- **Time** is the number of days until maturity.

The banker's discount is used to determine how much less than the face value the bank will pay for the instrument today, reflecting the interest that would have accrued if the instrument were held until maturity.

Banker's Gain

Banker's Gain is the actual profit a bank makes from the discounting process. It is the difference between the face value of the instrument and the amount the bank actually receives when it discounts the instrument. However, since the discount is computed on the face value rather than the present value, the gain can be calculated as:

Banker's Gain=Banker's Discount−(Face Value−Present Value)

Where:

- **Present Value** is the amount actually paid by the bank when it buys the bill.
- **Face Value** is the amount payable at maturity.

Key Difference: The banker's discount is computed using the face value of the instrument, while the banker's gain considers the actual present value received by the bank and the difference between this amount and the face value.

In essence, while the banker's discount represents the theoretical reduction in value, the banker's gain is the actual profit realized from the discounting process.

The formula can be represented as:

PW (1+RT/100) = F

BD= FRT/100

TD = PW X RT/100

BG = BD -TD

where PW = Present worth or present value

F= Face value

R =Rate of interest

T =Time

TD = True Discount

BG = Banker's gain

BD = Bankers Discount

Bankers gain will be determined after calculating true discount and present value.

So, present value, PW = 105400/(1 + 1/2 x (9/100))

= Rs. 104927

Also, true discount, TD = TR x PW/100

= 104927 x 1/2 x 9/100

= Rs. 4721.

So, Bankers gain = BD – TD = Rs. 4743 – Rs. 4721 = Rs. 22

Q. Find the present value, banker's discount, and banker's gain for a bill of Rs. 105400 which is due in 6 months at 9% interest rate per annum.

Here, F = Rs. 105400, T = 6/12 years = 1/2 years

R = 9%

So, Bankers discount, BD = FTR/100

= (105400 x 1/2 x 9)/100

=> 1054 x 1/2 x 9

So, the answer is Rs. 4743.

Simple Interest – Definition and Calculation

Simple interest is a method of calculating the interest charged or earned on a principal amount of money over a period of time. Unlike compound interest, which calculates interest on both the initial principal and the accumulated interest, simple interest is calculated only on the principal amount.

Formula for calculating simple interest:

$I=P\times r\times t$

Where,
I = interest

P = principal

r = interest rate (per year)

t = time (in years or fraction of a year)

Example 1: Basic Calculation

Problem: You invest $1,000 at an annual interest rate of 5% for 3 years. What is the simple interest earned?

Solution:

1. Principal (P) = $1,000
2. Annual interest rate (r) = 5% = 0.05 (as a decimal)
3. Time (t) = 3 years

Using the formula:

$SI=1000\times 0.05\times 3$

$SI=50x3$

$SI= 150$

So, the simple interest earned is Rs.150.

Example 2: Maturity Value

When dealing with simple interest and you want to determine the maturity value (the total amount you will have at the end of the investment or loan period), you can use the following approach:

Formula for Maturity Value:

Maturity Value=P+SI

where:

- Simple Interest=P×r×t
- P is the principal amount,
- r is the annual interest rate (as a decimal),
- t is the time in years.

You can also combine these into a single formula for the maturity value:

Maturity Value=P×(1+r×t)

Example:

Let's say you invest Rs.2,000 at an annual interest rate of 4% for 5 years.

1. Convert the interest rate to a decimal: 4% becomes 0.04.
2. Calculate the simple interest using the formula:

   Simple Interest=P×r×t

   Simple Interest=2000×0.04×5

   Simple Interest=2000×0.20

   Simple Interest=400
3. Use the maturity value formula:

   Maturity Value=P×(1+r×t)

   Maturity Value=2000×(1+0.04×5)

   Maturity Value=2000×(1+0.20)

   Maturity Value=2400

So, the maturity value of the investment after 5 years would be Rs.2,400. This amount includes the original principal plus the interest earned.

To find the time (t) when you know the maturity value, the principal, and the annual interest rate, you can rearrange the formula for the maturity value in terms of t.

Given:

- Maturity Value (MV)
- Principal (P)
- Annual Interest Rate (r, expressed as a decimal)

The formula for the maturity value is:

$MV=P\times(1+r\times t)$

To find t, rearrange the formula:

1. **Start by isolating 1+r×t:**

   $1+r\times t=MV/P$

2. **Subtract 1 from both sides:**

   $r\times t=MV/P-1$

3. **Solve for t:**

   $t=MV/P-1/r$

Example:

Suppose you have a principal of Rs.1,500, a maturity value of Rs.1,950, and an annual interest rate of 6% (0.06 as a decimal). You want to find out how long it will take to reach the maturity value.

1. Use the formula:

$$t=\frac{\frac{MV}{P}-1}{r}$$

$$t=\frac{\frac{1950}{1500}-1}{0.06}$$

2. Calculate the fraction inside the numerator:

$$\frac{1950}{1500}=1.3-1=0.3$$

3. Now divide by the interest rate:

$$t=\frac{0.3}{0.06}$$

$t=5$

So, it will take 5 years for the investment to grow from Rs.1,500 to Rs.1,950 at an annual interest rate of 6%.

Compound Interest:

Compound interest refers to the interest calculated on the initial principal and also on the accumulated interest from previous periods. This means that interest is earned not only on the original amount of money but also on the interest that has already been added to the principal. As a result, compound interest grows at a faster rate compared to simple interest, where interest is only calculated on the original principal.

Compound Interest Formula

As we have already discussed, the compound interest is the interest-based on the initial principal amount and the interest collected over the period of time. The compound interest formula is given below:

Compound Interest = Amount – Principal

Where,

- A = amount
- P = principal
- r = rate of interest
- n = number of times interest is compounded per year
- t = time (in years)

Alternatively, we can write the formula as given below:

CI = A – P

Examples 1

A town had 10,000 residents in 2000. Its population declines at a rate of 10% per annum. What will be its total population in 2005?

Solution:

The population of the town decreases by 10% every year. Thus, it has a new population every year. So the population for the next year is calculated on the current year population. For the decrease, we have the formula A = P(1 – R/100)n

Therefore, the population at the end of 5 years = 10000(1 – 10/100)5

= 10000(1 – 0.1)5 = 10000 x 0.95 = 5904 (Approx.)

Examples 2

The count of a certain breed of bacteria was found to increase at the rate of 2% per hour. Find the bacteria at the end of 2 hours if the count was initially 600000.

Solution:

Since the population of bacteria increases at the rate of 2% per hour, we use the formula

A = P(1 + R/100)n

Thus, the population at the end of 2 hours = 600000(1 + 2/100)2

= 600000(1 + 0.02)2 = 600000(1.02)2 = 624240

Examples 3

The price of a radio is Rs. 1400 and it depreciates by 8% per month. Find its value after 3 months.

Solution:

For the depreciation, we have the formula A = P(1 – R/100)n.

Thus, the price of the radio after 3 months = 1400(1 – 8/100)3

= 1400(1 – 0.08)3 = 1400(0.92)3 = Rs. 1090 (Approx.)

Compound Interest and Simple Interest

Now, let us understand the difference between the amount earned through compound interest and simple interest on a certain amount of money, say Rs. 100 in 3 years . and the rate of interest is 10% p.a.

Below table shows the process of calculating interest and total amount.

		Under Simple Interest	Under Compound Interest
First Year	Principal Interest 10% Year end Amount	Rs.100.00 Rs. 10.00 Rs.110.00	Rs.100.00 Rs. 10.00 Rs.110.00

		Under Simple Interest	Under Compound Interest
Second Year	Principal Interest 10% Year end Amount	Rs.100.00 Rs. 10.00 Rs.(110+10)=120.00	Rs.100.00 Rs. 11.00 Rs.121.00

		Under Simple Interest	Under Compound Interest
Third Year	Principal Interest 10% Year end Amount	Rs.100.00 Rs. 10.00 Rs.(120+10)=130	Rs.121.00 Rs.12.10 Rs.133.10

Progression

A progression (which is also known as a sequence) is nothing but a pattern of numbers. For example, 3, 6, 9, 12, ... is a progression because there is a pattern observed where every number here is obtained by adding 3 to its previous number. But this pattern doesn't need to be the same in every progression.

The pattern of a progression depends on its type. Let us see the types of progressions along with examples and their formulas.

Definition of a Progression

Progression is a list of numbers (or items) that exhibit a particular pattern. A progression is also known as a sequence. In a progression, every term is obtained by applying a specific rule on its previous term. In other words, every term of a progression is defined by a general term (or) nth term, which is denoted by an.

For example, the nth term of a progression 3, 5, 7, 9, ... is an = 2n + 1. Substituting n = 1, 2, 3, ... here, we get the 1st, 2nd, 3rd, ... terms. For example:

- When n = 1, first term = 2(1) + 1 = 3.
- When n = 2, second term = 2(2) + 1 = 5.
- When n = 3, third term = 2(3) + 1 = 7

and so on.

Types of Progressions

There are mainly 3 types of progressions in math. They are:

- Arithmetic Progression (AP)
- Geometric Progression (GP)
- Harmonic Progression (HP)

Arithmetic Progression (AP)

The differences between any two consecutive numbers are all same. example 1, 4, 7, 10, ...

Geometric Progression (GP)

The ratios of any two consecutive numbers are all same. Example 4, 16, 64, 256, ...

Harmonic Progression (HP)

The reciprocals of terms form an AP. example 1/2, 1/4, 1/6, ...

We will learn about each progression in detail in the upcoming sections.

Arithmetic Progression

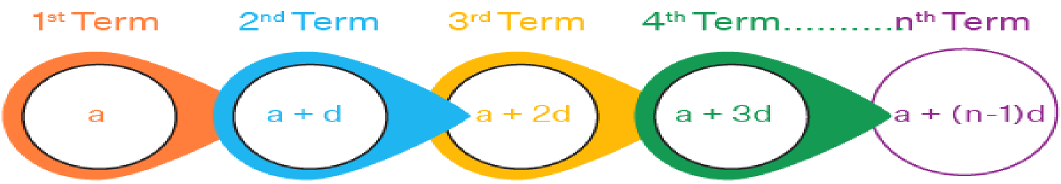
An arithmetic progression (AP) is a sequence of numbers in which each successive term is the sum of its preceding term and a fixed number. This fixed number is called the common difference. For example, 1, 4, 7, 10, ... is an AP as every number is obtained by adding a fixed number 3 to its previous term.

- 2nd term = 4 = 1 + 3 = 1st term + 3
- 3rd term = 7 = 4 + 3 = 2nd term + 3
- 4th term = 10 = 7 + 3 = 3rd term + 3

and so on.

In general, an arithmetic progression looks like this:

Arithmetic Progression



Here,

- 'a' is the first term and
- 'd' is the common difference (fixed number)

Arithmetic Progression Example

For example, Minnie put \$30 in her piggy bank when she was 7 years old. She increased the amount she put in her piggy bank on each successive birthday by \$3. So, the amount in her piggy bank follows the pattern of \$30, \$33, \$36, and so on. The succeeding terms are obtained by adding a fixed number, that is, \$3. This fixed number is called the common difference (It can be positive, negative, or zero). Hence the progression 30, 33, 36, ... is an AP.

Arithmetic Progression Formulas

Let the first term of the progression be a, the common difference be d, and the nth term be an. Then, the arithmetic progression formulas are given by:

- $a_n = a + (n - 1) d$
- $d = a_n - a_{n-1}$

- Sum of the first n terms, Sn = n/2(2a+(n-1)d) (or) Sn = n/2(a + l), where l = the last term = Tn.

An arithmetic progression is a sequence of numbers such that the difference d between each consecutive term is a constant.

$$a, a+d, a+2d, a+3d, \ldots\ldots\ldots$$

The n^{th} term, $a_n = a + (n-1)d$

Sum of first n terms, $S_n = \frac{n}{2}[2a + (n-1)d]$

$$= \frac{n}{2}[a + l]$$

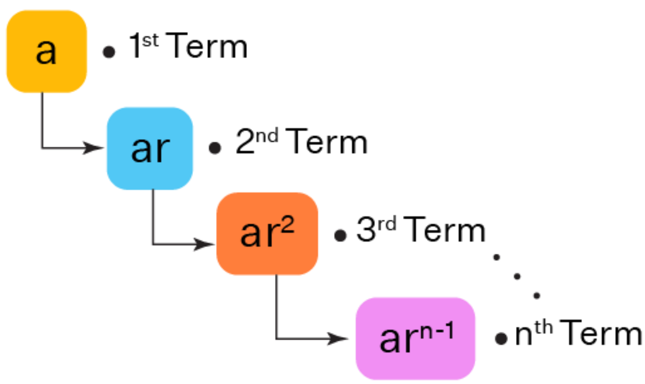
Geometric Progression

A geometric progression (GP) is a sequence of numbers in which each successive term is the product of its preceding term and a fixed number. This fixed number is called the common ratio. For example, 4, 16, 64, 256, ... is a GP as every number is obtained by multiplying a fixed number 4 to its previous term.

- 2nd term = 16 = 4(4) = 4(1st term)
- 3rd term = 64 = 4(16) = 4(2nd term)
- 4th term = 256 = 4(64) = 4(3rd term)

  and so on.

In general, a geometric progression looks like this:



Here,

- 'a' is the first term and
- 'r' is the common ratio (fixed number)

Geometric Progression Example

Consider an example of a geometric progression: 1, 4, 16, 64, ... Observe that 4/1 = 16/4 = 64/16 = ... = 4. All the ratios are same. Hence it is a GP.

Geometric Progression Formulas

Let the first term of the progression be a, the common ratio be r, and the nth term be an. Then, the geometric progression formulas are given by:

- a_n = arn - 1
- Sum of the first 'n' terms, S_n = a(rn - 1) / (r - 1) when r ≠ 1 and S_n = na when r = 1.
- Sum of infinite geometric series, S∞ = a / (1 - r) when |r| < 1 and S∞ diverges when |r| ≥ 1.

For a geometric progression a , ar , ar^2 ,........

- n^{th} term ,

$$a_n = ar^{n-1}$$

- Sum of n terms,

$$S_n = \begin{cases} \dfrac{a(r^n - 1)}{r - 1} & , \text{ when } r \neq 1 \\ na & , \text{ when } r = 1 \end{cases}$$

- Sum of infinite terms,

$$S_\infty = \begin{cases} \dfrac{a}{1 - r} & , \text{ when } |r| < 1 \\ \text{diverges} & , \text{ when } |r| \geq 1 \end{cases}$$

Harmonic Progression

A harmonic progression is a sequence obtained by taking the reciprocal of the terms of an arithmetic progression. The sequence of natural numbers is an arithmetic progression. So, taking reciprocals of each term we get 1,1/2,1/3,1/4,... This is harmonic progression.

Harmonic Progression Example

When a ball is dropped, the initial height reached by the ball is 1/2 units and after the first impact, the height attained by the ball is 1/4 units. After the second impact, the height attained by the ball is 1/6 units. After the third impact, the height attained by the ball is 1/8 units, and so on. Now the sequence of heights formed by the ball is: 1/2, 1/4, 1/6, 1/8, ...

This sequence is a harmonic progression because the reciprocal of all the terms of this progression form an arithmetic progression.

- Reciprocal of the sequence: 2, 4, 6, 8, ... → AP
- Hence, the sequence: 1/2, 1/4, 1/6, 1/8, ... → HP

Harmonic Progression Formulas

For a harmonic progression 1/a, 1/(a+d), 1/(a+2d), ...

- nth term, an = 1 / (a + (n - 1) d)
- Sum of the first n terms, Sn = 1/d ln [(2a + (2n - 1) d] / (2a - d)]

Sequence formula of n^{th} term

$$a_n = \frac{1}{a + (n-1)d}$$

Series formula for the sum of n terms

$$S_n = \frac{1}{d} \ln\left[\frac{2a + (2n-1)d}{2a - d}\right]$$

Problem-1. Find the common difference of the following AP and write the AP up to 8 terms.

7, 12, 17, 22, 27...

Solution- The common difference for an AP is the difference between any of its two consecutive terms. Thus for this AP, the common difference will be 12-7=5.

Since we are given the five terms of the AP, we can find the next terms by adding the common difference.

6th term= 27+5=32

7th term= 32+5=37

8th term= 37+5=42

Thus the AP is 7, 12, 17, 22, 27, 32, 37, 42

Problem-2 If a, b, c are in AP, find the value of a-2b+c.

Solution- Since a, b, c are in AP. Let's suppose the common difference of this AP is d.

Then the b=a+d, c=a+2d.

Now putting the value of b and c, in the expression a-2b+c, it will become a-2(a+d)+a+2d

=a-2a-2d+a+2d=0

Thus, the value of the expression a-2b+c will be zero if a, b, c are in AP.

Problem-3 The sum of the first six terms of an AP is 96, and the common difference is 4. Find the 4th term of the AP.

Solution- Let, the given AP has the first term as A, and it is given that total no. of terms=6, and the common difference is 4. Now, putting these values in the formula for the sum of an AP-

1⁄2(N)(2*A +(N-1)*D)= 6/2(2*A+5*4).

Since it is given that the sum of the AP is 96

=> 6/2(2*A+5*4)=96

=> A=6.

Thus, the required AP is 6, 10, 14, 18, 22. 26. The 4th term of the AP is 18.

Problem-4 Find the number of terms in the following GP and also find its sum.

2, 6, 18, …., 1458.

Solution- To find the number of terms in the GP, we first need to find the common ratio of the progression. Since common ratio(R) = Tn+1/Tn=6/2=3.

Now, we also know that Nth term of a GP is A*RN-1 where R is the common ratio.

Since the last term of the given GP is 1458, Now, equate it with the above formula to find the number of terms in the progression-

=> 1458= A*RN-1

=> 2*729= 2*3N-1

=> 2*36= 2*3N-1

On comparing LHS with RHS

=> 6=N-1

=> N=7

Problem-5 If the sum of reciprocals of the first 13 terms of an HP is 260, find the 7th term of the sequence.

Solution- Sum of reciprocals of an HP refers to the sum of AP used to form the HP,

Let' the first term of the AP be A, and common difference be D, then the sum of the AP will be 1⁄2(13)(2*A +(13-1)*D) Since the sum is given to be 260.

=> 1⁄2(13)(2*A +(13-1)*D)=260

=> 2*A +(13-1)*D=40

=> 2*A +12*D=40

=> A +6*D=20

Since the Nth term of an HP is 1/(A+(N-1)D)

Thus, the 7th term of given HP will be 1/(A+6*D)=1/20.

Annuities

An annuity is a series of payments, with one payment per period for a given number of periods.

The variables in an annuity are:

- N: the number of periods, in the term (equal to the number of payments)
- I/Y: the nominal interest rate
- PMT: the periodic payment
- PV: the present value of the annuity
- FV: the future value of the annuity

Annuities can be used for:

- amortizing a loan, where the loan value is the present value of the annuity. The variables are N, I/Y, PMT and PV.
- accumulating an amount, where the amount is the future value of the annuity. The variables are N, I/Y, PMT and FV.

Annuities can be solved in many ways:

- by formula
- by calculating interest period by period

- by focal date methods
- by calculator program, which is what we will concentrate on.

Previous/next navigation

Formula to Calculate Present Value Annuities

The formula for the present value of an ordinary annuity:

PV ordinary annuity = P * 1 - (1 + r)-n/ r

Where,

PV = present value of an ordinary annuity

P = value of each payment

R = interest rate/ period

N = total number of periods

The formula for calculating the present value of an annuity due is:

PV Annuity Due = C × [i1 − (1 + i)−n] × (1 + i)

Formula to Calculate Future Value Annuities

Instead of calculating each payment separately and then adding them all up, you can instead apply the following formula, which will tell you the amount of money you'd have in the end:

FV Ordinary Annuity = C × [i(1 + i)n −1]

Where:

C = cash flow/period

i = rate of interest

n = total number of payments

The formula for the future value of an annuity due is:

FV Annuity Due = C × [i(1 + i)n−1] × (1 + i)

Example:1

Calculate the future value of the ordinary annuity and the present value of an annuity due where cash flow per period amounts to rs. 1000 and interest rate is charged at 0.05%.

Solution:

Using the formula to calculate future value of ordinary annuity = C × [(1 + i)n – 1/i

= Rs. 1,000 × [0.05 (1 + 0.05)5−1]

=Rs.1, 000 × 5.53

=Rs. 5,525.63

Note that the one-cent difference in these outcomes, Rs. 5,525.64 vs. Rs. 5,525.63, is because of rounding in the first calculation.

Now to calculate the present value of an annuity due:

Use the formula

PV Annuity Due = C × [i1 − (1 + i)−n] × (1 + i)

Plugging in the values:

= Rs. 1,000 × [0.05(1− (1 + 0.05)−5] × (1 + 0.05)

= Rs. 1,000 × 4.33 × 1.05

= Rs. 4,545.95

Example 2

Assume that the formula will be used to calculate the future value of a two year ordinary annuity that offers an annual interest rate of 6%, monthly payments of $1000, and monthly compounding.

Define each of the variables but do not calculate the future value.

Solution

A = amount of annuity (Future Value) = unknown

P = periodic payment amount = $1000

where r = annual interest rate = 6% = 0.06

n = number of compounding periods per year = 12

t = time (in years) = 2 years

Note that the term of the annuity is 2 years. The interest calculation involves monthly compounding so n = 12 since there are 12 compounding periods in a year.

Example 3

Assume that the formula will be used to calculate the future value of a two year ordinary annuity that offers an annual interest rate of 6%, monthly payments of $1000, and monthly compounding.

Define each of the variables but do not calculate the future value.

Solution

A = amount of annuity (Future Value) = unknown

P = periodic payment amount = \$1000

where r = annual interest rate = 6% = 0.06

n = number of compounding periods per year = 12

t = time (in years) = 2 years

Note that the term of the annuity is 2 years. The interest calculation involves monthly compounding so n = 12 since there are 12 compounding periods in a year.

Example 4

Assume that the formula will be used to calculate the future value of a two year ordinary annuity that offers an annual interest rate of 6%, monthly payments of \$1000, and monthly compounding.

Define each of the variables but do not calculate the future value.

Solution

A = amount of annuity (Future Value) = unknown

P = periodic payment amount = \$1000

where r = annual interest rate = 6% = 0.06

n = number of compounding periods per year = 12

t = time (in years) = 2 years

Note that the term of the annuity is 2 years. The interest calculation involves monthly compounding so n = 12 since there are 12 compounding periods in a year.

Unit – III

Measures of Central Tendency

Average – Meaning

The number you get when you add two or more figures together and then divide the total by the number of figures you added.

Definition

A number that is calculated by adding quantities together and then dividing the total by the number of quantities.

Characteristics of a Good Average

The characteristics of a good average typically include:

1. **Relevance:** It should be applicable to the context, providing a meaningful summary of the data set.
2. **Simplicity:** A good average is easy to understand and interpret.
3. **Robustness:** It should not be overly influenced by outliers or extreme values, which can skew the result.
4. **Representativeness:** The average should reflect the central tendency of the data, giving a fair representation of the overall set.
5. **Consistency:** It should yield similar results under repeated sampling of the same population.
6. **Comparability:** A good average allows for easy comparison across different data sets or groups.
7. **Appropriateness:** Depending on the data distribution (e.g., normal, skewed), different types of averages (mean, median, mode) may be more appropriate.

These characteristics help ensure that the average provides a reliable and insightful summary of the data.

Uses of Average

Averages are used in various fields and contexts for several purposes, including:

1. **Data Summary:** Averages provide a quick way to summarize large data sets, giving a central value that represents the overall trend.
2. **Comparison:** Averages allow for easy comparison between different groups or time periods, helping to identify trends or disparities.

3. **Decision-Making:** In business and economics, averages inform decisions about budgeting, forecasting, and resource allocation.
4. **Performance Evaluation:** Averages are commonly used to assess individual or team performance, such as in grades or sales figures.
5. **Research Analysis:** In scientific studies, averages help interpret results and draw conclusions about populations or phenomena.
6. **Quality Control:** In manufacturing and service industries, averages monitor quality and consistency in production processes.
7. **Financial Analysis:** Averages, such as moving averages, are used in finance to analyze trends in stock prices and investment performance.
8. **Public Policy:** Averages can inform policy decisions by summarizing data on demographics, health, education, and economic conditions.
9. **Surveys and Polls:** Averages help interpret survey results, providing insights into public opinion and behavior.
10. **Sports Statistics:** In sports, averages (like batting average or points per game) help evaluate player and team performance.

Overall, averages serve as essential tools for understanding data and making informed decisions across various fields.

Functions

- The most important objective of calculating and measuring average is to determine a single figure of whole series.
- The average describes the characteristics of the entire group.
- Averages are helpful for taking and overview of a statistical data which ordinarily cannot be understood.

Objectives

The objectives of average are-

1.The most important objective of calculating and measuring average is to determine a single figure of whole series.

2.The average describes the characteristics of the entire group.

3.Averages are helpful for taking and overview of a statistical data which ordinarily cannot be understood.
4.Averages facilitates comparison between one group and another.

Types of Average

- Mathematical Average
- Locational Average
- Commercial Average

Arithmetic Mean

The arithmetic mean is the simple average, or sum of a series of numbers divided by the count of that series of numbers. In the world of finance, the arithmetic mean is not usually an appropriate method for calculating an average, especially when a single outlier can skew the mean by a large amount.

Merits of Arithmetic Mean

(i) It is rigidly defined, simple, easy to understand and easy to calculate.

(ii) It is based upon all the observations.

(iii) Its value being unique, we can use it to compare different sets of data.

(iv) It is least affected by sampling fluctuations.

Demerits of Arithmetic Mean

1. Effects of extreme value mean value may not figure in the series at all.

2. Arithmetic mean can not be computed when class intervals have open ends.

Individual Series

Find our mean from the following data

Roll No	1	2	3	4	5
Marks	35	40	60	75	90

Solution

Roll No	Marks (X)
1	35
2	40
3	60
4	75
5	90
N=5	$\sum X = 300$

Formula = X =∑X /N

X =300/5 = 60

The mean marks =60.

Discrete Series

Find the average of the following discrete series

X	4	6	8	10	12
Y	3	2	5	8	10

x	f	Fx
4	3	12
6	2	12
8	5	40
10	8	80
12	10	120
	∑f=28	∑fx=264

X = ∑fx /N = 264, N=28

X = 264 / 28 = 9.42

Continuous Series

Calculate Arithmetic Mean

Age in years	20-30	30-40	40-50	50-60	60-70
No of workers	8	15	12	9	6

Solution

Age(in years) f	No. of workers F	Mid –Value m	fm	d=m-A (A=45)	fd
20-30	8	25	200	-20	-160
30-40	15	35	525	-10	-150
40-50	12	45(A)	540	0	0
50-60	9	55	495	10	90
60-70	6	65	390	20	120
	∑f=50		∑fm=2150		∑fd=-100

Arithmetic Mean

X =∑fm / N =2150/50

The Arithmetic mean = 43.

Median

A median is a number that falls in the middle of a group. This is accomplished by ordering the numbers from smallest to largest and locating the one that falls in the middle. A mean is the average of a data set. It's also called the arithmetic mean. It's the average of the sum of the numbers in a group.

Merits of Median

1.Certainty

2.Simplicity

3.Unaffected by extreme values

4.Possible even if data is incomplete

5.Graphic presentation

6.Appropriate for qualitative data.

Demerits of Median

1.Not based on all observations

2.Lack of representative character

3.Effect of sampling fluctuations

4.Lack of algebraic treatment.

Individual Series

Arrange the data either ascending or descending order

$$\text{Median} = \frac{N+1}{2} \text{ th Item}$$

Find out the median from the following

Families	A	B	C	D	E	F	G	H	I	J	K	L
Monthly Income of Families	1,500	2,000	1,200	800	1,000	2,500	2,200	1,700	3,200	4,500	1,100	3,500

Solution

Families	A	B	C	D	E	F	G	H	I	J	K	L
Monthly Income of Families	800	1,000	1,100	1,200	1,500	1,700	2,000	2,200	2,500	3,200	3,500	4,500

$$\text{Median} = \frac{N+1}{2} \text{ th Item}$$

$$= \frac{12+1}{2} \text{ the item}$$

$= 13 / 2 = 6.5$th item

To find the median, add the 6th and 7th items together and then divide the sum by 2.

$$= \frac{1700+2000}{2} \text{ the item} = 1850$$

Discrete Series

Calculate the median of the following data:

Income (in Rs.)	1000	2000	2500	3000	4500
No.of workers	6	12	9	14	8

Solution

Income (in Rs.) (x)	**No.of workers (f)**	**Cumulative Frequency (c.f)**
1000	6	6
2000	12	18
2500	9	27
3000	14	41
4500	8	49
	N=∑f=49	

$$\text{Median} = \text{Size of } \frac{N+1}{2} \text{ th Item}$$

$$= \text{size of } \frac{49+1}{2} \text{ th Item}$$

= 25 which is nearer to 27

Cf of 27 = 2500

Median=2500.

Continuous series

Class Interval (x)	Frequency (f)
0-5	5
5-10	3
10-15	4

15-20	8
20-25	7
25-30	3

Solution

Class Interval (x)	Frequency (f)	Cumulative Frequency (c.f)
0-5	5	5
5-10	3	8
10-15	4	12
15-20	8	20
20-25	7	27
25-30	3	30
	N=∑f=30	

$$\text{Median} = \text{Size of } \frac{N}{2} \text{ th Item}$$

=30/2 =15th item

$$= \text{size of } L + \frac{N/2 - cf}{f} \times C$$

$$= 15 + \frac{15 - 12}{8} x\ 5 = 46.47$$

Median marks = 16.875.

Mode

Mode is the modal value in the value of the variable which occurs more number of times or most frequently is a distribution. Mode is tge value which occurs with the greatest number of frequency in a series

Types of modal

I Unimodal

If there is only one mode in series is called unimodal

II Bi-Modal

If there are two modes in the series, it is called bi-modal

III Tri-Modal

If there are three modes in the series, it is Relationship between different Averages

Symmetrical = `called Tri-modal

IV Multimodal

If there are more than three modes in the series it is called multi-mode.

Relationship among mean median and mode

The three averages are identical, when the distribution is symmetrical. In an asymmetrical distribution, the values of mean, median and mode are not equal.

Median = 1/3 (Mean - mode)

Mode = 2 median - 2 mode

Median = Mode * 2/3 (Mean – Mode)

Individual Series

Calculate the mode form the following data of the marks obtain by 10 students

Serial No	1	2	3	4	5	6	7	8	9	10
Marks obtained	60	77	74	62	77	77	70	68	65	80

Solution

Marks obtained by 10 students is here 77 is repeated three times

Therefore the Mode mark is 77.

Discrete Series

Calculate the mode form the following data of the wages of workers of are establishment. find the modal wages

Daily wages in Rs	3	4	6	7	9	10	12	13	15
No of wage earners	2	3	2	6	10	11	12	5	1

Solution

Grouping Table

Daily Wages is Rs.	Frequency of Wages Earners					
	1	2	3	4	5	6
3	2	5		7		
4	3		5		11	
6	2	8				18
7	6		16	27		
9	10	21			33	
10	11		23			28
12	12	17		18		
13	5		6			
15	1					

Analysis Table

Column	Size of Item						
	4	6	7	9	10	12	15
1						I	
2				I	I	I	
3					I	I	
4			I	I	I		
5				I	I	I	
6					I	I	I
			I	3	5	4	1

From the analysis table it is known that size10 has been repeated the maximum number of times, thus is, so the modal wages Rs.10.

Continuous series

Find out the mode from the following series

X	0-5	5-10	10-15	15-20	20-25	25-30	30-35
frequency	1	2	5	14	10	9	2

Grouping Table

<table>
<tr><th rowspan="2">X</th><th colspan="6">Frequency</th></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr>
<tr><td>0-5</td><td>1</td><td rowspan="2">3</td><td></td><td rowspan="3">8</td><td></td><td></td></tr>
<tr><td>5-10</td><td>2</td><td rowspan="2">7</td><td rowspan="3">21</td><td></td></tr>
<tr><td>10-15</td><td>5</td><td rowspan="2">19</td><td rowspan="3">29</td></tr>
<tr><td>15-20</td><td>14</td><td rowspan="2">24</td><td rowspan="3">33</td></tr>
<tr><td>20-25</td><td>10</td><td rowspan="2">19</td><td rowspan="3">21</td></tr>
<tr><td>25-30</td><td>9</td><td rowspan="2">11</td><td rowspan="2"></td></tr>
<tr><td>30-35</td><td>2</td><td></td><td></td></tr>
</table>

Analysis Table

Column	**Size of Item**						
	0-5	5-10	10-15	15-20	20-25	25-30	30-35
1				I			
2			I	I	I		
3				I	I	I	
4				I	I	I	
5		I	I	I	I	I	I
6			I				
		1	3	6	5	3	1

Modal value lies in 15-20 as it occurs most frequently

$$\text{Mode (Z)} = L + \frac{f1 - f0}{2\,f1 - f0 - f2} \text{ x}C$$

$$\text{Mode (z)} = 154 + \frac{14 - 5}{2(14) - 5 - 10} \text{ X } 5$$

$= 15 + 9/13 \text{ X } 5 = 15 + 45/13$

$= 15 + 3.46$

Mode = 18.46

Geometric Mean

The Geometric Mean (GM) is the average value or mean which signifies the central tendency of the set of numbers by finding the product of their values. Basically, we multiply the numbers altogether and take the nth root of the multiplied numbers, where n is the total number of data values.

Merits of geometric Mean

- It is rigidly defined.
- It is based upon all the observations and all terms of sequences very precisely.
- It gives comparatively more weight of values to small items or data.
- It is not that much affected by fluctuations of sampling.

Demerits of Geometirc mean

- It is very difficult to calculate
- It is impossible to use it when any item is zero or negative
- The value of the geometric mean may not correspond with any actual value in the distribution

Uses of Geometric mean

The geometric mean is also used for number sets, where the values that are multiplied together are exponential. Examples of this phenomenon include the interest rates that may be attached to any financial investments, or the statistical rates of human population growth.

Individual Series

$$G.M = \text{Anti log of } \frac{\sum \log X}{N}$$

Calculate Geometric Mean

50	72	54	82	93

Solution

X	LogX
50	1.6990
72	1.8573
54	1.7324
82	1.9238
93	1.9685

$$G.M = \text{Anti log of } \frac{\sum \log X}{N}$$

$$= \frac{9.1710}{5} = 1.8342$$

= Antilog of 1.8342 = 68.26

Discrete Series

Calculate Geometric mean from the following data

Size of Item	120	125	130	135	136	138	139	140	147
Frequency	2	3	3	1	1	7	4	2	8

Solution

Size of Item (X)	**Frequency (f)**	**Log X**	**F log x**
120	2	2.0792	4.1584
125	3	2.0969	6.2907
1360	3	2.1139	6.3417
135	1	2.1303	2.1303
136	2	2.1335	4.2670
138	7	2.1399	14.9793
139	4	2.1430	8.5720
140	2	2.1461	4.2922
147	8	2.1673	17.3384
	N=∑f=32		N==∑f log x = 68.3700

$$G.M = \text{Anti log of } \frac{\sum \log X}{N} = \text{Anti log of } \frac{68.375}{32}$$

= Antilog of 2.1366 = 137

Therefore G.M = 137

Continuous series

Geometric mean from the following data

Yield of wheat	7.54-10.5	10.5-13.5	13.5 -16.5	16.5-19.5	19.5-22.5	22.5-25.5	25.5-28.5
No of forms	5	9	19	23	7	4	1

Solution

Yield of wheat	**Mid-Value**	**Log M**	**No. of Forms**	**f log m**
7.54-10.5	9	0.9542	5	4.7710
10.5-13.5	12	1.0792	9	9.7128
13.5 -16.5	15	1.1761	19	22.3459
16.5-19.5	18	1.2553	23	28.8719
			N=68	∑flog m=81.9092

$$G.M = \text{Anti log of } \frac{\sum f \log m}{N}$$

= 81.9092/68 = 1.204547

= Antilog of 1.204547 = 16.02

G.M = 16.02

Harmonic Mean

Meaning

Harmonic Mean is the reciprocal of the arithmetic average of the reciprocal of values of various item in the in variable

Merits of Harmonic Mean

- It utilizes all values of a variable
- It is very important to small values
- It is amenable to further algebraic manipulation
- It provides consistent results in problems relating to tiome and rates than similar averages

Demerits of Harmonic Mean

- It is not very easy to understand
- The method of calculation is difficult
- The presence of both positive and negative items in a series makes it impossible to compute its value. The same difficulty is felt if one or more items are zero
- It is only a summary figure and may not be the actual item in the series.

Individual Series

$$H.M = \frac{N}{\sum 1/x}$$

Calculate harmonic Mean for the following data:

Items	14	36	45	70	105

Solution

Computation of Harmonic Mean

X	1/X
14	0.7142
36	0.2777
45	0.0222
70	0.0142
105	0.0095
Total	**1.0378**

$$H.M = \frac{N}{\sum 1/x}$$

$$= 5 \ /1.0378$$

H.M = 4.81

Discrete series:

When data is given along with their frequencies. Following is an example of discrete series:

Items	14	36	45	70	105
Frequency	2	5	1	3	2

Solution

Size of Item X	Frequency f	1/x	F 1/x
14	2	0.1428	0.2856
36	5	0.1388	0.6940
45	1	0.0222	0.0222
70	3	0.0428	0.1284
105	2	0.0190	0.0380
	N=∑f =13		∑f 1/x = 1.1682

$$H.M. = \frac{N}{\sum f(1/x)}$$

$$= \frac{13}{1.1682}$$

H.M= 11.13

Continuous Series

Calculate Harmonic Mean for the following continous data:

Wages (Rs.)	0-10	10-20	20-30	30-40
No of workers	2	5	1	3

Solution

Size of Item X	Mid point	1/m	f	f 1/m
2	5	0.2000	2	1.0000
5	15	0.0067	5	0.0335
1	25	0.0400	1	0.0400
3	35	0.0285	3	0.0855
			11	1.1590

$$H.M. = \frac{N}{\sum f 1/m}$$

$$= \frac{11}{1.1590}$$

H.M= 9.49

Measures of Dispersion and Variability

Dispersion Meaning

Dispersion is the study of scatterness around an average

Definition

Dispersion is the measures of the variation of the items --- A.L.Bowly

Dispersion is a measure of extent to which the individual items vary -- L.R.Connor

Importance of measuring variation or dispersion

- Testing the Reliability of the Measures of Central Tendency
- Comparing two or more series on the basis of theirvariability
- Enabling to control the variability
- Facilitating as a Basis for further statistical Analysis

Characteristics of a Measure of Variation

- It is easy to understand and simple ro calculate
- It should be rigidly defined
- It should be based on all observations and it should not be affected by extreme observations
- It should be amenable to further algebraic treatment
- It should have sampling stability

Methods of Measuring Dispersion

1. Range
2. Inter Quartile range
3. Quartile Deviation
4. Mean Deviation

5. Standard Deviation
6. Lorenz Curve

Range

Range is the difference between the largest am d the smallest value in the distribution. Ot os the simplest and crudest measure of dispersion

Uses of Range

- It is used in industries for the statistical quality control of the m infected product
- It is used to study the variations such as stock, shares and other commodities
- It facilitates the use of other statistical measures

Advantages of Range

- It is the simplest method of studying variation
- It is easy to understand and the easiest to compute
- It takes minimum time to calculate
- It is accurate

Disadvantages of Range

- Range is completely depended on the two extreme values
- It is subject to fluctuations of considerable magnitude from sample to sample
- It is not suitable for mathematical treatment
- It cannot be applied to open and classes
- Range cannot tell us anything about the character of the distribution

Quartile deviation

Quartile deviation is an absolute measure of dispersion. It is calculated on the basis of the difference of upper quartile and the lower Quartile divided by 2

In the series, four quartiles are there. By eliminating the lowest (25%) items and the highest (25%) items of a series, we can obtain a measure of dispersion and can find out half the distance between the first and the third quartiles.

$$\text{Quartile Deviation (Q.D)} = \frac{Q3 - Q1}{2}$$

$$\text{Co-efficient of Q.D} = \frac{Q3 - Q1}{Q3 + Q1}$$

Merits of Quartile Deviation

- It is simple to calculate and easy to understand
- Risk of extreme item variance is eliminated, as it depend upon the central 50 per cent items
- It can be applied to open and classes
- Demerits of quartile Deviation
- Items below Q1 and above Q3 are ignored
- It is not capable of further mathematical treatment
- It is affected much by the fluctuations of sampling
- It is not calculated from a computed average, but from a positional average.

Mean deviation

Mean deviation is the average difference between the items in a distribution computed from the mean, median or mode of that series counting all such deviation as positive. The mean deviation is also known as the average deviation

$$\text{Mean deviation} = \frac{\sum |D|}{N}$$

$$\text{Co – efficient of Mean Deviation (M.D)} = \frac{MD}{\bar{X} \text{ or } Z \text{ or } M}$$

Merits of Mean Deviation

- It is clear and easy to understand
- It is based on each and every item of the data It can be calculated from any measure of central tendency and as such as flexible too.

Demerits of mean Deviation

- It is not suitable for further mathematical processing
- It is rarely used in sociological studies
- It is mathematically unsound and illogical, because the signs are ignored in the calculation of mean deviation

Standard deviation

Standard deviation is the square root of the means of the stranded deviation from the Arithmetic mean. So, it is also known as Root Mean Square Deviation an Average of Second order. Standard deviation is denoted by the small Greek letter 'σ' the concept of standard deviation is introduced by Karl Pearson in 1893.

Uses of Standard deviation

- It is used in statistics because it possesses must of the characteristics of an ideal measure of dispersion
- It is widely used in sampling theory and by biologists.
- It is applied in co-efficient of correlation and in the study of symmetrical frequency distribution

Advantages of standard deviation

- It is rigidly defined determinate.
- It is based on all the observations of a series.
- It is less affected by fluctuations of sampling and hence stable.
- It is amenable to algebraic treatment and is less affected but fluctuations of sampling most other measures of dispersion.
- The standard deviation is more appropriate mathematically than the mean deviation, since the negative signs are removed b y squaring the deviations rather than by ignoring co efficient of variance.

Standard deviation is an absolute measure of dispersion. The corresponding relative measure is known as the co-efficient of variation. It is used to compare the variability of two or more series.

$$\text{Co –efficient of Standard deviation} = \frac{\sigma}{X}$$

$$\text{Co-efficient of Variance (C.V)} = \frac{\sigma}{x} \times 100$$

Graphic method of dispersion

Lorenz Curve

Lorenz Curve is a device used to show the measurement of economic inequalities as in the distribution of income and wealth. It can also be used in business to study the disparities of distribution of profit, wages, turnover, production and the like.

Range

Range = L-S

$$\text{Co-efficient of range} = \frac{L-S}{L+S}$$

Solved Problems

Find the range and co-efficient of range for the heights of 8 students of a class

158,160,165,168,170,173.

Solution

Range = L – S

Given Series the largest value of the series = 173

Smallest value of the series = 158

Range = 178 – 158 = 15

$$\text{Co-efficient of range} = \frac{L-S}{L+S} = 0.045$$

Quartile Deviation

$$\text{Quartile Deviation} = \frac{Q3 - Q1}{2}$$

Individual Series

Find out the value of quartile deviation and its co-efficient from the following data

Roll No	1	2	3	4	5	6	7
Marks	22	30	42	32	52	62	54

Solution

Marks arranged in ascending order 22,30,32,42,52,54,62

$$Q1 = \text{size of } \frac{N+1}{4} \text{ the item}$$

$$= \text{size of } \frac{7+1}{4} \text{ the item} = 8/4 \text{ th Item}$$

$$= \text{size of 2nd item} = 30$$

$$Q3 = \text{size of } \frac{3(N+1)}{4} \text{ th item}$$

$$= \frac{3(7+1)}{4} = 3 \times 8 / 4 = 24/4 = 6 \text{ th item}$$

Size of the 6th item = 54

$$Q.D = \frac{Q3 - Q1}{2} = \frac{54 - 30}{2} = 12$$

$$\text{Co-efficient of Q.D} = \frac{Q3 - Q1}{Q3 + Q1} = \frac{54-30}{54+30} = 24/84 = 0.29$$

Discrete Series

From the following data calculate Quartile deviation an d its co-efficient

Height (in cm)	153	155	157	159	161	163	165	167	169
No.of students	8	2	4	6	3	4	7	1	4

Solution

x	F	cf
153	8	8
155	2	10
157	4	14
159	6	20
161	3	23
163	4	27
165	7	34
167	1	35
169	4	39
	N=∑f=39	

$$Q1 = \text{size of } \frac{N+1}{4} \text{ the item}$$

$$= \text{size of } \frac{39+1}{4} \text{ th item} = 40/4$$

= size of 10 th item = 155 centimeters

$$Q3 = \text{size of } \frac{3(39+1)}{4} \text{ th item}$$

$$= \frac{3(39+1)}{4} = 30 \text{ th Item} = 165 \text{ centimeters}$$

$$\text{Co-efficient of Q.D} = \frac{165 - 155}{165 + 155} = \frac{10}{320} = 0.03$$

Continuous series

From the following table. Compute the quartile deviation as well as its co-efficient

Size	0-5	5-10	10-15	15-20	20-25	25-30
Frequency	6	18	30	46	60	40

Solution

Weekly wages (x)	**No of workers f**	**Cf**
0-5	6	6
5-10	18	24
10-15	30	54
15-20	46	100
20-25	60	160
25-30	40	200
	N=∑f=200	

Q1 = N/4 200/4 = 50

Q1 is lies between the class 10-15

$$Q1 = 10 + \frac{N/4 - cf}{F} \times C$$

$$= 10 + \frac{50-24}{30} \times 5$$

$$= 10 + \frac{26}{30} \times 5$$

$$= 10 + 130/30$$

$$= 10 + 4.33$$

$$Q1 = 14.33$$

$$Q3 = 3N/4 = 3(200)/4 = 150$$

Q3 lies between the class interval 20 – 25

$$Q3 = L + \frac{3N/4 - cf}{f} \times C$$

$$= 20 + \frac{150-100}{60} \times 5$$

$$= 20 + 4.17$$

$$Q3 = 24.17$$

$$Q.D = \frac{Q3 - Q1}{2} = \frac{24.17 - 14.33}{2} = \frac{9.84}{2} = 4.92$$

Mean Deviation

Individual Series

$$M.D = \frac{\sum |D|}{N}$$

$$\text{Co efficient of M.D} = \frac{M.D}{\text{Mean or Median or Mode}}$$

Calculate mean deviation a d its coefficient from Arithmetic mean for the following the data

150	200	250	300	350

Solution

X	IDI = X-X
150	100
200	50
250	0
300	50
350	100
∑x= 1250	∑I DI=300

Mean = ∑x / N = 1250/5 = 250

$$M.D = \frac{\sum |D|}{N} = 300/5 = 60$$

$$\text{Co-efficient of M.D} = \frac{\text{Mean Deviation}}{\text{Mean}} = \frac{60}{250}\ 0.24.$$

Discrete Series

Find mean deviation from median and its coefficient from the following data

X	10	11	12	13	14
F	5	14	20	14	5

Solution

x	f	Cf
10	5	5
11	14	19
12	20	39
13	14	53
14	5	58

$$\text{Median} = \text{Size of } \frac{N+1}{2} \text{ th item}$$

$$= \text{Size of } \frac{58+1}{2} \text{ th item}$$

= Size of 29.5 th item = 12

x	f	IDI (X –Median)= x-12	fIDI
10	5	2	10
11	14	1	14
12	20	0	0
13	14	1	14
14	5	2	10
			∑fI DI = 48

$$\text{Mean Deviation} = \frac{\sum f|D|}{N} = \frac{48}{58} = 0.83.$$

$$\text{Co – efficient of M.D} = \frac{\text{M.D}}{\text{Median}} = \frac{0.83}{12} = 0.07.$$

Continuous Series

Find the co-efficient of mean deviation from mean for the following data

Age in years	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No of persons	25	30	37	45	47	40	15	13

Solution

Age in years	M	No of persons f	D=m-A	fd	IDI=m-X	fIDI
0-10	5	25	-30	-750	32	800
10-20	15	30	-20	-600	22	660
20-30	25	37	-10	-370	12	444

30-40	35	45	0	0	2	90
40-50	45	47	10	470	8	376
50-60	55	40	20	800	18	720
60-70	65	15	30	450	28	420
70-80	75	13	40	520	38	494
		N = 252		**∑fd=520**		**∑fl Dl = 4004**

$$\bar{X} = A + \frac{\sum fd}{N} = 35 + \frac{520}{252} = 35 + 2 = 37.$$

$\bar{X}$ = 37 Years

$$M.D = \frac{\sum fd}{N} = \frac{4004}{252} = 15.89.$$

$$\text{Co-efficient of M.D} = \frac{M.D}{Mean} = \frac{15.89}{37} = 0.43\text{years.}$$

Standard Deviation

Individual Series

Compute standard Deviation form the following data of the income of 10 employees of a firm

Monthly income	600	620	640	620	680	670	680	640	700	650

Solution

X	X=X-¯X	X2
600	-50	2500
620	-30	900
640	-10	100
620	-30	900
680	30	900
670	20	400

680	30	900
640	-10	100
700	50	2500
650	0	0
∑X = 6500		**∑ X2 == 9200**

$$\bar{X} = \frac{\sum X}{N} = \frac{6500}{10}\ 650$$

$$\sigma = \sqrt{\frac{\sum X^2}{N}} = \sqrt{9200 / 10} = 30.3 \qquad \sigma = 30.3$$

Discrete Series

Calculate standard deviation form the following data

Marks (X)	10	20	30	40	50	60
No of Students(f)	10	14	22	12	9	5

x	**F**	**fx**	**x = X-X̄ = X-31.5**	**X2**	**Fx2**
10	10	100	-21.5	462.25	4622.5
20	14	280	-11.5	132.25	1851.5
30	22	660	-1.5	2.25	49.5
40	12	480	8.5	72.25	867
50	9	450	18.5	342.25	3080.25
60	5	300	28.5	812.25	4061.25
	N=72	**∑fX =2270**			**∑ fx 2 =14532**

$$\bar{X} = \frac{\sum fX}{N} = \frac{2270}{72} = 31.5$$

$$\sigma = \sqrt{\frac{\sum fX^2}{N}} = \sqrt{14532 / 70} = 14.41$$

σ = 14.41.

Continuous series

Calculate standard deviation form the following data

Class	0-10	10-20	20-30	30-40	40-50
Frequency	5	8	15	16	6

Solution

Class	Mid point	Frequency	$d = \frac{X-A}{C}$	d2	fd	fd2
0-10	5	5	-2	4	-10	20
10-20	15	8	-1	1	-8	8
20-30	25	15	0	0	0	0
30-40	35	16	1	1	16	16
40-50	45	6	2	4	12	24
		N=50			**∑fd=10**	**∑fd2 =68**

Assumed mean A= 25

Class interval C = 10

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \times C$$

$$= \sqrt{68/50 - (10/50)^2} \times 10$$

$$= \sqrt{1.36 - (0.2)^2} \times 10$$

$$= 11.49$$

SKEWNESS

Introduction

The term ‘Skewness’ refers to lack of symmetry, that is, when a distribution is not symmetrical it is called a skewed distribution. It the curve us normal or the data rte distributed symmetrically or uniformly. Spread will be the same on both sides of the center point and the means median and mode will all have the same value.

Definition

'Skewness or symmetry is the attribute of a frequency distribution that extends further on one side of the class with the highest frequency on the other.

Simpson and Kafka

When a series is not symmetrical it is said to be asymmetrical or skewed -Croxton and Cowden

Skewness of a Distribution

When a distribution is not symmetrical it is called a skewed Distribution

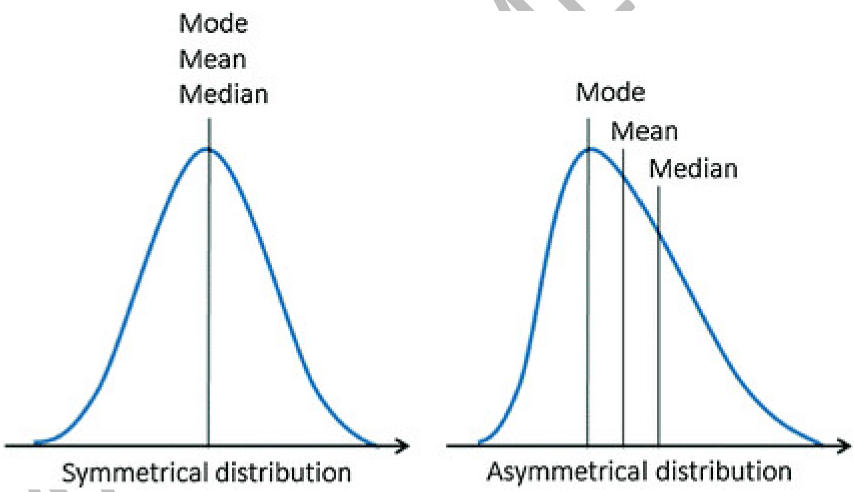
The analysis of presence of skewness in a distribution implies two main tasks. They are

I. Determination of the sign of skewness and testi9ng of skewness and

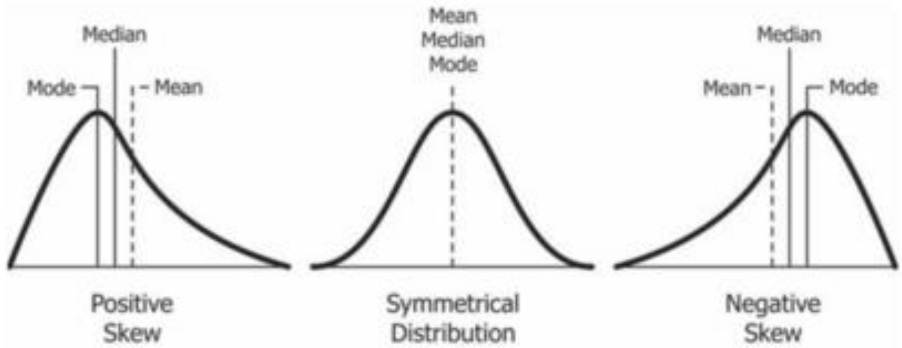
II. Determination of the extent of skewness

I Symmetrical Distribution

In a symmetrical distribution, the values of mean, median and mode are coinciding. The spread of the frequencies is the same on both sides of the centre point of the curve.



II Skewed Distribution



A distribution which is not symmetrical is called a skewed distributionit is called skewed distribution. it may be either positively Skewed or negatively skewed Distribution

I. **Positively Skewed Distribution**

In a frequency distribution positively skewed distribution the curve has longer tail to the rights and its value of the mean is highest and the made is least. The median lies in between the two. That is $\bar{X} > M > Z$

II. **Negatively Skewed Distribution**

In a frequency distribution if the curve has long tail to the left then it is negatively skewed distribution in which value ot mode is higher and mean is the least. The median lies in between the two. That is $\bar{X} < M < Z$

Various measures of Skewness

Skewness can be measured absolutely or relatively. Absolutely measures are called measures of skewness and relative measures are called the co-efficient of skewness.

Absolute measures of skewness

i. The Karl Peason's Coefficient of Skewness
ii. The Bowley's Co efficient of Skewness
iii. The Kelly's Coefficient of Skewness
iv. Measure of Skewness based on moments

Karl Pearson's Co-efficient of Skewness

This method is based upon the difference between mean and mode and the difference is divided by standard deviation to give a relative measures.

Bowley's Coefficient of Skewness

Bowelys measure is based on quartiles, In a symmetrical distribution first and third quartiles are equidistant from the median.

Objectives of Skewness

i. To find out the direction and extent of asymmetry in a series

ii. To compare two or more series with regards to skewness

iii. To study the nature of variation of the items about the central value

Karl Pearsons Coefficient of Skewness

Calculate Karl Pearson's coefficient of skewness for the following data

27	17	25	42	29	27	25	27	22

Solution

Size of the Item	Deviation d=X-A	D2
27	-2	4
17	-12	144
25	-4	16
42	13	169
29	0	0
27	-2	4
25	-4	16
27	-2	4
22	-7	49
	∑d = -20	**∑d2 = 406**

$$A + \frac{\sum d}{N} = 29 + \frac{-20}{9} = 29 - 2.22 = 26.78$$

$$(\sigma) = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} = \sqrt{406 / 9 - (-20/9)^2}$$

$$= \sqrt{45.11 - (2.2)^2} = 6.3$$

in the given series, 27 is repeated three times

Mode is 27

$$SKp = \frac{X - \text{Mode}}{\sigma} = \frac{26.78 - 27}{6.3} = -0.03$$

Bowley's Co-efficient of Skewness

Calculate the coefficient of Skewness

Age	0-10	10-20	20-30	30-40	40-50
No of Persons	8	11	26	9	6

Solution

Age	**No of Persons (f)**	**Cf**
0-10	8	8
10-20	11	19
20-30	26	45
30-40	9	54
40-50	6	60
	N = 60	

N/4 = 60/4 = 15 lies between the class interval 10 – 20

$$Q1 = L + \frac{N/4 - cf}{F} \times C$$

$$= 10 + \frac{15 - 8}{11} \times 10 = 10 + \frac{7}{11} \times 10 = 10 + 6.36 = 16.36$$

Q3 = 3N /4 = 3 (60)/4 = 45 lies between the CL 20 -30

$$Q3 = L + \frac{3N/4 - cf}{F} \times C = 20 + \frac{45 - 19}{26} \times 10 = 20 + \frac{28}{26} \times 10$$

$$= 20 + 10 = 30$$

N/2 = 60 /2 = 30 lies between the CL 20 - 30

$$Median = L + \frac{N/2 - cf}{F} \times C$$

$$= 20 + \frac{30 - 19}{26} \times 10 = 20 + \frac{11}{26} \times 10 = 20 + \frac{110}{26}$$

$$= 20 + 4.23 = 24.23$$

Bowley's Co – efficient of Skewness

$$SKp = \frac{Q3 + Q1 - 2Median}{Q3 - Q1} = \frac{30 + 16.36 - 2(24.23)}{30 - 16.3613.64} = \frac{48.38 – 48.46}{13.64} = \frac{-2.1}{}$$

= - 0.15

Kelly's Co efficient of Skewness

From the data given below calculate Kelly's co-efficient of skewness

Median =140

P10 = 29

P90 = 244

Solution

$$SKp = \frac{P90 + P10 - 2\ Median}{P90 - P10}$$

$$= \frac{244 + 29 - (2 X 140)}{244 - 29}$$

$$= \frac{273 - 280}{215}$$

$$= \frac{7}{215} = 0.033.$$

UNIT - IV

Correlation Analysis

Meaning

Correlation is the study of the natural relationship between two or more variables. Hence, it should be noted that the detection and analysis of correlation between two statistical variables requires relationship of some sort which associates the observation in pairs each of which is a value of the two variables.

Definition

The relationship that exists between two variables ---Smith

Correlation analysis deals with the association between two or more variables. –Tuite

Uses of Corrrelation

I. Correlation is very useful in physical and social sciences. Business and economics

II. Correlation analyses is very useful in economics to study the relationship between price and demand

III. It is also useful in business to estimates costs, value, price and other related variables

IV. Correlation is the basis of the concept of regression

V. Correlation analysis help in calculation the sampling once.

Types of Corrrelation

There are several types of correlation commonly used in statistical analysis:

1. **Positive Correlation:** Both variables move in the same direction. As one increases, the other also increases (e.g., height and weight).
2. **Negative Correlation:** The variables move in opposite directions. As one increases, the other decreases (e.g., temperature and heating costs).
3. **Zero Correlation:** There is no relationship between the variables; changes in one do not affect the other (e.g., shoe size and intelligence).
4. **Linear Correlation:** The relationship between the variables can be represented by a straight line. This is often assessed using Pearson's correlation coefficient.
5. **Non-Linear Correlation:** The relationship is not linear and may follow a different pattern, which can be assessed using methods like Spearman's rank correlation.

6. **Partial Correlation:** This measures the relationship between two variables while controlling for the effect of one or more additional variables.
7. **Multiple Correlation:** This assesses the relationship between one variable and two or more other variables, providing insights into complex relationships.

These types help researchers understand the nature and strength of relationships between variables in various contexts.

Methods of studying correlation

Graphical method

- Scatter diagram
- Simple graph method

Mathematical Methods

- Karl Pearson's Co-efficient of correlation
- Spearman's Rank Correlation coefficient
- Concurrent deviation method
- Method of least square

Scatter diagram method

It is a method of studying correlation between two related variables. The two variables X and Y will be taken upon the X and Y axes of a graph paper. For each part of X and Y values, we mark a dot and we go as many points as the numbers of observation.

Graphical method

In this method curves are drawn for separate series on a graph paper. By examining the direction and closeness of the two curves we can offer whether prompt variances are related. If both the curves are moving in the same direction correlation is said to be positive. On the contrary, if the curves are moving in the opposite directions is said to be negative

Karl Pearson's Co-efficient of correlation

Karl Pearson, a great statistician introduced a mathematical method for measuring the magnitude of relationship between two variables. This method. Known as Pearson Coefficient of correlation is widely used. It is denoted by the symbol "r'

Spearman's Rank Correlation Co-efficient

In 1904, a famous British psychologist Charles Edward Spearman found ou the method of Co-efficient of correlation of rank. Rank correlation is applicable to individual observation. This measure is useful in dealing with qualitative characteristics. The result, by using ran king method, is only approximate.

Concurrent deviation method

Under this method, the direction of change in X variable and y variable is taken into account to find out the deviation for each term the change in the value of the variable form its preceding value which may be + or –

Co-efficient of correlation

Find the Karl Pearson's Coefficient of Correlation

X	6	2	10	4	8
Y	9	11	5	8	7

Solution

X	Y	X2	Y2	XY
6	9	36	81	54
2	11	4	121	22
10	5	100	25	50
4	8	16	64	32
8	7	64	49	56
30	**40**	**220**	**340**	**214**

$$\text{Coefficient of Correlation} = \frac{N\sum XY - (\sum X)(\sum y)}{\sqrt{N\sum X2 - (\sum x)2} \cdot \sqrt{N\sum Y2 - (\sum Y)2}}$$

$$r = \frac{(5 \text{ X } 214) - (30 \text{ X } 40)}{\sqrt{5 \text{ X } 220 - (30)2}\ \sqrt{5 \text{ x } 340 - (40)2}}$$

$$= \frac{1070 - 1200}{\sqrt{1100-900}\ \sqrt{1700-1600}} = \frac{-130}{\sqrt{200}\ \sqrt{100}} = \frac{-130}{14.14(10)} =- 0.9194$$

Rank Correlation

Two judges ina beauty contest rank the 12 entries as follows

X	1	2	3	4	5	6	7	8	9	10
Y	6	5	3	1	2	4	7	10	9	8

Rank X	Rank Y	D=R(X)-(Y)	D2
1	6	-5	25
2	5	-3	9
3	3	0	0
4	1	3	9
5	2	3	9
6	4	2	4
7	7	0	0
8	10	-2	4
9	9	0	0
10	8	2	4
N=10			**∑D2 = 64**

$$R = 1 - \frac{6\sum D2}{N(N2-1)}$$

$$= 1 - \frac{6(64)}{10(102-1)} = 1 - \frac{384}{990} = 1 - 0.39 = 0.61$$

Calculate the co-efficient of concurrent deviation from the following data

X	65	50	45	60	40	55	30	40	60	60	55
Y	70	55	50	70	60	85	50	40	70	55	5

Solution

X	Direction of change (dx)	Y	Direction of change (dy)	Dxdy
65		70		
50	-	55	-	+
45	-	50	-	+
60	+	70	+	+
40	-	60	-	+
55	+	85	+	+
30	-	50	-	+
40	+	40	-	-
60	+	70	+	+
60	0	55	-	0
55	-	55	0	0
				C = 7

$$C = \frac{+\sqrt{2C - N}}{N}$$

$$= \sqrt{\frac{2X\,7 - 10}{10}}$$

= √4 / 10 = 0.6333

Regression Analysis

Meaning

The statistical method employed to estimate the unknown valued of one variable from the known value of the related variables is called regression

Definition

Regression is the measure of the average relationship between two or more variables in terms of the original units of the data ------ Blair

Regression analysis Meaning regression analysis is statistical device with which we estimate or predict the unknown values of one variable from known value of another variable

Regression analysis definition

One of the most frequently used techniques in economics and business research, top fin d a relation between two or more variables that are related causally, is regression analysis. --- Taro Famane

Uses of regression analysis

- It is useful to estimate the relationship between two variables
- It is useful for production of unknown value
- It is widely used in social sciences like economics, Natural and physical sciences
- It is useful to forecast the business situation
- It is useful to calculate correlation co-efficient and co-efficient of determinations

Methods of studying Regression

- Graphic method
- Algebraic method

Graphic method

Under the method the dots are plotted on a graph paper representing pair of values of thje given variables having a linear relationship The independent variable is taken in the X axis and the dependent variable taken on Y axis. The regression line of X on Y provides the most probable value of X given the most probable value of Y when the exact value of X is known. Thus we get two regression lines

Regression lines

I. Regression of X on Y
II. Regression of Y on X

Algebraic Method

Regression equation is an algebraic method. It is an algebraic expression of the regression line.

Regression Equations

Regression Equation of X on Y

Xc = a+ by

Regression Equation of Y on X

Yc = a + bx

Solved Problems

Find regression lines by using actual Mean

X	6	2	10	4	8
Y	9	11	5	8	7

Solution

X	X = (X-X)	X2	Y	Y = (Y-Y)	Y2	Xy
6	0	0	9	1	1	0
2	-4	16	11	3	9	-12
10	4	16	5	-3	9	-12
4	-2	4	8	0	0	0
8	2	4	7	-1	1	-2
∑X = 30	**∑x = 0**	**∑x2 = 40**	**∑Y = 40**	**∑y = 0**	**∑y2 =20**	**∑xy = -26**

$$\bar{X} = \frac{\sum x}{n} = 30 / 5 = 6$$

$$\bar{Y} = \frac{\sum y}{n} = 40 / 5 = 8$$

$$X - \bar{X} = r \frac{\sigma x}{\sigma y} (Y - \bar{Y})$$

$$r = \frac{\sum xy}{\sum y2}$$

$$r = \frac{-26}{20} = -1.3$$

x-6= -1.3 (y-8)

x=-1.3y+10.4+6

x=-1.3y +16.4

Regression equation of Y on X

$$(Y - \bar{Y}) = r = \frac{\Sigma y}{\sigma x} (X - \bar{X})$$

$$r = \frac{\sum xy}{\sum x^2}$$

$$r = \frac{-26}{40} = -0.65$$

y-8= -0.65 (x-6)

y-8=-0.65+3.9

y=-0.65x +3.9+8

y=-0.65x +11.9

Unit V

Index numbers

As index number is a specialized average designed to measure the change in a group of related variable over a period of time. It was first constructed in the year

Concept

In its simplest form on Index number is a Ratio of two numbers expressed as percent.

Definition

Index numbers are standardized numerical values that measure the relative change in a variable or a set of variables over time or between different groups. They are often used to compare levels of economic activity, prices, or quantities. An index number is typically expressed as a percentage of a base value, allowing for easy interpretation of trends and comparisons.

For example, if the price index for a basket of goods rises from 100 to 120, it indicates a 20% increase in prices compared to the base period. Index numbers are commonly used in areas like inflation measurement (Consumer Price Index) and economic performance (Gross Domestic Product index).

Characteristics of Index number

Index numbers have several key characteristics that make them useful for analysis in statistics. Here are some of the main characteristics:

1. **Relative Measure:** Index numbers express the change of a variable relative to a base value, allowing for easy comparison over time or between different groups.
2. **Base Year:** They are typically anchored to a specific base year or period, which serves as the reference point for comparison.
3. **Percentage Form:** Index numbers are usually presented as percentages, making it simple to interpret the magnitude of change.
4. **Aggregation:** They can aggregate multiple items into a single index, allowing for the analysis of overall trends in complex datasets (e.g., price indices for various goods).
5. **Variability:** Index numbers can represent different types of data, such as prices, quantities, or values, and can vary in their construction methods (e.g., simple vs. weighted indices).
6. **Comparative Analysis:** They facilitate the comparison of economic indicators across different time periods, regions, or sectors.
7. **Simplicity:** They simplify complex data into a single number, making it easier for decision-makers to understand trends and make informed choices.

8. **Sensitivity to Base Year:** The choice of base year can significantly affect the index number, so it's important to select it carefully.

These characteristics make index numbers a powerful tool in economic analysis, business metrics, and various other fields.

Uses of Index Number

- Index number is most widely used statistical devises
- Index numbers are used to measure the relative changes
- They are widely used in the evaluation of business and economic conditions
- It is useful for better comparison
- It is a good guide for the progress of every country
- It is useful for better comparison
- It is useful to know trends and techniques
- For forecasting future activities

Types of Index Numbers

- Price Index
- Quantity Index
- Value Index

Methods of Index Number

- Unweighted Index Number
- Simple Aggregative method
- Simple average of price Relative method

Weighted Index Number

1. Laspeyre's index number $P01 = \dfrac{\sum p1q0}{\sum P0q0} \times 100$

2.Paaschel's Method

. $P01 = \dfrac{\sum p1q1}{\sum P0q0} \times 100$

3.Bowley and Dorfish method

$$P01 = \frac{\frac{\sum p1q0}{\sum P0q0} + \frac{\sum p1q0}{\sum P0q0}}{2}$$

4.Fisher's Ideal method or Fishers Price Index Number

$P01 = \sqrt{L \times P}$

Consumer price Index Number (or) Cost of living index

Consumer price Index is designed to measure the change in the cost of living of workers because of change in the retail price. A change in the price level affects the cost of living of the people. People con some different types of commodities. So there is need to construct consumer's price index. Consumer price index can be used in different places for many purposes.

Uses of cost of living index

- It is useful in fixing the wages
- It is useful to know the purchasing power of money
- By using the cost of living index the Government determines the price and other variables
- It is useful I the analysis of price situations

Limitations of Index numbers

- If the chosen base year is not a normal one, the purpose is list
- Every index number has its own purpose. No index number can serve all puipose
- These are only appropriate indications of the relative level.

Solved Problems

Construct an index number for 2019 taking2020 as base from the following data

Commodity	price in 2019(Rs)	Price in 2020 (rs)
A	60	70
B	50	90
C	80	120
D	100	80
E	60	50

Solution

Commodity	price in 2019(Rs)	Price in 2020(rs)
A	60	70
B	50	90
C	80	120
D	100	80
E	60	50
	$\sum p0 = 350$	**$\sum p1 = 410$**

$$\text{Price Index } P01 = \frac{\sum p1}{\sum p0} \times 100 = \frac{410}{350} \times 100 = 117.$$

This means that we compared to 2013, in 2014 there is a net increase in the prices of commodities to the extent of 17%.

Compute Fisher's Ideal Index from the following data

Item	2005		2007	
	Price	Quantity	Price	Quantity
A	8	50	12	56
B	4	200	4	120
C	6	60	8	60
D	12	30	14	24
E	10	40	14	36

Solution

COMMODITY	P0	Q0	P1	Q1	p1q0	P0q0	p1q1	P0q1
A	8	50	12	56	500	400	672	448
B	4	100	4	120	200	400	480	480
C	6	60	8	60	360	360	480	360
D	12	30	14	24	360	360	336	288
E	10	40	14	36	480	400	504	360
					$\sum$ p1q0=1900	$\sum$ P0q0 =1920	$\sum$ p1q1 =2472	$\sum$ P0q1 =1936

P01 = √ L X P

$$P01 = \sqrt{\frac{\sum p1q0}{\sum P0q0} \times \frac{\sum p1q1}{\sum P0q1}} *100$$

$$P01 = \sqrt{\frac{1900}{1920} \times \frac{2472}{1936}} *100$$

= √ 0.9896 X 1.2769 X 100

= 2.2692 X 100

= 226.92.

Analysis of time Series

An arrangement of statistical data in accordance with time of occurrence or in chronological order is called a time series

Definition

A time series is a set of observation arranged in chronological order. Morris Hamberg Requirement of a time series

Data must be available for a long period of time

Data must consist of a homogeneous set of values belonging to different time periods

The time gap between the variables or composite of variables must be as For Passable equal

Uses of Time series

I. It helps in understanding the past behaviors and in establishing the future behavior
II. It helps in planning and forecasting the future operation
III. It facilitates comparison between data of one period with those of another period
IV. It helps in evaluating current accomplishment
V. It is useful in forecasting the trade cycles

Time Series Models

Mathematical Models and Multiplicative method

In classical analysis, it is assumed that some types of relationship exist among the four components of time series

1. Additive Model

According to this model, the time series is expressed as

Y = T + S + C + I

Y = The value of original time series

T = Time Value

S = Seasonal variation

C = Cyclical Variation

Irregular fluctuation

Multiplicative Model

According this model, the time series is expressed as

Y = Y X S X C X I

Time series Analysis

- Time series analysis is the analysis of identifying different components such as trend, seasonal, cyclical and irregular in a given time series data.
- Components of time series
- Time series data contain variations of the following types
- Secular Trend II) Seasonal Variation III) Cyclical Variations IV) Irregular variation

Secular Trend

A secular trend or long-term trend refers to the movements of the series reflecting continuous growth or decline over a long period of time. There are many types of trend. Some trends rise upward and some fall downward

Seasonal variation

Is that periodic investment in business activities within the year recurring periodically year after year

Generally, seasonal variation appear at weekly, monthly or quarterly intervals

Cyclical Variation

Up and down movements are different from seasonal fluctuations, in that they extend over longer period of time – usually two or more years.

Business time series is influenced by the wave-like changes of prosperity and depression

Causes Changes of property and depression Uses

I. Useful to study the character of business fluctuation
II. Useful to take timely decision in maintaining the business during different stages
III. Helps in facing recession and utilizing the booms

Unsecular Variation

Irregular variation refers to such variation in business activity which do not repeat in a definite pattern. They are also called 'erratic 'accidental or random variations which are generally non-recurring and unpredictable

Causes

War, food, revolution, strike, lockouts and the like.

Measurement of Secular of secular trend

a) Free hand Graphic Method
b) In this method we must plot the original data on the graph. Draw a smooth curve carefully which will show the direction of the trend. The ti9me is taken on the horizontal axis I(X) and the value of the variable on the vertical axis (Y)

Merits

(i) It is the simplest and easiest method
(ii) It can be applied to all types of trends
(iii) It is useful to understand the character of time series

Demerits

(i) It is subject to personal bias
(ii) Its results depend upon the judgments of the person who draw the time
(iii) It does not help to measures trend

Semi – Average Method

In this method the original data are divided into two equal parts and average are calculated for both the parts. These averages are called semi Average. trend line is drawn with the help of the semi averages

Merits

(i) It is simple and easi9er to understand
(ii) Everyone will get the same trend like
(iii) We can predict the future values based on the intermediate values

Demerits

(i) It is affected by the limitations of arithmetic mean

(ii) It is not enough for forecasting the future trend

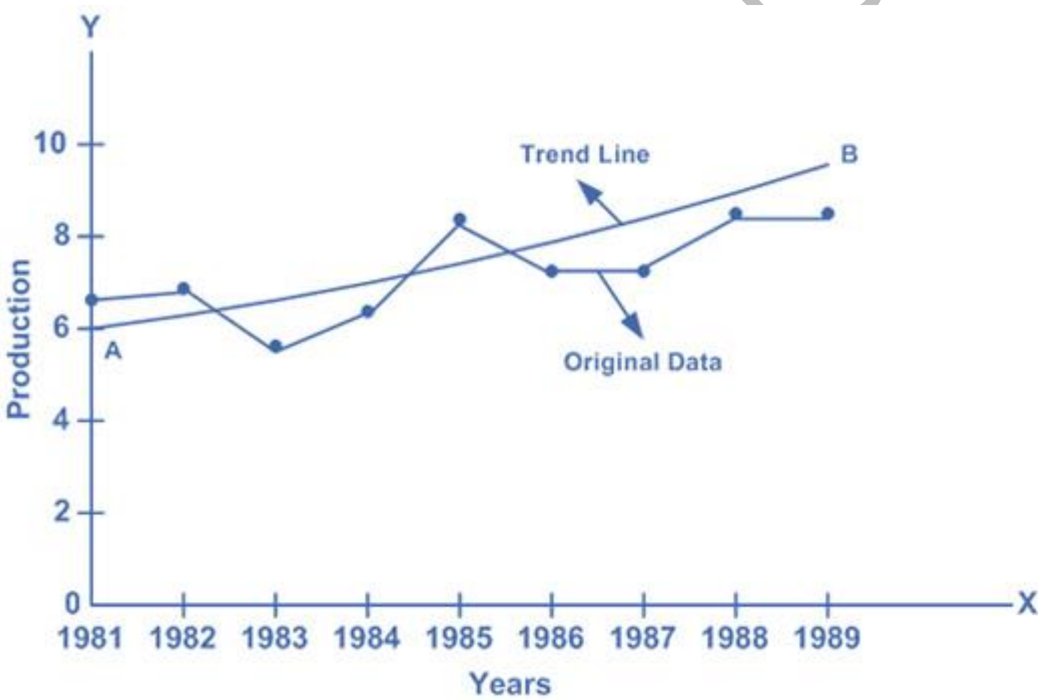
Moving average method

In this method, the average value of a number of years or months or weeks is taken into account and placed it at the centre of the time span and it nis is the normal or trend value for the middle period.

Solved Problems

Free-Hand Trend line1985

Year	1981	1982	1983	1984	1985	1986	1987	1988	1989
Production tons	0	2	40	6	8	10	6	8	8



Moving Average

1.Find the 3 yearly moving average from the following time series data

Year	2000	2001	2002	2003	2004	2005	2006	2007
Sales (In tones)	53	64	72	80	76	84	104	96

Solution

Year	Sales (in tons)	3yearly Moving Total	3 Yearly moving value
2000	53	-------	-------
2001	64	189	63
2002	72	216	72
2003	80	228	76
2004	76	240	80
2005	84	264	88
2006	104	284	95
2007	96	-------	--------

2. Calculate the 5 yearly moving average from the following data

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Number of students	705	685	703	687	705	689	715	685	725	730

Solution

Year	No of students	5 yearly moving total	Moving value
1998	705	--	--
1999	685	--	--
2000	703	3485	697.0
2001	687	3469	693.8
2002	705	3499	699.8
2003	689	3481	696.2
2004	715	3519	703.8

3. Calculate the Four – Yearly moving average for the following data

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Production (in '000Tons)	464	515	518	467	502	540	557	570	586	612

Solution

Year	Production in '000 Tons	4 yearly Moving	Combined Total	Moving Average
1998	464			
1999	515			
		1964		
2000	518		3966	495.75
		2002		
2001	467		4029	503.63
		2027		
2002	502		4093	511.63
		2066		
2003	540		4236	529.50
		2170		
2004	557		4424	553.00
		2254		
2005	570		4580	572.50
		2326		
2006	586			
2007	612			

4. Compute the trend from the following by the method of least square method

Year	2005	2006	2007	2008	2009
Production in Lakhs	70	74	80	86	90

Solution

Computation of trend Values

Year	Production in Lakhs (Y)	Deviation from 2002(X)	XY	X2
2005	70	-2	-140	4
2006	74	-1	-70	1
2007	80	0	0	0
2008	86	1	86	1
2009	90	2	180	4
N=5	**∑y = 400**	**∑x = 0**	**∑xy=52**	**∑x2 =10**

Since ∑x = 0

$a = \frac{\sum y}{N} = 400 / 5 = 80$

$b = \frac{\sum xy}{\sum x^2} = 52/10 = 5.2$

Y = a + bx =80 + 5.2x

Yc = 80 + 5.2 x

When x= -2

Y2005 = 80 + 5.2 (-2) = 80 -10.4 = 69.6

When X = -1

Y2006= 80 + 5.2 (-1) = 80 -5.2 = 74.8

When X = 0

Y2007 =80 + 5.2 (0) = 80 -0 = 80

When X =1

Y2008= 80 + 5.2 (-1) = 80 +5.2 = 85.2

When X=2

Y2009 = 80 + 5.2 (2) = 80 + 10.4 = 90.4.