

Fundamentals of Business Mathematics and Statictics

PAPER - IV

THE INSTITUTE OF COST ACCOUNTANTS OF INDIA
(Statutory body under an Act of Parliament)
www.icmai.in

SYLLABUS - 2016

WORK BOOK

FUNDAMENTALS OF BUSINESS MATHEMATICS AND STATISTICS

FOUNDATION

PAPER – 4

The Institute of Cost Accountants of India
(Statutory body under an Act of Parliament)
www.icmai.in

First Edition : May 2019

Published By :

Directorate of Studies

The Institute of Cost Accountants of India

CMA Bhawan, 12, Sudder Street, Kolkata – 700 016

www.icmai.in

Copyright of these study notes is reserved by the Institute of Cost Accountants of India and prior permission from the Institute is necessary for reproduction of the whole or any part thereof.

Preface

Professional education systems around the world are experiencing great change brought about by the global demand. Towards this end, we feel, it is our duty to make our students fully aware about their curriculum and to make them more efficient.

Although it might be easy to think of the habits as a set of behaviours that we want students to have so that we can get on with the curriculum that we need to cover. It becomes apparent that we need to provide specific opportunities for students to practice the habits. Habits are formed only through continuous practice. And to practice the habits, our curriculum, instruction, and assessments must provide generative, rich, and provocative opportunities for using them.

The main purpose of this volume is to disseminate knowledge and motivate our students to perform better. Thus, we are delighted to inform about the **e-distribution of the first edition of our 'Work book' for Foundation level.**

This book has been written to meet the needs of students as it offers the practising format that will appeal to the students to read smoothly. Each chapter includes unique features to aid in developing a deeper under-standing of the chapter contents for the readers. The unique features provide a consistent reading path throughout the book, making readers more efficient to reach their goal.

Discussing each chapter with illustrations integrate the key components of the subjects. In the second edition, we expanded the coverage in some areas and condensed others.

It is our hope and expectation that this second edition of work book will provide further an effective learning experience to the students like the first edition.

The Directorate of Studies,

The Institute of Cost Accountants of India

FUNDAMENTALS OF BUSINESS MATHEMATICS AND STATISTICS

FOUNDATION

PAPER – 4

INDEX

Sl. No.	Section – A : Business Mathematics	Page No.
1	Arithmetic	1 – 16
2	Algebra	17 – 28
	Section – B : Statics	
3	Statistical Representation of Data	29 – 40
4	Measures of central Tendency and Dispersion	41 – 102
5	Correlation and Regression	103 – 118
6	Probability	119 – 140

Suggested Marks Distribution from Examination Point of View

Only for Practice Purpose

Total 100 Marks	3 Hours	MCQ = 60 Marks
		Others = 40 Marks

Objective Question

60 Marks (1 Mark each questions)	MCQ
	True/False
	Matching

Short Notes

Minimum Marks for each Questions	4 Marks
Maximum Marks for each Questions	7 Marks

Practical Problem

Minimum Marks for each Questions	6 Marks
Maximum Marks for each Questions	8 Marks

SECTION – A : BUSINESS MATHEMATICS

Study Note – 1

Arithmetic

Learning Objective:

After studying this chapter, the student will be able to:

- ***Appreciate the patterns in numbers***
- ***Find the first term and common difference***
- ***Cite examples for Arithmetic Progression***
- ***Indicate given fractions on a diagram and find equivalent fractions for a given fraction***
- ***Choose appropriate units for a given purpose.***

1. Multiple-Choice Questions:

(i) **If $\frac{1}{2}$ of money of A = $\frac{1}{3}$ money of B = $\frac{1}{4}$ of money of C, then the continued ratio of money of A, B and C**

(a) 2 : 3 : 4 (b) 6: 4 :3 (c) 4 :3 : 2 (d) 3 : 2 : 1

Answer: (a)

L.C.M of 2, 3, 4 is 12 and let A's amount is ₹ 12. Then the amount of B and C are respectively ₹ 18 and ₹ 24. Hence A : B : C = 2 : 3 : 4.

(ii) **The ratio of two numbers is 11:15. The sum of 3 times the first number and twice the second number is 1260. The H. C. F. of the number is:**

(a) 10 (b) 12 (c) 15 (d) None of these

Answer: (d)

We have a : b = 11 : 15, ⇒ 15a = 11b. It also given that 3a + 2b = 1260

⇒ 15a + 10b = 1260 × 5, ⇒ 11b + 10b = 1260 × 5, ⇒ b = 300. Hence a = 220.

So, H.C.F. of the given numbers is 20.

(iii) **Of the four numbers in proportion, if the product of two extreme numbers is 51, the other numbers are:**

(a) 32, 16 (b) 18, 30 (c) 3, 17 (d) 6, 24

Answer: (c)

Let four numbers a, b, c, d are in proportion, then a : b = c : d, ⇒ad = bc, ⇒bc = 51
Hence the numbers are 3 and 17.

(iv) The ratio of the present age of mother to her daughter is 7 : 3. Ten years hence the ratio would be 9 : 5. Then their present ages are:

(a) 49 years; 21 years (b) 35 years; 15 years (c) 42 years; 18 years (d) 56 years; 24 years

Answer: (b)

Let the present age of mother and her daughter is 7x and 3x years.
As per sum (7x + 10) : (3x + 10) = 9 : 5, ⇒ 5(7x + 10)
= 9(3x + 10), ⇒ 8x = 40, ⇒ x = 5.
Their present ages are 35 years and 15 years.

(v) If $\frac{a+b}{a-b} = 2$ then the value of $\frac{a^2 - ab + b^2}{a^2 + ab + b^2}$ is :

(a) $\frac{13}{7}$ **(b)** $\frac{5}{13}$ **(c)** $\frac{7}{13}$ **(d) None of these.**

Answer: (c)

Given that $\frac{a+b}{a-b} = 2 \Rightarrow \frac{(a+b)+(a-b)}{(a+b)-(a-b)} = \frac{2+1}{2-1}$ **(Componendo- Divedendo).** $\Rightarrow \frac{a}{b} = 3$

Now, $\frac{a^2 - ab + b^2}{a^2 + ab + b^2} = \frac{\left(\frac{a}{b}\right)^2 - \left(\frac{a}{b}\right) + 1}{\left(\frac{a}{b}\right)^2 + \left(\frac{a}{b}\right) + 1} = \frac{7}{13}$.

(vi) The simple interest on ₹10 for 4 months at the rate of 3 paise per rupee per month is:
(a) ₹ 1.20 (b) ₹ 12 (c) ₹ 120 (d) ₹ 1200

Answer: (a)

We have $I = \frac{PRT}{100}$, Here P = ₹ 10, T = 4 months = $\frac{1}{3}$ years

and $R = \left(\frac{3}{100} \times 100 \times 12\right)\% = 36\%$. Hence $I = ₹\left(\frac{10 \times 1 \times 36}{100 \times 3}\right)$ ₹120

(vii) On what sum of money will the simple interest for 3 years at 10% per annum be half of the compound interest on ₹400 for 2 years at 6% per annum?

(a) ₹ 104.6 (b) ₹ 82.4 (c) ₹ 98.2 (d) ₹ 120.8

Answer: (b)

Here P = ₹ 400, n = 2 Years and i = 0.06 then Amount = $P(1+i)^n$

$= 400(1.06)^2 = 449.44$. So, Compound Interest = $(449.44 - 400) = 49.44$.

In case of simple interest let, Principal = ₹ Y. As per sum, $\frac{Y \cdot 3 \cdot 10}{100} = \frac{49.44}{2} \Rightarrow Y = 82.4$.

(viii) A tradesman marks his goods with two prices, one for ready money and the other for 4 month's credit. What ratio should two prices bear to each other, allowing 10% simple interest?

(a) 20 : 11 (b) 35 : 36 (c) 40 : 41 (d) 30 : 31

Answer: (d)

Late the price of goods for ready money and for 4 month's credit are respectively ₹x and ₹y.

The interest on ₹x for 4 months at 10% p.a. $\frac{x \times 10}{100 \times 3} = \frac{x}{30}$

Thus, as per sum, $y = x + \frac{x}{30} = \frac{31}{30}x \rightarrow x : y = 30 : 31$.

(ix) A man deposited a sum of money to a bank at 12% simple interest p.a. The total interest that he will get at the end of 4 years is ₹1620. The deposit amount is

(a) ₹3600 (b) ₹3025 (c) ₹4400 (d) ₹3375

Answer: (d)

Here T = 4 years, R = rate of interest p.a. =12, Total Interest = 1620.

We know that, $\frac{100 \times 1}{T \times R} = \frac{100 \times 1620}{4 \times 12} = 3375$.

(x) The year by which a sum of rupees would be 1.36 times of itself at 8% p.a. C.I. is

(a) 4 years (b) 5 years (c) 3 years (d) 6 years

Answer: (a)

Let the Principal is ₹P. Here A = Amount = 1.36P and i = interest on Rupee 1 for 1 year = 0.08. We know, amount $(A) = P(1+i)^n = P \times (1.08)^n$

$\Rightarrow 1.36P \; P \times (10.8)^n. \Rightarrow 1.36P = (10.8)^n. \Rightarrow n = \frac{\log 1.36}{\log 1.08} = 3.995 \approx 4$

(xi) Some money is distributed between A and B in the ratio 2: 3. If A receives ₹72, then b receives
(a) ₹90 (b) ₹144 (c) ₹108 (d) None of these

Answer: (c)

Let A and B respectively receives ₹2x and ₹3x. It also be given than A receives ₹72. So, x = 36. Hence B will receives ₹3 × 36 = ₹108

(xii) If the terms – 1 + 2x , 5, 5 + x from an A. P., then the value of x is
(a) 2 (b) 3 (c) 4 (d) 5

Answer: (b)

Since – 1 + 2x, 5 and 5 + x forms an A. P. So, 5 – (-1 + x) = (5 + x) – 5
⇨ 6 – x = x
⇨ X = 3.

(xiii) If the sum of three numbers in A. P is 18, then the middle term is
(a) 6 (b) 4 (c) 8 (d) 10

Answer: (a)

Let the three terms of the A. P. are (a – d), a and (a+ d)
As per Sum, (a- d) + a + (a + d) = 18 =>a = 6.

(xiv) You save 1p to-day, 2p to-morrow and 3p day after tomorrow. How much you can save in 1 year (1 year = 365 days)?
(a) 696.70 (b) 667.95 (c) 766.94 (d) 676.75

Answer: (b)

Actually, we have to find the sum 1 + 2 +3 + ….. upto 365 terms.
We know that the sum of n terms of an A. P. having 1st term and common difference are a and d respectively is $\frac{n}{2}\{2a+(n-1)d\}$.

Total saving is $\frac{365}{2}\{2\times1+364\times1\}=365\times183=66795.$

Hence, I can save ₹667.95 in 1 year.

(xv) If $b\propto a^3$ and a increases in the ratio 3:2 then b increases in the ratio
(a) 8:27 (b) 27:8 (c) 2:3 (d) None of these

Answer: (b)

Let b increases in the ratio k:1 i.e. b will be kb when a increases in the ratio 3:2.
Now b α a^3 => b = ka^3, Let presently a = 2x and it will be 3x after increase.

Hence b = $k8x^3$ => $\frac{27}{8}b = k(3x)^3 \Rightarrow k = \frac{27}{8}$. creases in the ratio 27:8.

(xvi) A.M and G.M of two positive integers a and b (a<b) are respectively 5 and 4; value of a and b are

(a) 2,8 (b) 6,9 (c) 8,10 (d) 4,6

Answer: (a)

As per sum, a + b = 10 and ab = 16.

$\Rightarrow (a-b)^2 = (a+b)^2 - 4ab \Rightarrow a-b = \pm 6$.

As a > b, a – b = - 6 => a = 2, b = 8.

(xvii) The 2nd term of a G.P is b and common ratio is r. If the product of the first three terms is 64, the value of b is

(a) 4 (b) 3 (c) 5 (d) 6

Answer: (a)

Let three numbers in G.P. are $\frac{b}{r}$, b and br.

It is given that their product is 64 $\Rightarrow b^3 = 64 \Rightarrow b = 4$.

(xviii) The first and last terms of an A.P are -4 and 146 and the sum is 7171. The number of terms in this A.P. is

(a) 98 (b) 100 (c) 101 (d) 110

Answer: (c)

Let the number of terms be n.

We know that sum of n terms of an A.P. = $\frac{n}{2}$(First term + Last term).

Thus, $\frac{n}{2}(-4+146) = 7171 \Rightarrow \frac{n}{2} \times 142 = 7171 \Rightarrow n = \frac{7171}{71} = 101$.

(xix) A.M. of two integral numbers exceeds their G.M by 2 and the ratio of the numbers is 1:4. Then the numbers are

(a) 5,20 (b) 1,4 (c) 2,8 (d) 4,16

Answer: (d)

Ratio of two numbers is 1:4. Let they are a and 4a.

Given that A.M. = G. M. + 2 $\Rightarrow \frac{5a}{2} = \sqrt{4a^2 + 2} \Rightarrow 5a = \pm 4a + 4 \Rightarrow a = 4$ (since the numbers are integer).

Hence the numbers are 4 and 16.

2. State whether the following statements are TRUE or FALSE:

(i) The ratio of two numbers is 2: 3. If 6 is subtracted from the second number then the number which is subtracted from the first number so that the new ratio becomes the same as that of the previous, is 4.

Answer: True.

Given that a : b = 2 : 3, $\Rightarrow$ 3a = 2b. Let x be the number subtracted from 'a'

So that (a – x): (b-6) = 2: 3, $\Rightarrow$ 3a – 3x = 2b – 12, $\Rightarrow$ 2b – 3x = 2b – 12, $\Rightarrow$ 3x = 12, $\Rightarrow$ x = 4.

(ii) If 0.5 of A = 0.6 of B = 0.75 of C and A+B+C = 60, then (A + 2), B and C will be in continued proportion.

Answer: False.

We have $\frac{A}{2} = \frac{3B}{5} = \frac{3C}{4} \Rightarrow \frac{A}{6} = \frac{B}{5} = \frac{C}{4} = \frac{A+B+C}{6+5+4}$ (Addendo). $\Rightarrow \frac{A}{6} = \frac{B}{5} = \frac{C}{4} = \frac{60}{15}$

Hence, A = 24, B = 20, C = 16. Let (24 + x), 20 and 16 are in continued proportion.

So, (24 + x).16 = 400, $\Rightarrow$ 24 + x = 25, $\Rightarrow$ x = 1.

(iii) If $\frac{\sqrt{a} - \sqrt{b}}{\sqrt{a} + \sqrt{b}} = \frac{1}{2}$ then a:b is 1:9.

Answer: False.

Given that $\frac{\sqrt{a} - \sqrt{b}}{\sqrt{a} + \sqrt{b}} = \frac{1}{2}, \rightarrow \frac{(\sqrt{a} + \sqrt{b}) + (\sqrt{a} - \sqrt{b})}{(\sqrt{a} + \sqrt{b}) - (\sqrt{a} - \sqrt{b})} = \frac{2+1}{2-1}$ **(Componendo- Divedendo)**

$\Rightarrow \frac{\sqrt{a}}{\sqrt{b}} = \frac{3}{1}$, $\Rightarrow$ a : b = 9 : 1.

(iv) The simple interest at X% for x years will be ₹X on a sum of ₹ $\frac{100}{x^2}$:

Answer: False

Let ₹P be the principal. As per sum x = $\frac{p.x.x}{100} \Rightarrow P = \frac{100}{x}$.

(v) ₹1600 earns ₹176 as compound interest for the second year at 10% per annum?

Answer: True.

Let ₹P be the principal. The amount after 1 year = P(1+ i), where i = $\frac{R}{100}$. Hence the compound interest for the second year = P(1 + i)i

So, P(1 + 0.1)0.1 = 176 $\Rightarrow$ P = 1600

(vi) A sum of money placed at compound interest doubles itself in 5 years. It will amount to eight times itself in 15 years.

Answer: True.

Let ₹P be the principal and it will amount to 8P in n yeas.

As per Sum 2P = P $(1 + i)^5 \Rightarrow \log 2 = 5 \log (1+i) \Rightarrow \log(1+i) = \frac{\log 2}{5}$,

and $8P = P(1+i)^n \Rightarrow \log 8 = n\log(1+i) \Rightarrow n = \frac{\log 8}{\log(1+i)} = 5\frac{\log 8}{\log 2}$ 5.3 = 15 Years.

(vii) The simple interest on a sum of money at the end of 6 years is $\frac{3}{10}$ *th*of the sum itself. The rate percent p.a. is 4%

Answer: False.

Let P be the sum of money. Here T = 6 years. $I = \frac{3}{10}P$. We have to find rate of interest per annum (R). We know $R = \frac{100I}{PT} = \frac{100 \times 3P}{10 \times p \times 6} = 5$.

(viii) A man borrows ₹1600 on the understanding that it is to be paid back in 6 equal installments at intervals of six months, the first payment to be made six months after the money was borrowed. The value of each installment, if the money is worth 6% p.a. is ₹295.

Answer: True.

Let P be the installment amount.

Here n = 3 years, p = no. of installment in a year = 2 and i = 0.06.

Total amount paid in a year = 2P.

Amount (A) = $\frac{2P}{0.06}\left\{\left(1+\frac{0.06}{2}\right)^{6}-1\right\}$.

Also ₹1600 amounted to 16000 $\left(1+\frac{0.06}{2}\right)^{6}$ in 3 years. (If the interest compounded half-yearly).

So, $\frac{2P}{0.06}\left\{\left(1+\frac{0.06}{2}\right)^{6}-1\right\} = 1600\left(1+\frac{0.06}{2}\right)^{6} \Rightarrow \frac{2P}{0.06}\times 0.194 = 1600\times 1.194$

$\Rightarrow P = \frac{1910.4\times 0.03}{0.194} = 295.4 \approx 295$

(ix) The ratio of work done by (x – 1) men in (x + 1) days to that of (x + 2) men in (x – 1) days is 9: 10, then value of x is 8.

Answer: True

As per the sum $\frac{(x-1)(x+1)}{(x+2)(x-1)} = \frac{9}{10} \Rightarrow 10(x+1) = 9(x+2) \Rightarrow x = 8.$

(x) Divide 69 into three parts which are in A. P and are such that the product of the 1st two parts is483, then the third part is 23.

Answer: False

Let 69 is divided into (a – d), a and (a + d).
As per sum, (a – d) + a + (a + d) = 69. => 3a = 69 => a = 23.
Also a(a -d) = 483 => a - d = 21
=> d = 2.
Hence the numbers are 21, 23, 25.

(xi) Whatever be the value of x, the terms x, x+3 and x+6 will be in A.P.

Answer: True

As (x + 3) – x = (x + 6) – (x + 3),
so whatever may be the value of x the terms x, (x + 3) and (x + 6) are in A.P.

(xii) Given c∝ (ax + b), value of c is 3 when a = 1, b = 2 and value of c is 5, when a = 2, b = 3. Value of x is 0.

Answer: False

As c ∝ (ax +b), so c = k (ax + b), k is a non zero constant.
For a = 1, b = 2, we get c = k (x + 2) ….. (i)
a = 2, b = 3, we get c = k (2x + 3) …… (ii)
Subtracting (i) from (ii), 0 = k (x + 1) or, x + 1 = 0 as $k \neq o$.Hence x = – 1.

(xiii) If $x^2 \alpha$ yz, $y^2 \alpha$ zx and $z^2 \alpha$ xy then the product of three constant of variation is 3.

Answer: False

As per sum, $x^2 \alpha$ Ayz
$y^2 \alpha$Bzx and
$z^2 \alpha$Cxy
$\Rightarrow x^2y^2z^2 = Ayz \times Bzx \times Cxy$
=>ABC = 1

(xiv) A man can finish a piece of work, working 8 hours a day in 5 days. If he works now 10 hours daily, then he can finish the same work in 6 days?

Answer: False.

Let A (= days) = 5, B (= hours) = 8, it is clear that $A \alpha \frac{1}{B} \Rightarrow A = \frac{k}{B}$, where k is a non-zero constant.

It is clear that A = 5 when B = 8 => k = 40

To find A, when B = 10, we have $A = \frac{40}{10} = 4$

(xv) If the nth term of a series be 2n+5, then the series is in A.P.

Answer: True

The n th term of a series is (2n + 5)

So, $T_n = (2n+5) \Rightarrow T_{n-1} = (2n+3)$ and $T_{n+1} = (2n+7)$.

Thus we have $2T_n = T_{n-1} + T_{n+1}$

Hence the series is an A.P.

Answer the following questions:

3. In a certain test, the number of successful candidates was three times than that of unsuccessful candidates. If there had been 16 fewer candidates and if 6 more would have been unsuccessful, the numbers would have been as 2 to 1. Find the number of candidates.

Answer: 136.

Let the number of successful candidates be 3s, then's' be the number of unsuccessful candidates. So, total number of candidates 4s.
Now as per sum {(4s – 16) – (s + 6)} : (s + 6) = 2 : 1, ⇒ (3s – 22) = 2(s + 6), ⇒ s = 34,
⇒ 4s = 136.

4. A certain product C is made by two ingredients A and B in the proportion of 2 : 7. The price of A is three times that of B. The overall cost of C is ₹ 5.90 per tonne including labour charges of 70 paise per tonne. Find the cost A and B per tonne.

Answer: ₹10.80, ₹3.60.

Let the price of A is ₹ 3t per tonne, then that of B is ₹ t per tonne.
As per sum 9 tonnes of C is made by 2 tonnes of A and 7 tonnes of B.
So, the overall cost of 1 tonne of C is
{₹ (3t × 2 + t × 7) / 9} + labour charges. ⇒ 13t = 9(5.90 – 0.70), ⇒ t = 3.60.
Hence, cost of A and B per tonne are respectively ₹10.80 and ₹3.60.

5. Rahul borrowed ₹830 from Mr. Lal at 12% p.a. S.I. for 3 years. He then added some more money to the borrowed sum and lent it to Shobha for the same period at 14% p.a. rate of interest. If Rahul gains ₹93.90 in the whole transaction, how much money did he add from his side ?

Answer: 105.

Let Rahul added ₹ P with the borrowed sum. We have $I = \frac{PRT}{100}$,

$\Rightarrow 93.90 = \frac{(830+P)\times 3\times 14}{100} - \frac{830\times 3\times 12}{100} \Rightarrow$ 830 × 3 × 2 + P x 3 x 14 = 9,390

$\Rightarrow 42P = 4410 \Rightarrow P = 105$

6. The difference between compound interest and simple interest on a sum for 2 years at 10% per annum, when the interest is compounded annually is ₹1600. If the interest were compounded half yearly, then what would be the difference in two interests?

Answer: 2481.

Let the principal be ₹ P. In case of simple interest, We have $I = \frac{PRT}{100}$,

Where T = 2 years and R = 10% . Hence $I = ₹\frac{P \times 2 \times 10}{100} = ₹\frac{P}{5} = ₹\ 0.20P$

When the interest Compounded annually,

$I = P\{(1+i)^n - 1\} = P\{(1+0.1)^2 - 1\} = 0.21P$. As per sum 0.01P = 1600,

P = 160000.

When the interest Compounded half - yearly,

$$I = P\left\{\left(1+\frac{i}{2}\right)^{2n} - 1\right\} = P\left\{(1+0.05)^4 - 1\right\} = 0.21550625P.$$

The difference of two interest = ₹0.01550625 × 160000 = ₹ 2481.

7. A man takes a loan of ₹25,000 at the rate of 8% S.I. with the understanding that it will be repaid with interest in 12 equal annual installments, at the end of every year. How much he is to pay in each installment?

Answer : ₹1388.89.

Here, Loan Amount (P) = ₹25000; Rate of interest (I) 8% p.a. and Time period is 12 years. So, total interest for that = $\frac{25000 \times 12 \times 8}{100}$ = ₹ 24000.

Let the amount with the above interest will repaid in 12 equal installment of ₹ P

So, $24000 = P(1+11\times.08) + P(1+10\times.08) + \ldots\ldots\ldots + P(1+1\times.08) + P$

$= 12P + P \times 0.08\,(11+10+\ldots\ldots\ldots\ldots+2+1+0)$

$= 12P + P \times 0.08 \times \frac{11}{2}(11+1) = 12P + 5.28P = 17.28P.$

$\Rightarrow P = 1388.89$

8. The wear and tear of a machine is taken each year to be one-tenth of the value at the beginning of the year for the first ten years and one-fifteenth each year for the next five years. Find its scrap value after 15 years. (Given that log (0.9) = -0.0458; log 14 = 1.1461 and log 15 = 1.1761.

Answer: 24.66%

Let the value of the machine is 100P and A be the value of the machine after 10 years. It is given that for the first ten years rate of wear & tear of the machine is 10%.

Then $A = 100P\,(1 - 0.1)^{10} = 100P \times (0.9)^{10}$

For the next five years rate of wear & tear of the machine is $\frac{100}{15}\%$ or $6\frac{2}{3}\%$.

Let B be the value of the machine after 15 Years.

Then $B = 100P \times (0.9)^{10}\left(1-\frac{1}{15}\right)^{5} = 100P \times (0.9)^{10}\left(\frac{14}{15}\right)^{5}$

$$\Rightarrow \log B = \log 100 + \log P + 10\log(0.9) + 5(\log 14 - \log 15)$$

$$\Rightarrow \log\left(\frac{B}{P}\right) = 2 - 0.458 - 0.15 = 1.392 \Rightarrow \frac{B}{P} = 24.66 \Rightarrow B = 24.66P$$

Scrap value after 15 years is 24.66%.

9. A person purchases a house worth ₹150000 on a hire purchase scheme. At the time of gaining possession he has to pay 30% of the cost of the house and the rest amount is to be paid in 15 equal annual instilment. If the compound interest is reckoned at 8% p.a. What should be the value of each installment?

Answer : ₹12267

Let ₹P be annual installment. Now, 70% of ₹150000 = ₹105000.

Here, n = 15 years and i = 0.08

Therefore, amount (A) $\frac{P}{i}\left\{(1+i)^{15} - 1\right\} = \frac{P}{0.08}\left\{(1.08)^{15} - 1\right\} = \frac{P}{0.08} \times 2.1722.$

Again ₹ 105000 is amounted to $105000(1+i)^{15} = 105000(1.08)^{15} = 105000 \times 3.1722 = 333081.$ in 15 years, if the interest compounded yearly.

Hence, $\frac{P}{0.08} \times 2.1722 = 333081, \Rightarrow P = \frac{333081 \times 0.08}{2.1722} = 12267.05 \approx 12267.$

Thus, value of each instilment is ₹12267.

10. Compute compound interest on ₹1500 for 3 years at 10% p.a.

Answer: ₹1815.

Here P = Principal = 1500, n = number of years = 3, i = interest on Rupee 1 for 1 year = 0.1. We know, amount (A) $= P(1+i)^{n} = 1500 \times (1.1)^{3} = 1500 \times 1.21 = 1815$.

11. If the population of a town increases every year by 25 per thousand, in how many years the population will be doubled? (Given: log 2 = 0.3010 and log 1.025 = 0.0107).

Answer : 28 years (nearly).

Let P be the present population of the town and after n years it will be 2P.

i = increase of population every year in respect of 1 person = 0.025.

As per sum, $2P = P(1+i)^n = P(1.025)^n, => \log\log 2 = n\log 1.025, => n = \frac{\log\log 2}{\log 1.025} \approx 28$

12. The difference between the compound interest and the simple interest on a sum put out for 3 years at 8% was ₹500. Find the sum.

Answer : ₹25365.

Let the principal be ₹ P. In case of simple interest, We have $I = \frac{PTR}{100}$,

Where T = 3 years and R = 8% . Hence I =₹ $\frac{P \times 3 \times 8}{100}$ ₹ 0.24P

When the interest Compounded annually,

$I = P\left\{(1+i)^n - 1\right\} = P\left\{(1+0.08)^3 - 1\right\} = 0.259712P.$

As per sum 0.019712P = 500, $=> P = 25365.26 \approx 25365$.

13. If $\frac{b+c}{a} = \frac{c+a}{b} = \frac{a+b}{c}$ and a + b + c ≠ 0 then show that each of these ratio is equal to 2.Also prove that $a^2+b^2+c^2 = ab+bc+ca$.

Answer : $=> a^2+b^2+c^2 = ab+bc+ca$.

Here $\frac{b+c}{a} = \frac{c+a}{b} = \frac{a+b}{c}$, Each of these ratio is equal to $\frac{(b+c)+(c+a)+(a+b)}{a+b+c}$ (By Addendo) i.e. 2.

Thus

$b+c = 2a, c+a = 2b, a+b = 2c$

$=> 4\left(a^2+b^2+c^2\right) = (b+c)^2 + (c+a)^2 + (a+b)^2 = 2\left(a^2+b^2+c^2+bc+ca+ab\right)$

$=> a^2+b^2+c^2 = ab+bc+ca.$

14. There has been increment in the wages of labourers in a factory in the ratio of 22:25, but there has also been a reduction in the number of labourers in the ratio of 15:11. Find out in what ratio the total wage bill of the factory would be increased or decreased.

Answer: 6:5

Let Presently there are 11x number of labourers and₹22y be the wages per labour. So, wages bill = ₹15x × 22y = 330xy.

After increment in the wages it becomes ₹25y per head and number of labour reduces to 11x. In this case, total wages bill = 11x × 25y = 275xy.
Hence, the total wages bill of the factory decrease in the ratio 6:5.

15. If 5^{th} and 12^{th} terms of an A.P are 14 and 35 respectively find the A.P

Answer : 2, 5, 8, 11, 14, 17, ……

Let the first term and common difference of the A.P. are a and d respectively.
Thus, a + 4d = 14 and a + 11d = 35 =>7d = 21 => d = 3.
So, a = 14 – 4d => a = 2.
Hence the A.P. is 2, 5, 8, 11, 14, 17, ...

16. Insert 4 arithmetic means between 4 and 324.

Answer: 68, 132, 196, 260.

If we insert 4 Arithmetic Mean (A.M.) between 4 and 324 then, the 1^{st} and 6^{th} terms of the A.P. are respectively 4 and 324. Let d be the common difference.
Then 4 + 5d = 324 => 5d = 320 => d = 64.
Hence the A.M.'s are 68, 132, 196 and 260.

17. Insert 3 geometric means between 1/9 and 9

Answer: $\frac{1}{3}, 1 \text{ and } 3$

If we insert 3 Geometric Mean (G.M.) between $\frac{1}{9}$ and 9 then, the 1^{st} and 5^{th} terms of the G.P. are respectively $\frac{1}{9}$ and 9. Let r be the common ratio.

Then $\frac{1}{9} \times r^4 = 9 \Rightarrow r = 3$.

Hence the G.M.'s are $\frac{1}{3}$, 1 and 3.

18. X is proportional directly to Y and inversely with Z. X = $\frac{1}{6}$, when Y = 5 and Z = 9. Find the relation among X, Y and Z.

Answer: $X = \frac{3Y}{10Z}$

As per sum, $X \alpha Y$ and $X \alpha \frac{1}{2} \Rightarrow X \alpha \frac{Y}{Z} \Rightarrow X = k\frac{Y}{Z}$.

It is given that, $X = \frac{1}{6}, when Y = 5 and Z = 9. \Rightarrow \frac{1}{6} = k\frac{5}{9} \Rightarrow k = \frac{3}{10} \Rightarrow X = \frac{3Y}{10Z}$.

19. The expenses of a boarding house are partly fixed and partly varies with the number of boarders. The charge is ₹70 per head when there are 25 boarders and ₹60 per head when there are 50 boarders. Find the charge per head when there are 100 boarders.

Answer : ₹ 55.

Let x = fixed monthly expense, y = variable expense, n = no. of boarders.
Now y ∝ n or, y = k n, k is non-zero constant. The monthly expenses for 25 and 50 boarders are respectively ₹ 1,750 and ₹ 3,000.
Now total expense = fixed expenses + variable expenses.
i.e., T = x + y = x + kn, where T = total expenses.
So, from T = x + kn, we get, 1,750 = x + 25k ….. (1)
and 3,000 = x + 50k……. (2)
Subtracting (1) from (2), 25k = 1,250, or k = 50.
Again, putting the value of k in (1), we find x = ₹500.
Now, charge for 100 boarders = x + 100 k = 500 + 100 x 50 = ₹5,500
So, Charge per head = 5,500/100 = ₹ 55.

20. Find three numbers in GP. Whose sum is 19& Product is 216?

Answer: 4, 6, 9

Let three numbers in G.P. are $\frac{a}{r}, a$ and ar.

It is given that their product is 216 $\Rightarrow a^3 = 216 \Rightarrow a = 6$

Also their sum is 19. $\Rightarrow \frac{6}{r} + 6 + 6r = 19 \Rightarrow 6r^2 - 13r + 6 = 0$

$\Rightarrow (2r - 3)(3r - 2) = 0 \Rightarrow r = \frac{3}{2} or \frac{2}{3}$

When $r = \frac{3}{2}$, numbers are 4, 6, 9 and when $r = \frac{2}{3}$, numbers are 9, 6, 4.

Hence the numbers are 4, 6, 9.

21. Three numbers are in A.P and their sum is 15. If 1,3,9 are added to them respectively, they form a G.P. Find the numbers.

Answer : 3, 5, 7 OR 15, 5, -5.

Let three numbers in A.P. are (a – d), a, (a +d).

Also their sum is 15. $\Rightarrow 3a = 15 \Rightarrow a = 5$

If we add 1, 3, 9 respectively with the above three then they are (6 – d), 8 and (14 + d).

As per sum, $(14+d)(6-d) = 64 \Rightarrow d^2 + 8d - 20 = 0 \Rightarrow (d+10)(d-2) = 0.$

$\Rightarrow$ d=2, -10.

When d = 2, numbers are 3, 5, 7 and when d = -10, numbers are 15, 5, -5.

Study Note – 2

Algebra

Learning Objective:

After studying this chapter, you should be able to:

- ***Determine whether an object is an element of a set.***
- ***Determine whether a set is finite or infinite.***
- ***Determine whether two sets are equal, equivalent, or neither.***
- ***Recognise technical terms and appreciate some of the uses of algebra***
- ***To gather knowledge about Indices and Logarithms***
- ***Learn about Permutations and Combinations***
- ***Solve simple quadratic equations.***

1. Choose the Correct Answer:

(i) The set A={1,2,3} and the set B={1,2}, then A-B is
(a).{0} (b).{2} (c).{3} (d).Φ

Answer: (c)

Difference of two sets A and B i.e. A-B contain only those element of A which are not in B.

(ii) A={x : x∈ N and x^2 <36}, where N is the set of natural number.
(a).{1,2,2,……,35} (b).{1,2,3,4,5} (c).A (d).N

Answer: (b)

Square of 1,2,3,4 and 5 are less than 36.

(iii) If a set V contain 4 elements the power set of V i.e. P(V) contain
(a).4 elements (b).0 element (c).16 elements (d).32 elements

Answer: (c)

If a set contain n element then its power set contain 2^n element.
Since V contain 4 elements so P(V) contain 2^4=16 elements.

(iv) Let A={1,2,3,4,8} B={2,4,6,7}, then A∆B is equal to
(a).{1,3,6,7,8} (b).{1,3,8} (c).{6,7} (d).{1,2,3,4,6,7,8}

Answer: (a)

AΔB=(A-B)∪(B-A)
A-B={1,3,8} and B-A={6,7} therefore AΔB={1,3,6,7,8}

(v) If $p\sqrt{p} = \sqrt[p]{p}$, then p is equal to
(a).$\frac{2}{3}$ (b).$\frac{3}{2}$ (c).2 (d).3

Answer: (a)

Given that, $p\sqrt{p} = \sqrt[p]{p}$
=> $p^{1+\frac{1}{2}} = p^{\frac{1}{p}}$ **=>** $\frac{3}{2} = \frac{1}{p}$ **=>** $p = \frac{2}{3}$

(vi) $3^{x-1} - 3^{x-3} = 8$ Then the value of $x^2 - x$+4 is
(a).10 (b).8 (c).12 (d).16

Answer: (a)

Given that, $3^{x-1} - 3^{x-3} = 8$
$\Rightarrow 3^x 3^{-1} - 3^x 3^{-3} = 8 \Rightarrow 3^x(3^{-1} - 3^{-3}) = 8$
$\Rightarrow 3^x\left(\frac{1}{3} - \frac{1}{27}\right) = 8 \Rightarrow 3^x = 8 \times \frac{27}{8} \Rightarrow 3^x = 3^3 \Rightarrow x = 3$
Putting x=3 in x^2-x+4
3^2-3+4=10

(vii) ($\log_b a$. $\log_c b$. $\log_a c$) is equal to
(a).0 (b).1 (c).abc (d).(a+b+c)

Answer: (b)

We know that $\log_b a$. $\log_c b$. $\log_a c = \frac{log_{10} a}{log_{10} b} \cdot \frac{log_{10} b}{log_{10} c} \cdot \frac{log_{10} c}{log_{10} a} = 1$

(viii) The value of $\log_e 1$ is
(a).e (b).1 (c).0 (d).10

Answer: (c)

let $log_e 1$=N
=> $e^N = 1 \Rightarrow e^N = e^0 \Rightarrow N = 0 \Rightarrow log_e 1 = 0$

(ix) If $a = b^2 = c^3$, then the value of $\log_a(abc)$ is

(a).$\frac{13}{3}$ (b).$\frac{11}{6}$ (c).$\frac{8}{3}$ (d).$\frac{b}{c}$

Answer: (b)

By the property of logarithm, we have $log_a abc = log_a a.a^{\frac{1}{2}}.a^{\frac{1}{3}}$

$= log_a a^{1+\frac{1}{2}+\frac{1}{3}} = log_a a^{\frac{11}{6}} = \frac{11}{6} log_a a = \frac{11}{6}$.

(x) If $n_{P_3} = 120$ *then* n is

(a).8 (b).4 (c).6 (d).0

Answer: (c)

We have, $n_{P_3} = 120$

$\Rightarrow \frac{n!}{(n-3)!} = 120 \Rightarrow n(n-1)(n-2) = 120$

$=>n(n-1)(n-2) = 6.5.4$ $(product\ of\ three\ consucative\ positive\ integer)$

⇒n=6

(The value of n can also be find by solving cubic equation in n)

(xi) If $^nP_4$=30 x $^nP_2$, then the value of n is

(a).10 (b).8 (c).6 (d).5

Answer: (b)

Given, $^nP_4 = 30 \times {}^nP_2$

$\Rightarrow \frac{n!}{(n-4)!} = 30.\frac{n!}{(n-2)!} \Rightarrow (n-2)(n-3) = 30$

$=>n^2 - 5n - 24 = 0 \Rightarrow (n-8)(n+3) = 0 \Rightarrow n = 8, -3$

Neglecting negative value n=8

(xii) Numbers of permutation made out of letters of the words TRIANGLE

(a).8 (b).40320 (c).16 (d).40340

Answer: (b)

The word TRIANGLE contain non repeated 8 letters so the number of permutation is $^8P_8 = 8! = 40320$

(xiii) Everybody in a room shakes hands with everybody else. The total number of hands shakes is 66. The total number of person in the room is

(a).11 (b).12 (c).10 (d).14

Answer: (b)

Let total number of person be n
Number of hand shake is ${}^{n}C_2 = 66 \Rightarrow n(n-1) = 132$
$\Rightarrow (n-12)(n+11) = 0$ neglecting negative value n = 12

(xiv) If $x^2 - 5x + 6 = 0$, then the roots of the equation are -
(a).3 and 3 (b).2 and 3 (c).2 and 2 (d).1 and 2

Answer: (b)

Given equation is, $x^2 - 5x + 6 = 0$
$\Rightarrow (x-2)(x-3) = 0 \quad \Rightarrow x = 2, 3$

(xv) If $2 - \sqrt{3}$ be one root of a quadratic equation then other root is -
(a).2 (b).$\sqrt{3}$ (c).$2 + \sqrt{3}$ (d).2+i$\sqrt{3}$

Answer: (c)

Irrational (surd) roots occur in conjugate pair.

(xvi) If $b^2-4ac > 0$ then the roots of the equation $ax^2 + bx + c = 0$
(a).Real and equal (b).Real and unequal (c).Imaginary (d).Imaginary and equal

Answer: (b)

Since the discriminate (b^2-4ac) is greater than 0, So roots are real and unequal.

(xvii) If one root of the equation $x^2 - 3x + m = 0$ exceeds the other by 5 then the value of m equal to -
(a). – 6 (b). – 4 (c).12 (d).18

Answer: (b)

Let the roots be **a** and **a+5**
Then sum of the roots = a + a + 5 = - 3 or a = - 4
Product of the roots = a(a + 5) = m => - 4.(- 4 + 5) = m => **m= - 4**

(xviii) If c = 0 then one of the root of the equation $ax^2 + bx + c = 0$ is –
(a).a (b).b (c).a+b (d).0

Answer: (d)

If c = 0 then $ax^2 + bx = 0$
$\Rightarrow x(ax+b)=0 \quad \Rightarrow$ **x = 0** and –b /a

2. State whether the following statements are TRUE or FALSE:

(i) {ϕ} is a null set.

Answer: False

{ϕ} contain one element ϕ, so it is not a null set.

(ii) Null set is a superset of every set.

Answer: False

Null set is subset of all sets.

(iii) If a set X has 10 elements and the set Y has 9 elements, then the number of element in the set X∩Y is 9, when X⊃Y.

Answer: True

Here X⊃Y, hence X∩Y=Y, so the number of element in the set X∩Y is 9.

(iv) Set of integer Z is the subset of Natural number N.

Answer: False

Set of natural number N = {1,2,3,4,……..},
Set of integer Z={…….-4,-3,-2,-1,0,1,2,3,4,…….}
Hence, N is the subset of Z

(v) The value of $(32)^{\frac{1}{5}}$ is $\frac{1}{2}$.

Answer: False

We know, $(32)^{\frac{1}{5}} = (2^5)^{\frac{1}{5}} = 2.$

(vi) Logarithm of 1 with any base is zero.

Answer: True

let $log_X 1$=N (where X is any base)
=>$X^N = 1 \Rightarrow X^N = X^0 \Rightarrow N = 0 \Rightarrow log_X 1 = 0$

(vii) The logarithm of same number for different base is different.

Answer: True

The statement is True. As for example, $64=4^3$ => **$\log_4 64=3$** and $64=8^2$ => **$\log_8 64=2$**

(viii) The logarithms with base e are called common logarithm.

Answer: False

The logarithms with base e are called natural logarithm.

(ix) Logarithm of a number (≠1) with itself as a base is unity.

Answer: True

Let a ≠1. Now, $\log_a a = \log_e a / \log_e a = 1$.

(x) The value of 5! is equal to 1020.

Answer: False

It is clear that, 5!=5x4x3x2x1=120

(xi) The total number of arrangements of the letters in the expression $x^3y^2z^4$ when written in full length is 1260.

Answer: True

We know that, $x^3 y^2 z^4$ = x x x y y z z z z
The total number of arrangements=$\frac{9!}{3!2!4!} = \frac{9.8.7.6.5.4!}{3.2.1.2.1.4!} = 1260$

(xii) The value of $^5C_2$ is equal to $^5C_3$

Answer: True

We have, $^5C_2 = \frac{5!}{3!\times(5-3)!} = \frac{5!}{3!\times 2!} = 10 = {}^5C_3$

(xii) $^nC_r = {}^nC_{n-r}$ (n≥r)

Answer: True

It is clear that, $^nC_{n-r} = \frac{n!}{(n-n+r)!(n-r)!} = \frac{n!}{(n-r)!r!} = {}^nC_r$

(xiv) If ${}^nP_1 = {}^nC_1$, then ${}^nP_3 = {}^nC_3$

Answer: False

${}^nP_3 = \frac{n!}{(n-3)!}$ *but* ${}^nC_3 = \frac{n!}{(n-3)!3!}$. Hence, ${}^nP_3 \neq {}^nC_3$

(xv) If the roots of the equations $ax^2+bx+c= 0$ are equal the then $b^2=4ac$.

Answer: True

Roots of the equation $ax^2 + bx + c = 0$ are $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$. when roots are equal then $\sqrt{b^2-4ac}=0$ i.e $b^2 = 4ac$.

(xvi) If one roots of a quadratic equation is a + ib the other root $a - \sqrt{b}$.

Answer: False

Imaginary roots occur in conjugate pair therefore other root is a - ib.

(xvii) If 2 and 3 be the roots of an equation then the equation can be written as (y + 2)(y + 3) = 0

Answer: False

If 2 and 3 be the roots of an equation then the equation can be written as (y-2)(y-3)=0.

(xvii) If in the equation $ax^2 + bx + c = 0$ coefficient of x vanishes then the roots are of equal in magnitude and opposite in sign.

Answer: True

If coefficient of x vanishes then $ax^2 + 0.x + c = 0$

$\Rightarrow ax^2+c= 0 \Rightarrow x = \pm\sqrt{\frac{-c}{a}}$

Hence the roots are of equal in magnitude and opposite in sign.

(xviii) The roots of the equation $(x-4)^2(x-2)(x+2)=0$ are 4, 4, 2, - 2

Answer: True

Given that, $(x - 4)^2(x - 2)(x + 2) = 0$

$\Rightarrow (x - 4)(x - 4)(x - 2)(x + 2) = 0$

$\Rightarrow x=4, 4, 2, - 2.$

3. Answer the following questions:

(i) If A and B are two sets such that n(A)=27, n(B)=35 and n(A∪B)=50, find n(A∩B).

Answer: 12

We know, n(A∪B)=n(A)+n(B)- n(A∩B)
=> n(A∩B)= n(A)+n(B)- n(A∪B)=27+35-50=12

(ii) In a group of 850 people, 600 can speak Bengali and 340 can speak Nepali. Find How many can speak Nepali only.

Answer: 200

Let A be the set of people speak Bengali and B be the set of people speak Nepali
So, n(A)=600, n(B)=340 and n(A∪B)=800
=>n(A∩B)= n(A)+n(B)- n(A∪B)=600+340-800=140
Number of people speak both Bengali and Nepali i.e. n(A∩B)=140
Hence, Number of people speak Nepali only =n(B)- n(A∩B)=340-140=200

(iii) If A={x : x∈ N, x≤7}, B={x:x is prime, x<8} and C={x: x∈ N , x is odd and x<10} verify that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Answer:

A={1,2,3,4,5,6,7}, B={2,3,5,7} and C={1,3,5,7,9}
$(A \cup B)$={1, 2, 3, 4, 5, 6, 7} = A
$(A \cup C)$={1, 2, 3, 4, 5, 6, 7, 9} and $B \cap C$={3, 5, 7}
Now $A \cup (B \cap C)$ ={1, 2, 3, 4, 5, 6, 7} = A
and $(A \cup B) \cap (A \cup C)$={1, 2, 3, 4, 5, 6, 7}=A
Hence, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(iv) If A and B be two sets containing 3 and 6 elements respectively, what can be the maximum number of element in $A \cup B$? Under what condition the maximum numbers of elements in $A \cup B$ occur?

Answer: 9, $A \cap B = \phi$

Since the union of two sets contain element of both the sets, taking common element only once, so the maximum number of element in $A \cup B$ is 9.
Maximum number of element occur in $A \cup B$, when two sets has no common elements, i,e. $A \cap B = \phi$.

(v) $2^X + \frac{1}{2^X} = \frac{5}{2}$, **find the value of x.**

Answer: ± 1

Given, $2^X + \frac{1}{2^X} = \frac{5}{2}$

$\Rightarrow \frac{2^{2X}+1}{2^X} = \frac{5}{2} \Rightarrow 2 \times 2^{2X} + 2 = 5 \times 2^X$, let 2^x=y

Then, $2y^2 - 5y + 2 = 0 \Rightarrow (y-2)(2y-1) = 0 \Rightarrow y = 2, \frac{1}{2}$

$=>2^X = 2 \Rightarrow 2^X = 2^1 \Rightarrow X = 1 \; and \; 2^X = \frac{1}{2} \Rightarrow 2^X = 2^{-1} \Rightarrow X = -1$

Hence, X=±1

(vi) **Show that** $log_3\sqrt{3\sqrt{3\sqrt{3} \dots \dots \infty}} = 1$

Answer:

$Let\ x = \sqrt{3\sqrt{3\sqrt{3} \dots \dots \infty}} \Rightarrow log_3 x = log_3\sqrt{3\sqrt{3\sqrt{3} \dots \dots \infty}}$

$\Rightarrow log_3 x = \frac{1}{2} log_3 3\sqrt{3\sqrt{3\sqrt{3} \dots \dots \infty}} \Rightarrow log_3 x = \frac{1}{2} log_3 3x$

=>2$log_3 x = log_3 3 + log_3 x \Rightarrow log_3 x = log_3 3 \Rightarrow log_3 x = 1$

$\Rightarrow log_3\sqrt{3\sqrt{3\sqrt{3} \dots \dots \infty}} = 1.$

(vii) **Evaluate $\log_2\log_2(\log_2 4)$**

Answer: 0

We have, $\log_2\log_2(\log_2 4) = \log_2\log_2\log_2 2^2 = \log_2\log_2 2 \log_2 2 = \log_2\log_2 2.1 = \log_2 1 = 0$

(viii) **Find the value of** $log 324$ **to the base** $3\sqrt{2}$**.**

Answer: 4.

We know that $324 = 4 \times 81 = 2^2 \times 3^4$

$\log_{3\sqrt{2}} 324 = \log_{3\sqrt{2}} 2^2.3^4 = \log_{3\sqrt{2}} (3\sqrt{2})^4 = 4\log_{3\sqrt{2}} (3\sqrt{2}) = 4.1 = 4$

(ix) **Prove that $\log_a y.\log_x a = \log_x y$**

Answer:

Let $\log_a y = p$, $\log_x a = q$ and $\log_x y = r$ => $y = a^p$, $a = x^q$ and $y = x^r$

Now, $y = a^p = (x^q)^p = x^{pq}$ => $pq = \log_x y$ => $\log_a y.\log_x a = \log_x y$.

(x) Three persons enter a railway carriage, when there are 5 vacant seats. In how many ways can they seat themselves?

Answer: 60

Clearly the first person can occupy any of the 5 seats in 5 ways. Now, the second person can occupy any of the remaining 4 seats in 4 ways and finally the last person can occupy remaining 3 seats in 3 ways.
Hence by the fundamental principal of counting, the required number of ways =5 x 4 x 3 = 60

(xi) How many words can be formed from the letter of the word DAUGHTER so that the vowels always come together?

Answer: 4320

The given words contain 8 different letters.
When the vowels A, U, E are always together, they can be treated as one letter,
So we have to arrange D, G, H, T, R, (AUE). These 6 letters can be arranged in ${}^6P_6 = 6! = 720$ ways.
And the letters A U E can be arrange in ${}^3P_3 = 3! = 6$ ways.
Hence, the required number of words = 720 X 6 = 4320.

(xii) In how many ways can a committee of 5 members be selected from 6 men and 5 ladies, consisting of 3 men and 2 ladies?

Answer: 200 ways.

3 men out of 6 can be selected in ${}^6C_3$ ways.
2 ladies out of 5 can be selected in ${}^5C_2$ ways.
Hence, a five members committee can form in ${}^6C_3 \times {}^5C_2 = 200$ ways consisting of 3 men and 2 ladies.

(xiii) In how many ways can a boy invite one or more of 5 friends?

Answer: 31 ways.

The boy can invite 1 friend, 2 friends, 3 friends, 4 friends or all of his friends.
He can invite n friends out of 5 in ${}^5C_n$ ways.
Hence, he can invite his friends in ${}^5C_1 + {}^5C_2 + {}^5C_3 + {}^5C_4 + {}^5C_5 = 31$ ways.

(xiv) Form the equation whose roots are (1 ± i).

Answer: $x^2 - 2x + 2 = 0$.

Sum of the roots=1 + i + 1 – i = 2
Product of the roots = (1 + i)(1 - i) = $1^2 - i^2 = 1 + 1 = 2$
The required equation is $x^2 - 2x + 2 = 0$

(xv) Solve $py^2 + qy + r = 0$

Answer: $\frac{-q \pm \sqrt{q^2-4pr}}{2p}$.

Given equation is $py^2 + qy + r = 0$
$\Rightarrow 4p^2y^2 + 4pqy + 4pr = 0$ (multiplying both side by 4p)
$\Rightarrow (2py)^2 + 2.2py.q + q2 = q^2 - 4pr$ (adding both side q^2)
$\Rightarrow (2py + q)^2 = q^2 - 4pr$
$\Rightarrow 2py + q = \pm\sqrt{q^2 - 4pr}$
$\Rightarrow y = \frac{-q \pm \sqrt{q^2-4pr}}{2p}$

(xvi) Find the least positive value of m for which the equation $x^2 + mx + 4 = 0$ has real roots.

Answer: 4.

For real roots discriminate greater than or equal to zero.
i.e., $m^2 - 4.1.4 \geq 0 \Rightarrow m^2 \geq 16 \Rightarrow m \geq 4$.
Hence, 4 is the least positive value of m.

(xvii) Solve: $2\left(x + \frac{1}{x}\right)^2 - 7\left(x + \frac{1}{x}\right) + 5 = 0$.

Answer: $2, \frac{1}{2}, \frac{1 \pm i\sqrt{3}}{2}$.

Let $\left(x + \frac{1}{x}\right) = y$, then the given equation becomes $2y^2 - 7y + 5 = 0$
$\Rightarrow$ (y - 1)(2y - 5) = 0
$\Rightarrow y = 1 \; and \; \frac{5}{2}$
When y=1, $\left(x + \frac{1}{x}\right) = 1$ **=>**x^2-x+1=0 **=>**$x = \frac{1 \pm \sqrt{1-4}}{2} = \frac{1 \pm i\sqrt{3}}{2}$
When $y = \frac{5}{2}$, $x + \frac{1}{x} = \frac{5}{2}$ $\Rightarrow 2x^2 - 5x + 2 = 0$
$\Rightarrow$ (x - 2)(2x - 1) = 0
$\Rightarrow x = 2 \; and \; \frac{1}{2}$
Therefore the required solution is $2, \frac{1}{2}, \frac{1 \pm i\sqrt{3}}{2}$.

(xviii) Solve: $2^{2x} - 3.2^{x+2} + 32 = 0$.

Answer: $2, \frac{1}{2}, \frac{1 \pm i\sqrt{3}}{2}$.

Let $2^x = y$, then the given equation becomes $y^2 - 3.y.2^2 + 32 = 0$

$\Rightarrow y^2 - 12y + 32 = 0 \quad \Rightarrow (y - 8)(y - 4) = 0 \quad \Rightarrow y - 4 = 0$

$\Rightarrow 2^x = 2^2 \quad \Rightarrow \mathbf{x=2}$

And $y - 4 = 0 \quad \Rightarrow 2^x - 8 = 0 \quad \Rightarrow 2^x = 2^3 \quad \Rightarrow \mathbf{x=3}$

Hence, x = 2 and 3.

SECTION – B : STATISTICS

Study Note – 3

Statistical Representation of Data

Learning Objective:

After studying this Chapter, the student will be able to:

- ***Define data and range.***
- ***Examine a real-world problem in which the range of a set of whole numbers is found.***
- ***Examine a real-world problem in which the range of a set of decimals is found.***
- ***Examine a real-world problem in which the range of a set of integers is found.***
- ***Compute the range of a set of numbers.***
- ***Describe the procedure for computing the range of a set of numbers.***

I. Multiple Choice Questions: (In each question choose the appropriate option out of four options given)

1. Data are

(a) Collections of any number of observations
(b) Collections of any number of related observations
(c) Compilation of any number of observations
(d) Arrangement of information of related observations

2. Data point is

(a) An unrelated observation
(b) An unrelated information
(c) A point on a graph
(d) A single observation

3. A population is

(a) A collection of all the elements which are to be studied & about which attempts are to be made to draw conclusion
(b) A collection of all the elements on which conclusions are already made & further studies are to be made
(c) A collection of all the elements on which studies are already done & about which attempts are to be made to draw conclusion
(d) Combination of samples on which studies and conclusions are all made after carrying out hypothesis test

4. We study data in a sample in order to be able to
   (a) Describe the coherence of the data
   (b) Study the population
   (c) Present sample in a statistical manner
   (d) Do statistical analysis of few data

5. A representative sample contains
   (a) The relevant characteristics of the population as they are included in the population
   (b) The relevant characteristics of the population in the same proportion as they are included in the population
   (c) The relevant characteristics of the standard sample constructed on the basis of standard set by population
   (d) The relevant characteristics of the population which is made on the basis of some standard already checked.

6. Method of organizing data into similar classes and counting the number of observations produces
   (a) Frequency Polygon
   (b) Frequency Chart
   (c) Histogram
   (d) Frequency Distribution

7. Information before it is arranged and analyzed is called
   (a) Raw data
   (b) Raw sample
   (c) Primary sample
   (d) Raw population

8. Data array arranges data
   (a) In ascending order
   (b) In descending order
   (c) In ascending or descending order
   (d) In the order required by the study

9. Which one of the following is not an advantage of a data array?
   (a) Data array helps to identify easily lowest and highest values in the data
   (b) Data array facilitates to divide the data into sections
   (c) Data array helps to identify any value appearing more than once in the data
   (d) Data array presents large quantity of data in a compressed way

10. A frequency distribution is a table
(a) That organizes data into groups of values describing one characteristic of the data
(b) That organizes data into groups of values in descending order of the data
(c) That organizes data into groups of values in ascending order of the data
(d) That organizes data into groups of values describing characteristics of the population

11. A relative frequency distribution pairs
(a) Each class with its total number of data points that falls within that class
(b) Each class with a fraction or percentage of the total data
(c) Each class with its appropriate fraction or percentage of the total data
(d) Each class with equal fraction or percentage of the total data

12. A relative frequency table is given below:

Class	Relative Frequency
2.0 to 2.5	0.05
2.6 to 3.1	0.00
3.2 to 3.7	0.10
3.8 to 4.3	0.40
4.4 to 4.9	0.25
5.0 to 5.5	0.20
	1.00

Which one of the following is not an appropriate statement in relation to the above table?
(a) The classes are all inclusive
(b) All the data fit into one category or another
(c) Classes are not mutually exclusive
(d) Classes are not overlapping

13. A frequency distribution is given below:

Class (Age)	Frequency
Birth to 7	8873
8 to 15	9246
16 to 23	12060
24 to 31	11949
32 to 39	9853
40 to 47	8439
48 to 55	8267
56 to 63	7430
64 to 71	7283
72 and older	6192

With respect to the above distribution

(a) Birth to 7 is the open ended class

(b) 72 and older is the open ended class

(c) Birth to 7 & 72 and older both are open ended classes

(d) All Classes are closed

14. Discrete classes are

(a) Entities that progress from one class to the next without a break

(b) Entities that progress from one class to the next open ended class

(c) Entities that progress from one open ended class to the next closed class

(d) Entities that do not progress from one class to the next without a break

15. Which one of the following is discrete data?

(a) High school Grades

(b) Pound of pressure in pressure cooker

(c) No of children in each family

(d) Height of a person as per age

16. In construction of a frequency distribution number of classes depends on

(a) The number of data points

(b) The range of the data collected

(c) The number of data points and the range of the data collected

(d) The number of data points and the variability of the data collected

17. Width of class intervals is

(a) $\frac{\text{Next unit value after largest value in data - Smallest value in data}}{\text{Total number of class intervals}}$

(b) $\frac{\text{Smallest value in data + 0.5}}{\text{Total number of items}}$

(c) $\frac{\text{Next unit value after largest value in data - 0.1}}{\text{Total number of class intervals}}$

(d) $\frac{\text{Next unit value after largest value in data + Smallest value in data}}{\text{Total number of class intervals}}$

18. Following is data on chlorine levels in ppm of 30 gallons of treated water

16.2	15.8	15.8	15.8	16.3	15.6
15.7	16.0	16.2	16.1	16.8	16.0
16.4	15.2	15.9	15.9	15.9	16.8
15.4	15.7	15.9	16.0	16.3	16.0
16.4	16.6	15.6	15.6	16.9	16.3

To construct a frequency distribution on above with 6 classes, the width of a class is

(a) 0.6

(b) 0.3

(c) 0.2

(d) 0.5

19. An all-inclusive and mutually exclusive classes in a frequency distribution means

(a) Every data point fits into at least one class & smallest data corresponds to the first class

(b)Every data point fits into at least one class & No data point fits into more than one class

(c) No data point fits into more than one class & largest data corresponds to the last class

(d) Smallest data corresponds to the first class & largest data corresponds to the last class

20. A frequency distribution has first two classes as 15.2 to 15.4 & 15.5 to 15.7. Real limit of the first class is

(a) 15.3 to 15.6

(b) 15.05 to 15.55

(c) 15.25 to 15.45

(d) 15.15 to 15.45

21. In a frequency distribution representing car sales of a region in millions, the first class is 0-5. 5 is the

(a) Real upper limit of the class

(b) Stated upper limit of the class

(c) Real and Stated upper limit of the class

(d) Real upper interval of the class

22. In a frequency distribution Class marks describe

(a) The mid points of the classes

(b) The stated class boundaries of the classes

(c) The real class boundaries of the classes

(d) The class intervals

23. Class mark of a class for continuous variable is

(a) $\frac{\text{Stated lower limit of the class + Stated lower limit of the next class}}{2}$

(b) $\frac{\text{Stated lower limit of the class + Stated upper limit of the class}}{2}$

(c) $\frac{\text{Stated lower limit of the class + Stated lower limit of the next class}}{\text{Total no of Items}}$

(d) $\frac{\text{Stated lower limit of the class + Stated lower limit of the next class}}{\text{Total no of Items}}$

24. Class mark of a class for discrete variable is

(a) $\frac{\text{Stated lower limit of the class + Stated lower limit of the next class}}{2}$

(b) $\frac{\text{Stated lower limit of the class + Stated upper limit of the class}}{2}$

(c) $\frac{\text{Stated lower limit of the class + Stated lower limit of the next class}}{\text{Total no of Items}}$

(d) $\frac{\text{Stated lower limit of the class + Stated lower limit of the next class}}{\text{Total no of Items}}$

25. Which one of the following is not a characteristic of Histogram?

(a) Rectangles in histogram are proportional in width to the range of values within a class

(b) Rectangles in histogram are proportional in height to the number of items falling in the class

(c) Rectangles in histogram are of equal width if the classes in distribution are of equal width

(d) Area of a Rectangle in histogram in percentage of area of all the rectangles is greater than relative frequency of that class

26. Relative frequency histogram is

(a) That uses relative width of data points in each of the classes rather than actual width

(b) That uses relative frequency of data points in each of the classes rather than actual number of points

(c) That uses relative frequency of data points in each of the classes rather than total number of points

(d) That uses relative width of data points in each of the classes rather than total width of all the classes

27. A frequency polygon is

(a) Line graph that connects the mid points of all the bars in a histogram

(b) Bar chart that is made with class marks of all the classes in a distribution

(c) Line graph that connects the class boundaries of all the classes in a distribution

(d) Line graph that connects the class marks of all the continuous classes in a distribution

28. Which one of the following is not an advantage of a Histogram?

(a) The rectangle clearly shows each separate class in the distribution

(b) The area of each rectangle shows the proportion of the total number of observations that occur in that class

(c) Rectangles in histogram are proportional in width to the range of values within a class

(d) The histogram becomes increasingly smooth as we increase the number of classes and number of observations

29. In a frequency distribution with addition of more classes and data points
(a) Frequency polygon becomes Histogram
(b) Frequency polygon becomes Frequency Curve
(c) Histogram becomes frequency curve
(d) Frequency curve becomes frequency polygon

30. A graph of a cumulative frequency distribution is
(a) Frequency Polygon
(b) Ogive
(c) Frequency curve
(d) Cumulative Ogive

31. More than type Ogive
(a) Slopes up to the right
(b) Slopes down to left
(c) Slopes down to right
(d) Slopes up to the left

32. Relative frequency distribution is
(a) The display of a data set that shows the percentage of the total variability of data that is due to different classes
(b) The display of a data set that shows the percentage of the total data set that falls into each of a set of mutually exclusive classes
(c) The display of a data set that shows the percentage of the total deviation of the data from data that falls into each class
(d) the display of a data set that shows the percentage of width of each class in comparison to total range of the data

33. First two classes of three frequency distribution is given:

Frequency Distribution	1st class	2nd class
A	50-60	60-70
B	50-59	60-69
C	Below 60	60-70

Which of the following is correct?
(a) A is exclusive type, B open ended, C inclusive type distribution
(b) A is exclusive type, B inclusive type, C open ended type distribution
(c) A is inclusive type, B exclusive type, C open ended type distribution
(d) A open ended type, B exclusive type, C inclusive type distribution

34. Necessity of open ended classes arises when

(a) There are few very high values or few very low values which are far apart from the majority of observations

(b) There are one very high value or one very low value which are far apart from the all other observations

(c) There are few very high values or few very low values which are far apart from the average of observations

(d) There are one very high value or one very low value which are far apart from the average of observations

35. Need for obtaining class boundaries arises for which of the following

(a) When class limits having a gap between upper limit of one class and lower limit of next class

(b) When successive class intervals have one or more values common

(c) When there is an overlap of class intervals

(d) All the above

36. Frequency density of a class in a frequency distribution is

(a) Frequency density = $\frac{\text{Width of the class}}{\text{Class frequency}}$

(b) Frequency density = $\frac{\text{Class frequency}}{\text{Width of the class}}$

(c) Frequency density = Class frequency × Width of the class

(d) Frequency density = $\frac{\text{Class Boundary}}{\text{Class frequency}}$

37. If the class mid points in a frequency distribution of age of a group of persons are 25, 32, 39, 46, 53 and 60 size of the class interval is

(a) 3.5

(b) 4.5

(c) 7

(d) 5

38. Cumulative frequency of a given class in cumulative frequency distribution

(a) Represents the total of all the previous class frequencies

(b) Represents the total of the previous class frequencies and the class against which it is written

(c) Represents the total of all the previous class frequencies including the class against which it is written

(d) Represents the total of all the previous class frequencies except that of open ended class

39. Less than type cumulative frequency in a distribution represent
(a) Total frequency of all values less than and equal to the class value to which it relates
(b) Total frequency of all values less than the class value to which it relates
(c) Total frequency of all values equal to the class value to which it relates
(d) Total frequency of all values less than or equal to the class mark to which it relates

40. In a cumulative frequency distribution more than cumulative frequencies correspond to
(a) Lower limit or boundaries
(b) Upper limit or boundaries
(c) Lower class mark
(d) Upper class mark

41. Histogram is a
(a) Volume diagram
(b) Width diagram
(c) Area diagram
(d) Perimeter diagram

42. In a histogram the area of a rectangle drawn on a class interval is
(a) Equal to the number of frequencies in that class multiplied by its cumulative frequency
(b) Equal to the number of cumulative frequencies in that class multiplied by the range of the class interval
(c) Equal to the number of frequencies in that class multiplied by its class mark
(d) Equal to the number of frequencies in that class multiplied by the range of the class interval

43. In Histogram with unequal class interval the height of each rectangle is proportional to
(a) Class mark
(b) Frequency density
(c) Frequency multiplied by Class mark
(d) Width of class interval

44. Frequency polygon is drawn on the presumption that
(a) The frequencies in a class interval are concentrated at ends
(b) The frequencies in a class interval always increase in steps
(c) The frequencies in a class interval are evenly distributed
(d) The frequencies in a class interval increase evenly.

45. The total area under a frequency curve is
(a) Approximately equal to area under the original Histogram
(b) Approximately greater than the area under the original Histogram
(c) Approximately less than the area under the original Histogram
(d) Exactly equal to area under the original Histogram

46. Cumulative frequency polygon is

(a) Rectangle

(b) Square

(c) Bar

(d) Ogive

47. With respect to a frequency distribution which one of the following is not correct?

(a) The smallest and largest values that go into any given class of a frequency distribution are referred to as the class limits

(b) In comparison to a data array, the frequency distribution has the advantage of representing data in compressed form

(c) As a general rule statisticians regard a frequency distribution as incomplete if it has fewer than 20 classes

(d) The distinction between real class limits and stated class limits is made only when we are dealing with continuous variable

48. Which of the following represents the most accurate scheme of classifying data

(a) Quantitative methods

(b) Qualitative methods

(c) A combination of quantitative and qualitative methods

(d) A scheme that is determined only with specific information about the situation

49. Which one of the following is not an example of compressed data?

(a) Frequency distribution

(b) Data array

(c) Histogram

(d) Ogive

50. When constructing a frequency distribution the first step is

(a) Calculate the class marks for the data

(b) Sort the data points into classes and count the number of points in each class

(c) Decide on the type and number of classes for dividing the data

(d) None of these

51. As the number of observations and classes increase, the shape of a frequency polygon

(a) Tends to become increasingly smooth

(b) Tends to become jagged

(c) Stays the same

(d) Varies only if data become more reliable

MCQ Answers:

1	b	11	c	21	c	31	c	41	c	51	a
2	d	12	c	22	a	32	b	42	d		
3	a	13	b	23	b	33	b	43	b		
4	b	14	d	24	a	34	a	44	c		
5	b	15	c	25	d	35	d	45	a		
6	d	16	c	26	b	36	b	46	d		
7	a	17	a	27	a	37	c	47	c		
8	c	18	b	28	d	38	c	48	d		
9	d	19	b	29	b	39	a	49	b		
10	a	20	d	30	b	40	a	50	c		

II. Exercise Problems:

1. Following the rise in cost of living, the D.A. of the employees of a firm was increased at the following rates

₹20 for the pay range up to ₹199
₹25 for the pay range up to ₹200-399
₹38 for the pay range up to ₹400-599
₹45 for the pay range up to ₹600-799
₹52 for the pay range up to ₹800-999

No increase in the grade with pay ₹1000 or more. What will be the additional amount required to be paid by a department in a year which has 25 employees with the following pay in rupees

872	**164**	**632**	**792**	**285**	**169**	**684**	**485**	**535**	**932**	**123**
785	**338**	**219**	**542**	**185**	**992**	**508**	**180**	**380**	**428**	**272**
136	**180**	**429**								

Answer:

The given information is shown in the following frequency distribution

Pay Range	No. of Students
Up to 199	7
200 – 399	5
400 – 599	6
600 – 799	4
800 – 900	3
Total	25

Detail computation for the answer is as below:

Pay Range	No. of Students (f)	DA in ₹ (x)	xf
Up to 199	7	20	140
200 – 399	5	25	125
400 – 599	6	38	228
600 – 799	4	45	180
800 – 900	3	52	156
Total	25		829

Additional Amount required is ₹829

2. What are the differences between Histogram and bar diagram?

Answer:

- A histogram is a two dimensional diagram where both the width (base) and the length(height of the rectangle) are important whereas bar diagram is one dimensional in which only length (height of the bar) matters while width is arbitrary;
- In a histogram the bars (rectangle) are adjacent to each other whereas in bar diagram proper spacing is given between different bars;
- In a histogram the class frequencies are represented by the area of the rectangles while in a bar diagram they are represented by the heights of the corresponding bars.

III. State the following statements are true or false

1. **A sample is a collection of some but not all, of the elements of the population.**
2. **Frequency table is an uncompressed way of presenting data**
3. **A frequency distribution shows the number of observations from the data set that fall into each of the classes**
4. **Open ended class allows either the upper or lower end of a quantitative classification scheme to be limitless**
5. **The distinction between real limits and stated limits is made only with discrete variables**
6. **Graphs of frequency distributions and relative frequency distributions are useful because they emphasize and clarify trends**
7. **The relative frequency histogram has the same shape as an absolute frequency histogram made from the same data set**
8. **Cumulative frequency of the first class interval is the frequency itself in case of less than type cumulative frequency distribution only**
9. **One disadvantage of the data array is that it does not allow us to easily find the highest and lowest values in the data set**
10. **If we were to connect the midpoints of the consecutive bars of a frequency histogram with a series of lines, we would be graphing a frequency polygon**

Answers:

1/(T), 2/(F), 3/(T),4/(T), 5/(F), 6/(T), 7/(T), 8/(F), 9/(F), 10/(T)

Study Note – 4

Measures of Central Tendency and Dispersion

Learning Objective:

After studying this chapter, the student will be able to:

- *Recognise central tendency and various measures of central tendency,*
- *Explains and evaluates various measures of central tendency,*
- *Evaluates and interprets partition values – Quartiles, Deciles and Percentiles,*
- *Recognizes the importance of measuring dispersion,*
- *Explains and evaluates the measures of dispersion-Range,*
- *Quartile deviation, Mean deviation, Standard deviation,*
- *Recognizes the importance of measuring dispersion,*
- *Explains and evaluates the measures of dispersion-Range, Quartile deviation, Mean deviation, Standard deviation etc.*
- *Discusses about absolute and relative measures of dispersion.*

I. Multiple Choice Questions: (In each question choose the appropriate option out of four options given)

1. Numerical ways of describing quantitative data is named
(a) Measures of Location
(b) Measures of Uncertainties
(c) Measures of Shapes
(d) Measures of errors

2. An average is a
(a) Measure of Dispersion that shows the central value of the data
(b) Measure of Shapes that shows the inclination of data to a specific value
(c) Measure of uncertainties that make the data set ambiguous
(d) Measure that shows the central value of the data set.

3. If we consider several sets of data using only central value, we may draw
(a) A probabilistic conclusion
(b) A leptokurtic conclusion
(c) An unbiased conclusion
(d) An erroneous conclusion

4. Dispersion is also called
   (a) Spread
   (b) Variation
   (c) Inclination
   (d) (a) & (b) not (c)

5. Average annual income of executives of two companies A & B are equal at ₹1,00,000. Therefore with this information only we could safely conclude that
   (a) Distributions of the salaries of executives in A and B are same
   (b) Distributions of the salaries of executives in A and B are mesokurtic
   (c) Distributions of the salaries of executives in A and B cannot be commented upon
   (d) Distributions of the salaries of executives in A and B are different

6. Which one of the following is not a measure of location?
   (a) Weighted Mean
   (b) Geometric Mean
   (c) Mean Deviation
   (d) Median

7. In a data set the sum of deviations of each value from the Arithmetic mean is always
   (a) 0
   (b) > 0
   (c) ≥ 0
   (d) ≤ 0

8. Which of the following is appropriate w.r.t Arithmetic Mean
   (a) Arithmetic mean uses the value of all non-zero items in a sample or population
   (b) Arithmetic mean uses the value of all items in a sample or population
   (c) Arithmetic mean uses the value of all non-negative items in a sample or population
   (d) Arithmetic mean uses the value of items except zero in a sample or population

9. There are 180 students in a college. In the last annual examination total marks obtained by all the students were collected and mean is 76. Which one of the following statement is correct?
   (a) Population parameter is 76
   (b) Sample statistic is 76
   (c) Sample parameter is 76
   (d) Population statistic is 76

10. There are 5 salesperson in a Maruti Car showroom. The number of new cars sold in the last month by the respective salesperson are 9, 120, 5, 8, and 11. For such a data set measure of location is better described by

a) Mean
b) Coefficient of variation
c) Median
d) Standard deviation

11. Median is defined as
a) Average of the maximum and minimum values in the data set considered
b) Midpoint of the items after all the items in the data set are arranged from minimum to maximum values
c) Average of all the values corresponding to items in a data set excluding the items with minimum and maximum values
d) Midpoint of the items after deleting the items in the data set with minimum and maximum values

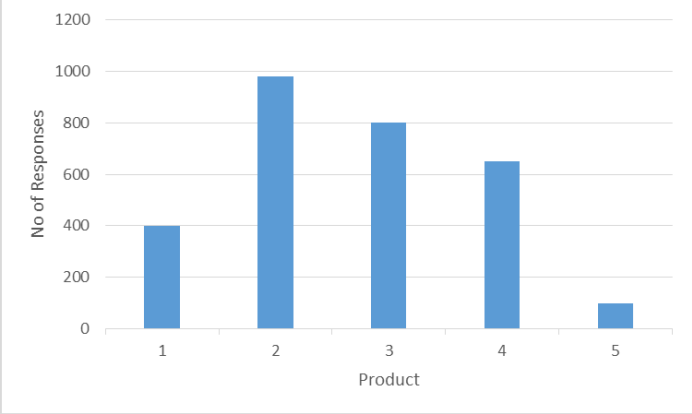
12. Which one of the following is correct?
a) Median and Mean of a data set will never be one of the given values
b) Median of a data set with even number of items will never be one of the given values
c) Median of a data set with odd number of items will never be one of the given values
d) Median and Mean of a data set will always be one of the given values

13. A sample of 10 adults revealed they spent the following number of hours last month using Facebook.

3	5	7	5	9	1	3	9	17	10

The median number of hours is
(a) 7
(b) 9
(c) 6
(d) 5

14. A survey was taken to find which of the five products 1,2,3,4 or 5 has the highest consumer preference. The result is shown below.



From above we could conclude

(a) Product 5 is the mode

(b) Product 2, 3, 4 all are mode

(c) Product 1 & 5 are mode

(d) Product 2 is mode

15. The ages of the individuals in a sporting club are 22,26,27,27,28,29,30,31,31,38,39. The mode of this set is

(a) 27 & 31

(b) 27

(c) 31

(d) 29

16. A set of observations has 12 observations and no two values are the same. The mode of this set is

(a) 12

(b) 6

(c) 0

(d) Cannot be found with the given information

17. A set of observations has 8 observations and they are all same. The mode of this set is

(a) 0

(b) The given value

(c) Sum of the given values

(d) 8

18. The following is the percentage change in net income from last year to this year for a sample of 12 construction companies in Chennai

5	1	-10	-6	5	12	7	8	6	5	-1	11

Median of the set is

(a) 1

(b) 6

(c) 5.5

(d) 5

19. Several indicators of long-term economic growth in a country and their annual percentage change are listed below:

Economic Indicator	Percentage change (%)	Economic Indicator	Percentage change (%)
Inflation	4.5	Real GNP	2.9
Exports	4.7	Investment(domestic)	3.6
Imports	2.3	Investment (foreign)	2.1
Real domestic income	2.9	Productivity	1.4
Consumption	2.7	Manufacturing Productivity	5.2

Modal percentage change is

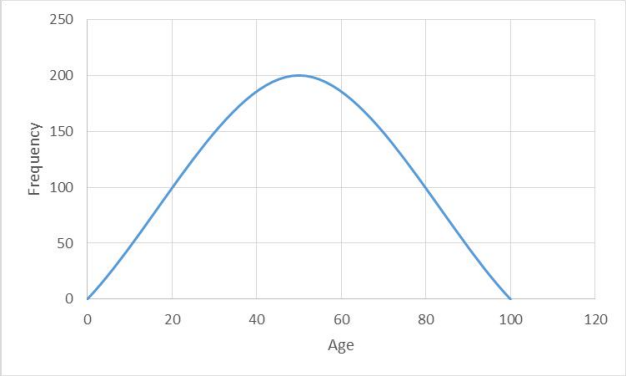
(a) 5.2

(b) 2.9

(c) 3.23

(d) 1.4

20. For the following frequency distribution, we could state



(a) Mean< Median<Mode

(b) Mean>Median>Mode

(c) Mean = Median = Mode

(d) Median = Mode

21. Which one of the following is not a mathematical average?

(a) Geometric mean

(b) Mode

(c) Harmonic mean

(d) Arithmetic mean

22. Arithmetic average ($\overline{X}$) is characterized as the Centre of Gravity because

(a) $\Sigma(X - \overline{X}) > 0$

(b) $\Sigma X > 0$

(c) $\Sigma(\overline{X} - X) < 0$

(d) $\Sigma(X - \overline{X}) = 0$

23. A data set contains 5 items as 3, 5, 8, 12, and 15. With respect to the data set which one of the following is correct?

(a) $\Sigma(X - 12)^2 < \Sigma(X - 8.6)^2$**, where X is the value of items**

(b) $\Sigma(X - 15)^2 < \Sigma(X - 8.6)^2$**, where X is the value of items**

(c) $\Sigma(X - 8.6)^2$ **is minimum, where X is the value of items**

(d) $\Sigma(X - 3)^2$ **is minimum, where X is the value of items**

24. In a grouped data $\Sigma f = 60, \Sigma fd = -7$, where d is the deviation of each value in data set from assumed mean and f is the frequency of each value. If assumed mean is 62, the arithmetic average of the frequency distribution is

(a) -0.117

(b) 61.88

(c) 0.117

(d) 62.117

25. In a frequency distribution$\Sigma f = 63, \Sigma fm = 1780$, where m is the mid value of each class and f is the frequency of each class. If assumed mean of the distribution is 35, the arithmetic average of the frequency distribution is

(a) 63.25

(b) 28.25

(c) -15.86

(d) 6.75

26. The numbers 3.2, 5.8, 7.9 and 4.5 have frequencies $x, (x + 2), (x - 3)$ and $(x + 6)$. If the arithmetic mean is 4.876, the value of x is

(a) 1.63

(b) 2.98

(c) 8

(d) 5

27. The average monthly salaries of different categories of employees of a firm are as follows:

Designation	Number	Mean monthly salary (₹)in lakhs
Senior executives	2	4
Junior executives	5	2
Non-executives	8	1

The overall average monthly salary of the firm in lakhs is

(a) ₹8.67

(b) ₹1.73

(c) ₹2.33

(d) ₹0.47

28. From a population, random samples are drawn and for each sample different averages like arithmetic average, median and mode are computed. Which one of the following statement is correct?

(a) Medians of different samples show least variation

(b) Medians of different samples show least variation

(c) Arithmetic means of different samples show least variation

(d) Medians and Modes of different samples show same minimum variation

29. The following table shows a frequency distribution

Class	Frequency
Below 10	2
10-20	5
20-30	3
30-40	6
40-50	3
Above 50	1

The arithmetic average of the distribution

(a) Can be easily determined by the formula $\bar{X} = \frac{\Sigma fx}{\Sigma f}$, where x is the mid value of each class and f is the frequency of each class

(b) Can be easily determined by the formula $\bar{X} = A + \frac{\Sigma fd}{\Sigma f}$, where A is the assumed mean, and d is the deviation of each item from A

(c) Can never be determined

(d) is always greater than Median but less than Mode.

30. Mean of 20 values is 45. If one of these values is to be taken 64 instead of 46, the corrected mean is

(a) 46.4

(b) 44

(c) 44.1

(d) 45.9

31. The average monthly sales for the first 11 months of the year in respect of a certain salesman were ₹12000 but due to his illness during the last month the average sales for the whole year came down to ₹11375. The value of his sale during the last month is

(a)₹2250

(b)₹4500

(c)₹2255

(d)₹346

32. There were 500 workers working in a factory. Their mean was calculated as ₹200. Later on it was discovered that the wages of two workers were misread as ₹180 and ₹20 in place of ₹80 and ₹120. The correct mean is

(a) ₹200

(b) ₹200.8

(c) ₹199.6

(d) ₹200.4

33. The sum of the deviations of a certain number of observations measured from 4 is 72 and the sum of the deviations of the observations from 7 is -3. No of observations and mean is
(a) 23, 7.13
(b) 24, 7
(c) 25, 6.88
(d) 72, -3

34. The mean of a certain number of items is 42. If one more item is 64 is added to the data the mean becomes 44. The number of items in the original data is
(a) 10
(b) 22
(c) 32
(d) 54

35. Following data shows annual income per head of different class of people in a town

Class of people	Number of families	Annual income/head(₹)
Professor	3	5000
School teacher	10	3000
Student		-1000

Geometric mean of this group is
(a) ₹ 3,000
(b) 0
(c) Imaginary
(d) Infinite

36. In a data set when it is desired to give more weightage to smaller items and small weightage to large items then most appropriate average to be used is
(a) Median
(b) Mode
(c) Arithmetic mean
(d) Geometric mean

37. A person walks 8 km at 4 km/hour, 6 km at 3km/hour and 4 km at 2km/hour. The average speed can be obtained by
(a) Geometric mean
(b) Median
(c) Mode
(d) Harmonic mean

38. The following data pertains to the number of members in a family

No of members	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	1	3	5	6	10	13	9	5	3	2	2	1

The median size of the families is

(a) 7

(b) 8

(c) 6

(d) 5

39. You are given below a certain statistical distribution

Values	Less than 100	100-200	200-300	300-400	400 & above
Cumulative Frequency	40	129	277	341	380

Median of the distribution is

(a) 315.541

(b) 415.541

(c) 141.216

(d) 241.216

40. Median of data set 1 containing items 10,20,50,60, 70 is equal to Median of data set 2 containing items 15, 25,50,65,75. Because of this property it is sometimes said

(a) Median is insensitive

(b) Median is sensitive

(c) Median is biased

(d) Median is neutral

41. Which one of the following is correct?

(a) 3rd quartile has 25% of total number of observations below it and 75% of the observations above it

(b) 2nddecile has 90% of total number of observations above it and 10% of the observations below it

(c) 35th Percentile has 35% of total number of observations below it and 65% of the observations above it

(d) For a certain data set Percentile Decile and Quartile never become equal to each other

42. With respect to a data set if P_k indicates k^{th}Percentile, $D_k$indicatesk^{th} decile & Q_2 the median, then which one of the following is correct?

(a) $P_{50} = D_5 = Q_2$

(b) $P_{50} > D_5 = Q_2$

(c) $P_{50} > D_5 > Q_2$

(d) $P_{50} = D_5 > Q_2$

43. From a data set the following frequency distribution is constructed

Class Interval	Cumulative Frequency
50-60	1
60-70	5
70-80	10
80-90	13
90-100	18
100-110	24
110-120	28
120-130	34
130-140	38
140-150	42
150-160	44

P_{35} of the above distribution is

(a) 94.8

(b) 72.8

(c) 112

(d) 146.8

44. In a frequency distribution the 50th percentile belongs to class 40-42. If there are 250 items and median is 40.295 then for the same distribution 5thdecile is

(a) 8.095

(b) 40.295

(c) 4.0295

(d) 0.40925

45. Marks obtained by 50 students in Statistics are placed in a frequency distribution and from there percentiles are obtained as $P_{10} = 11.67, P_{50} = 27.5, P_{80} = 40, P_{90} = 47.14$. The range of marks obtained by the middle 80% of the students is

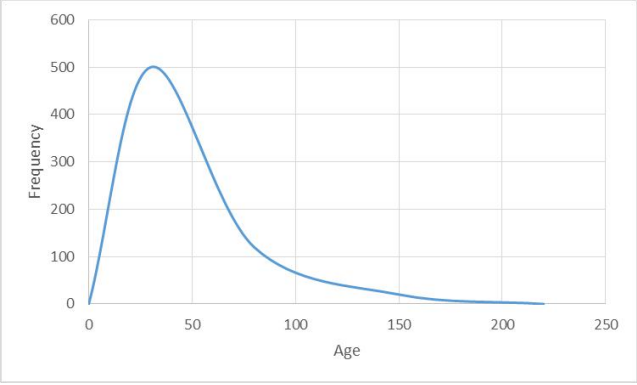
(a) 40

(b) 28.33

(c) 35.47

(d) 12.5

46. For the following frequency distribution, we could state



(a) Mean< Median<Mode
(b) Mean>Median>Mode
(c) Mean = Median = Mode
(d) Median = Mode

47. The weekly sales from a sample of hi tech electronic supply stores were organized into a frequency distribution. The mean of weekly sales was computed to be ₹100, 000, the median ₹115, 000 and the mode ₹119, 000. The distribution is
(a) Positively skewed
(b) Symmetrical
(c) Negatively skewed
(d) Open ended

48. The unemployment rate in India in thousands million by month of year 2018 is given in the table below

Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
8.7	8.8	8.7	7.8	7.3	7.8	6.6	6.5	6.5	6.8	7.3	7.6

Mode of the unemployed rate is
(a) 8.8
(b) 7.2
(c) 8.7, 7.8, 6.5
(d) 8.7, 7.8, 7.3, 6.5

49. KPC hospital employs 200 persons on the nursing staff. Fifty are nurse's aides, 50 are practical nurses and 100 are registered nurses. Nurse's aides receives 1200 an hour, practical nurses 1500 an hour and registered nurses 2000 an hour. The weighted mean hourly wage rate is
(a) ₹1567
(b) ₹1675
(c) ₹4700
(d) None of the above

50. For finding out the percentage changes in sales YOY basis of a business we have to use
(a) Arithmetic mean
(b) Median
(c) Harmonic mean
(d) Geometric mean

51. H.M of 3, 6, 12, and 24 whose weights are 6,2,4,8 respectively is
(a) 6
(b) 6.67
(c) 0.33
(d) 3.33

52. Unless all the values in the given data is positive, this mean cannot be calculated. The mean is
(a) Arithmetic mean
(b) Harmonic mean
(c) Geometric mean
(d) Mode

53. For a given set of data $\mathbf{AM} = \frac{14}{3}$, and $\mathbf{GM} = 4$. HM is
(a) 14/12
(b) 3/16
(c) 24/7
(d) 12/14

54. The following frequency distribution table has some missing value indicated as (*)

Class Boundary	Frequency	Cumulative frequency
2-4	3	3
4-6	*	*
6-8	*	*
8-10	*	10

If median class is 4-6 and median is 5, the simple frequency of the median class is
(a) 4
(b) 5
(c) 7
(d) 3

55. For which of the following distribution $\mathbf{Q_1 + Q_3 = 2Q_2}$ is true where Q represents Quartile
(a) Positively skewed
(b) Negatively skewed
(c) Symmetrical
(d) Median

56. Quartiles of the numbers 29, 12, 26, 19, 24, 36, 21, 33, 35 are
(a) 12, 24, 33
(b) 19, 26, 35
(c) 21, 26, 33
(d) 20, 26, 34

57. Which one of the following is correct?
(a) Three times of excess of mean over median is greater than excess of mean over mode
(b) Three times of excess of mean over median is less than excess of mean over mode
(c) Three times of excess of mean over median is equal to the excess of mean over mode
(d) Three times of excess of mean over mode is greater than excess of mean over median

58. The harmonic mean of the numbers $1, \frac{1}{2}, \frac{1}{3} \ldots \ldots \ldots \frac{1}{n}$ is
(a) 1/n
(b) 2/ (n+1)
(c) n
(d) 1/(n+1)

59. For any set of positive values of a variable we can write
(a) AM $\geq$ GM $\geq$ HM
(b) AM $=$ GM $=$ HM
(c) AM $\leq$ GM $\leq$ HM
(d) AM $\geq$ GM $=$ HM

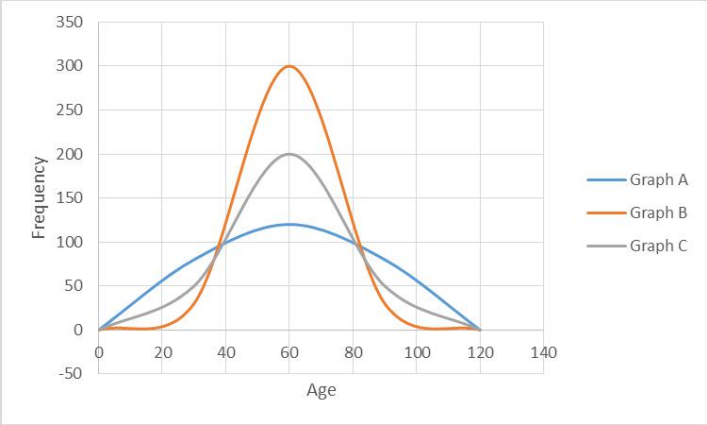
60. For a business unit when for a particular parameter when ratio changes are more important than the absolute changes then it is better to use
(a) Geometric mean
(b) Mode
(c) Arithmetic mean
(d) Median

61. Difference between the maximum and minimum values in a data set is named
(a) Variance
(b) Standard deviation
(c) Mean absolute deviation
(d) Range

62. The arithmetic mean of the squared deviations of respective data in a data set from the mean of the data is named
(a) Standard deviation
(b) Variance
(c) Quartile deviation
(d) Range

63. The financial analyst of a business unit is concerned about the variability of the unit's earning. Under this context to operate effectively he should have adequate knowledge on
(a) Central tendencies
(b) Dispersion
(c) Central tendencies & Dispersion
(d) None of the above.

64. Study the following graphs



In which of the distribution Mean is more representative of the data?

(a) B

(b) C

(c) A

(d) Cannot be said with the given information

65. Which of the following is not one of the reasons for measuring the dispersions of a distribution?

(a) It provides an indication of the reliability of the central tendency measure

(b) It enables us to compare several samples with similar averages

(c) It uses more data in describing a distribution

(d) It draws attention to problems associated with very small or very large variability in distributions

66. Which one of the following is appropriate?

(a) The greater the dispersion of the data in a distribution the lesser the reliability of the central value of the data

(b) The lesser the dispersion of the data in a distribution the greater the reliability of the central value of the data

(c) The dispersion of the data in a distribution has impact on measuring the reliability of the central value of the data

(d) All the above

67. Dispersion is called average of the second order because

(a) It is known by squaring the deviations in individual values from some representative value

(b) It is known by summing up the deviations in individual values from some representative value

(c) It is known by averaging the deviations in individual values from some representative value

(d)It is known by taking square root of the deviations in individual values from some representative value

68. Which one of the following is not a characteristic of coefficient of dispersion?

(a) It is a relative measure of dispersion

(b) It is dependent on units of measurement

(c) It is a ratio of absolute dispersion to an appropriate average

(d) It can be used to compare relative accuracy of the data

69. Which of the following measure of dispersion measures spread of data on the basis of distance between the selected values in a data set?

(a) Quartile Deviation

(b) Mean Deviation

(c) Variance

(d) Range

70. A data set contains values $x_1, x_2, x_3, x_4 \ldots\ldots\ldots\ldots x_n$. Quasi range of the data set is

(a) $x_{n-1} - x_2$

(b) $x_n - x_1$

(c) $x_n - x_{n-1}$

(d) $x_2 - x_1$

71. Frequency distribution of a data set is given below

Size of items	60-62	63-65	66-68	69-71	72-74
No of items	5	18	42	27	8

Range of the above distribution is

(a) 15

(b) 12

(c) 13.5

(d) 37

72. Frequency distribution of a data set is given below

Size of items	10-12	13-15	16-18	19-21	22-24
No of items	22	8	35	23	12

Quasi Range of the above distribution is

(a) 15

(b) 12

(c) 13.5

(d) 27

73. Two series A with items (5,6,8,10,12,16,19,21,25) and B with items (25,30, 35,40,45) are given. Which series shows greater variability?
(a) Series A as it has a coefficient of variation 0.066
(b) Series B as it has a range 20
(c) Series B as it has a coefficient of variation 0.20
(d) Series A as it has a quasi-range of 15

74. Range does not fulfill one of the characteristic of a good measure of dispersion. The characteristic is
(a) Good measure of dispersion should be hardly comprehensible
(b) Good measure of dispersion should be based on all the observation
(c) Good measure of dispersion should be affected by sampling variation
(d) Good measure of dispersion should not be amenable to algebraic treatment

75. Following frequency distribution is given

Days of week	1	2	3	4	5	6	7
No of units sold	12	45	20	28	40	15	30

Quartile deviation of the distribution is
(a) 0.455
(b) 12.5
(c) 2.2
(d) 0.227

76. For a frequency distribution with open ended classes the most appropriate measure of dispersion is
(a) Mean Deviation
(b) Range
(c) Quartile Deviation
(d) Standard Deviation

77. Quartile deviation ignores
(a) 50% of the observation
(b) 25% of the observation
(c) Extreme values of the observation
(d) 37.5% of the observation

78. In true sense Quartile deviation is not a measure of dispersion as
(a) It does not lend itself to further mathematical treatment
(b) It is affected considerably by fluctuations of sampling
(c) It really does not show scatter around the average
(d) It excludes the extremes

79. Following data is given.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Sales(₹'000)	78	80	80	82	82	84	84	86	86	88	88	90

Quartile deviation of the above data set is

(a) 7.0

(b) 0.04

(c) 0.02

(d) 3.5

80. Following frequency distribution relates wages and no of workers

Wages (₹)	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No of workers	20	45	85	160	70	55	35	30

The percentage of workers getting wages between ₹22 and ₹58 is

(a) 68.4%

(b) 50%

(c) 37.5%

(d) 75%

81. Following frequency distribution is given

Size of items	59.5-62.5	62.5-65.5	65.5-68.5	68.5-71.5	71.5-74.5
No of items	5	18	42	27	8

With respect to above which one of the following is correct?

(a) Range > Quasi range> Percentile range $(P_{90} - P_{10})$

(b) Range < Quasi range < Percentile range $(P_{90} - P_{10})$

(c) Range > Quasi range =Percentile range $(P_{90} - P_{10})$

(d) Range > Quasi range < Percentile range $(P_{90} - P_{10})$

82. Mean deviation is

(a) Based on position of certain items in a distribution

(b) Based on all values except extremes

(c) Based on all values

(d) Based on only on extreme values

83. Mean deviation gives a relative measure of dispersion and compares

(a) Two or more series having different means

(b) Two or more series of different quartile deviation

(c) Two or more series expressed in same units but of different orders of magnitude

(d) Two or more series with different mean but of same quartile deviation

84. Coefficient of Mean deviation is

(a) $\frac{\text{Average absolute amount of scatter of the items in a distribution from mean}}{\text{Mean}}$

(b) $\frac{\text{Average absolute amount of scatter of the items in a distribution from median}}{\text{Median}}$

(c) $\frac{\text{Average absolute amount of scatter of the items in a distribution from mode}}{\text{Mode}}$

d) $\frac{\text{Average absolute amount of scatter of the items in a distribution from mean or median or mode}}{\text{Mean or median or mode}}$

85. For a frequency distribution which one of the following formula is correct?

(a) $\frac{\Sigma f|X-M|}{N}$, where Mis mean, X is respective item in a distribution, f is the frequency, N is the total item

(b) $\frac{\Sigma f|X+M|}{N}$, where Mis mean, X is respective item in a distribution, f is the frequency, N is the total item

(c) $\frac{\Sigma f|X*M|}{N}$, where Mis mean, X is respective item in a distribution, f is the frequency, N is the total item

(d) $\frac{\Sigma f|X/M|}{N}$, where Mis mean, X is respective item in a distribution, f is the frequency, N is the total item

86. For measuring dispersion which one of the following is most preferable?

(a) Range

(b) Quartile deviation

(c) Mean deviation

(d) Quasi Range

87. For a frequency distribution $\Sigma fx = 576, \Sigma|x - \bar{x}| = 36$ &$\bar{x} = 12$, where x is the value of respective items in the distribution, f is the frequency &$\bar{x}$ is the mean of the distribution. The mean deviation of the distribution is

(a) 3

(b) 0.75

(c) 16

(d) 0.0625

88. For a set of ungrouped data $N = 15, \Sigma x = 480, \Sigma x^2 = 15735$. For the data set, Standard deviation is

(a) 5

(b) 32

(c) 6.4

(d) 25

89. Interquartile range measures

(a) Approximately how far from the median we must go on either side before we can include one half values of the data set

(b) Approximately how far from the median we must go on either side before we can include three fourth values of the data set

(c) Approximately how far from the median we must go on either side before we can include one fourth values of the data set

(d) Approximately how far from the mean we must go on either side before we can include one half the values of the data set

90. Find the interquartile range from the daily wages (in ₹) 7 persons given below 12, 7, 15, 10, 19, 17, and 25

(a) ₹13

(b) ₹18

(c) ₹9

(d) ₹4.5

91. Which one of the following is affected by extreme values?

(a) Interquartile range

(b) Range

(c) Median

(d) Mode

92. Average absolute deviation or mean deviation is better measure of dispersion than the ranges because

(a) It is simplest to calculate

(b) It is based on one single value

(c) It is least influenced by extreme values

(d) It takes every observation into account

93. The formula $\Sigma\frac{x^2}{N} - \mu^2$ indicates

(a) Sample variance where x is the item or observation, μ is the population mean and N total number of items in the population

(b) Population standard deviation where x is the item or observation, μ is the population mean and N total number of items in the population

(c) Mean deviation of the population where x is the item or observation, μ is the population mean and N total number of items in the population

(d) Population variance where x is the item or observation, μ is the population mean and N total number of items in the population

94. The original data is in units "thousands of ₹". Which of the following measure has not the same unit with the original data?
(a) Variance
(b) Quartile deviation
(c) Mean
(d) Mode

95. The standard deviation determine
(a) Where the values of a frequency distribution are located in relation to a central tendency
(b) Where the values of a frequency distribution are located in relation to the variance
(c) Where the values of a frequency distribution are located in relation to the mean
(d) Where the values of a frequency distribution are located in relation to coefficient of variation

96. Standard score measures
(a) How far individual items in a distribution depart from the standard deviation of the distribution
(b) How far individual items in a distribution depart from the mean of the distribution
(c) How far individual items in a distribution depart from the median of the distribution
(d) How far individual items in a distribution depart from the symmetry of the distribution

97. The mean and variance of a population are 0.166 and 0.0034. An observed data 0.282 in the population
(a) Deviates from the mean by 0.116 units
(b) Deviates from the median by 2 units
(c) Deviates from the mean by 2 units
(d) Deviates from the median by 0.116 units

98. For a sample of 12 items, $\bar{x} = \text{mean} = 1.351, \Sigma x^2 = 23.496.182$. Sample variance is
(a) 380.64
(b) 1.825201
(c) 17391.697
(d) 144,888

99. Two samples of size 50 each, the mean value for the first sample (1) is 1.16 with a standard deviation of 0.21. The second sample (2) has a mean of 1.75 and a standard deviation of 0.35. Hence
(a) First Sample has less relative dispersion since $s_1 < s_2$, where s is the standard deviation
(b) Second Sample has greater relative dispersion since $s_1 > s_2$, where s is the standard deviation
(c) First Sample has less relative dispersion since $\overline{x_1} < \overline{x_2}$, where $\bar{x}$ is the mean
(d) Second Sample has greater relative dispersion since standard deviation as a percentage of mean in sample 2 is greater than that in sample 1

100. Fractiles that divide data into ……equal parts are called……. Which pair of phrases best completes the above sentence?
(a) 100, deciles
(b) 4, quartiles
(c) 10, percentiles
(d) 16, octiles

101. Why is it necessary to square the differences from the mean when computing the population variance?
(a) So that extreme values will not affect the calculation
(b) Because it is possible that N i.e., total number of items could be very small
(c) Some of the differences will be positive and some will be negative
(d) All the above

102. Assume that a population has $\mu = \text{mean} = 100$, $\sigma = \text{standard deviation} = 10$. If a particular observation has a standard score of 1, it can be concluded that
(a) Its value is 110
(b) It lies between 90 and 110 but its exact value cannot be determined
(c) Its value is greater than 110
(d) Nothing can be determined with the given information

103. If one were to divide the standard deviation of a population by the mean of the same population and multiply this value by 100, one would have calculated
(a) Population standard score
(b) Population coefficient of variation
(c) Population variance
(d) Average absolute deviation

104. The measure of the average squared distance between the mean and each item in the population is the……..and the positive square root of this value is the…….
Which pair of phrases best completes the above sentence?
(a) Average absolute deviation, Dispersion
(b) Mean squared deviation, Root mean square
(c) Variance, Standard deviation
(d) Quartile deviation, Coefficient of variation

105. Which one of the following is correct?
(a) Population variance is the mean of the difference between each value and mean
(b) For populations whose values are near the mean, the variance will be large
(c) If items in a population are expressed in ₹, variance of the population will also be in ₹
(d) For populations whose values are very dispersed from the mean, the variance will be large

106. While computing variance of population squared differences between each value and the specific central tendency is done to avoid
(a) $\Sigma(x - \mu) = 0$
(b) $\Sigma(\mu - x) = 0$
(c) $\Sigma x - \mu = 0$
(d) $\mu - \Sigma x = 0$

107. Square root of the mean of the squares of deviations from mean is
(a) Absolute Mean Deviation
(b) Quartile deviation
(c) Standard score
(d) Standard deviation

108. Which one of the following is correct?
(a) Standard deviation is dependent on the choice of origin
(b) Standard deviation is dependent on the change of scale
(c) Standard deviation is dependent on the range of data considered
(d) Standard deviation is independent on standard score

109. Sampling fluctuations has least effect on
(a) Quartile deviation
(b) Interquartile range
(c) Standard deviation
(d) Mean deviation

110. Variance of a composite set, comprising N number of sets with σ_i as the s.d of 'i' th set and d is the deviation of "i" th set's mean from composite mean is calculated from
(a) $N\sigma^2 = \sum_{i=1}^{N}(n_i{\sigma_i}^2 + n_i{d_i}^2)$
(b) $\sigma^2 = \sum_{i=1}^{N}(n_i{\sigma_i}^2 + n_i{d_i}^2)$
(c) $\sigma^2 = \sum_{i=1}^{N}(n_i{\sigma_i}^2 + n_i{d_i}^2) * N$
(d) $N^2\sigma^2 = \sum_{i=1}^{N}(n_i{\sigma_i}^2)$

111. Percentage variation in the mean is
(a) Standard deviation
(b) Mean deviation
(c) Coefficient of variation
(d) Coefficient of mean dispersion

112. For a distribution Coefficient of quartile deviation is 0.04 and Interquartile range is 7. Third quartile of the distribution is
(a) 84
(b) 91
(c) 70
(d) 77

113. The extract of a frequency distribution is given

Class Interval	10-15	15-20	20-25	25-30	30-40	40-50	50-60	60-70
Frequency	4	12	16	22	10	NA	NA	NA

If 41st item is the median then interquartile range of the distribution is
(a) 8.05
(b) 5.64
(c) 10.45
(d) 16.1

114. Quartile deviation of a distribution is 11. It means that
(a) 1st& 3rd Quartiles differ from the median by 11
(b) 1st& 3rd Quartiles differ from the median by 11on an average
(c) Spread between 1st and 3rd quartile is 11
(d) Spread between 1st and 3rd quartile is 11 on an average

115. A sample is taken from a population containing annual earnings, in ₹, of different class of professionals in a locality. S.D (s) of the sample means
(a) An observation in the sample is at an average distance (s) from the mean of the sample
(b) An observation in the sample is at an average distance (s %) from the mean of the sample
(c) An observation in the sample is at a distance (s) from the mean of the sample
(d) An observation in the sample is at an average distance (₹ s) from the mean of the sample

116. Two samples A and B are given whose means are equal. A is the sample of wage rates in ₹/hour for 20 different classes of workers in India & B is the sample of wage rates in $/hour for same class of workers in USA. Standard deviation of the samples are ₹300/hour (s_A) and $50/hour($s_B$). Which one of the following is correct?
(a) Sample A is more dispersed than sample B as $s_A > s_B$
(b) Sample B is less dispersed than sample A as $ is more valuable than ₹
(c) Dispersion of the samples cannot be compared as population standard deviations are not given
(d) Sample A is more dispersed than sample B as $s_A > s_B$& mean of the samples are equal

117. A person walks 8 km at 4 km/hour, 6 km at 3km/hour and 4 km at 2km/hour. The average speed of the person is
(a) 3 km/hour
(b) 4.5km/hour
(c) 2 km/hour
(d) 6km/hour

118. The values of mode and median for a moderately skewed distribution are 64.2 and 68.6 respectively. The value of the mean is

(a) 66.4
(b) 70.8
(c) 199.2
(d) 6.6

119. In a moderately asymmetrical distribution the values of mode and mean are 32.1 and 35.4 respectively. Median is

(a) 64.2
(b) 34.3
(c) 17.2
(d) 70.8

120. If the mean and median of a moderately asymmetrical series are 26.8 and 27.9 respectively, most probable mode is

(a) 55.8
(b) 53.6
(c) 27.9
(d) 30.1

121. The mean of 200 items is 48 and their standard deviation is 3. Sum of squares of all items is

(a) 9600
(b) 462,600
(c) 360,000
(d) 40,000

122. Which one of the following is not a characteristics of skewness?

(a) A distribution where the value of mean, median and mode do not coincide
(b) $Q_3 - M_d$ is not equal to $M_d - Q_1$, where M_d is median and Q is quartiles
(c) Sum of positive deviations from the median is equal to sum of the negative deviation
(d) Values on a graph paper do not yield a normal bell shaped curve

123. Which one of the following is dispersion?

(a) It deals with the symmetry of distribution of values on both sides of the central value
(b) It helps in finding out whether in a distribution concentration is in higher or lower values
(c) It is not an average although measured by the use of various types of averages
(d) It is the average of the deviations around a central value

124. The value of median in both positively and negatively skewed distribution lies between those of mean and mode almost at a point

(a) 2/3rds of the gap between mean and mode
(b) 1/3rds of the gap between mean and mode
(c) 3/4th of the gap between mean and mode
(d) 1/2 of the gap between mean and mode

125. For a group of 10 items $\Sigma X = 452, \Sigma X^2$= 24,270, and Mode = 43.7. Pearson Coefficient of skewness is
(a) 19.59
(b) 45.2
(c) 0.08
(d) 0.97

126. A distribution is made with prices of houses sold in a city are the items. From the distribution it is observed that peak of the distribution data is right of the average value. It means that
(a) Distribution is positively skewed and houses were being sold at more than the average value
(b) Distribution is negatively skewed and houses were being sold at more than the average value
(c) Distribution is positively skewed and houses were being sold at less than the average value
(d) Distribution is symmetrical but houses were being sold at more than the average value

127. If skewness is between −1 and −½ or between +½ and +1, the distribution is
(a) Moderately skewed.
(b) Highly skewed
(c) Approximately symmetric
(d) Symmetric

128. If skewness is less than −1 or greater than +1, the distribution is
(a) Moderately skewed.
(b) Highly skewed
(c) Approximately symmetric
(d) Symmetric

129. If in a distribution the bulk of the data is at the left and the right tail is longer, we say that
(a) The distribution is skewed right or positively skewed
(b) The distribution is skewed left or negatively skewed.
(c) The distribution is skewed right or negatively skewed
(d) The distribution is skewed left or positively skewed.

130. If in a distribution the peak is toward the right and the left tail is longer, we say that the distribution is
(a) Positively skewed
(b) Symmetric
(c) Right skewed
(d) Negatively skewed

MCQ Answers:

1	a	11	b	21	b	31	b	41	c	51	b	121	b
2	d	12	b	22	d	32	a	42	a	52	c	122	c
3	d	13	c	23	c	33	c	43	a	53	c	123	d
4	d	14	d	24	b	34	a	44	b	54	a	124	a
5	c	15	a	25	b	35	c	45	c	55	c	125	c
6	c	16	c	26	d	36	d	46	b	56	d	126	b
7	a	17	b	27	a	37	d	47	c	57	c	127	a
8	b	18	d	28	c	38	c	48	d	58	b	128	b
9	a	19	b	29	c	39	d	49	b	59	a	129	a
10	c	20	c	30	c	40	a	50	d	60	a	130	d
61	d	71	a	81	a	91	b	101	c	111	c		
62	b	72	b	82	c	92	d	102	a	112	b		
63	c	73	c	83	c	93	d	103	b	113	d		
64	a	74	b	84	d	94	a	104	c	114	b		
65	c	75	a	85	a	95	c	105	d	115	d		
66	d	76	c	86	c	96	b	106	a	116	d		
67	c	77	a	87	b	97	a	107	d	117	a		
68	b	78	c	88	a	98	d	108	b	118	b		
69	d	79	d	89	a	99	d	109	c	119	b		
70	a	80	a	90	c	100	b	110	a	120	d		

II. Exercise Problems:

1. The distribution of wages in two factories X and Y are given below:

Wages in ₹	No of workers	
	Factory X	Factory Y
50-100	2	6
100-150	9	11
150-200	29	18
200-250	54	32
250-300	11	27
300-350	5	11

State in which factory the wages are more variable.

Answer:

In this type problem we have to compute for each factory coefficient of variation.

Coefficient of variation = $\frac{\text{Standard Devaiation}}{\text{Mean}} \text{ x } 100$

We have used the following formula for mean and standard deviation

Mean = $\text{Assumed mean} + \frac{\Sigma fd'}{\Sigma f} * h$,

Standard deviation = $\sqrt{\frac{\Sigma f(d')^2}{\Sigma f} - (\frac{\Sigma fd'}{\Sigma f})^2} * h$

Where d' is the deviation of respective mid values of classes from the assumed mean/width, f is the frequency of the respective class and h is the width of classes.

Detail computation is given in the following table with assumed mean as 225

	Mid Value	$d' = \frac{m-225}{50}$	Factory X			Factory Y		
Wages	m		f_x	$f_x d'$	$f_x (d')^2$	f_y	$f_y d'$	$f_y (d')^2$
50-100	75	-3	2	-6	18	6	-18	54
100-150	125	-2	9	-18	36	11	-22	44
150-200	175	-1	29	-29	29	18	-18	18
200-250	225	0	54	0	0	32	0	0
250-300	275	1	11	11	11	27	27	27
300-350	325	2	5	10	20	11	22	44
Total			110	-32	114	105	-9	187
Factroy X				Factory Y				
	$\bar{X}$	210.4545455			$\bar{Y}$	220.7143		
	σ_x	48.778467			σ_y	66.58839		
	Cof. Of variation	23.17767331			Cof of variation	30.1695		

From the above it is observed that Coefficient of variation of Factory Y is more than that of Factory X. So we can conclude that distribution of wages in factory Y is more variable than that in X

2. Calculate average profits from the following data in regard to profits and losses of 120 firms:

Profits (₹)	Number of firms	Loss (₹)	Number of firms
5000-6000	**15**	**0-1000**	**5**
4000-5000	**20**	**1000-2000**	**6**
3000-4000	**35**	**2000-3000**	**8**
2000-3000	**15**	**3000-4000**	**3**
1000-2000	**5**	**4000-5000**	**2**
0-1000	**6**		

Answer:

Mean = $\text{Assumed mean} + \frac{\Sigma fd'}{\Sigma f} * h$,

Where d' is the deviation of respective mid values of classes from the assumed mean/width, f is the frequency of the respective class and h is the width of classes.

In this problem Assumed mean is 500
Detail computation is given in the following table:

Profit Losses	Mid Value m	$d' = \frac{m-500}{1000}$	f	fd'
5000 to 6000	5500	5	15	75
4000 to 5000	4500	4	20	80
3000 to 4000	3500	3	35	105
2000 to 3000	2500	2	15	30
1000 to 2000	1500	1	5	5
0 to -1000	500	0	6	0
-1000 to 0	-500	-1	5	-5
-2000 to -1000	-1500	-2	6	-12
-3000 to -2000	-2500	-3	8	-24
-4000 to -3000	-3500	-4	3	-12
-5000 to -4000	-4500	-5	2	-10
Total			120	232

So mean = $\overline{X} = 500 + \frac{232}{120} * 1000 = 2433.33$ ₹

3. Given below is the distribution of 150 candidates obtaining marks X or higher in a certain examination (all marks are given in whole numbers):

X	10	20	30	40	50	60	70	80	90	100
c.f	150	135	120	102	77	47	25	9	2	0

Calculate mean marks obtained by the candidates.

Answer:

From the problem it is observed that cumulative frequency is given. From here we have to construct the following frequency distribution in respect of different classes. From the table it is clear that 150 items are more than 10 whereas 135 items are more than 20. So class 10-20 should have a frequency of 150-135 =15. Accordingly other classes are built and the result is given below in the first table.

From the first table we calculate the mean by formula

Mean = $\text{Assumed mean} + \frac{\Sigma fd'}{\Sigma f} * h$,

Where d' is the deviation of respective mid values of classes from the assumed mean/width, f is the frequency of the respective class and h is the width of classes.

Class	Frequency
	f
10-20	15
20-30	15
30-40	18
40-50	25
50-60	30
60-70	22
70-80	16
80-90	7
90-100	2

Detail computation is given in the following table

Class	Frequency	Mid Value	$d' = \frac{m-55}{10}$	
	f	m		fd'
10-20	15	15	-4	-60
20-30	15	25	-3	-45
30-40	18	35	-2	-36
40-50	25	45	-1	-25
50-60	30	55	0	0
60-70	22	65	1	22
70-80	16	75	2	32
80-90	7	85	3	21
90-100	2	95	4	8
Total	150			-83

So mean = $\bar{X} = 55 - \frac{83}{150} * 10 = 49.47$

4. The following table gives the life time in hours of 500 radio tubes of a certain make. Find the mean life time of the radio tubes

Life time (in hours)	Number of tubes	Life time (in hours)	Number of tubes
Less than 300	**0**	**Less than 800**	**267**
Less than 400	**22**	**Less than 900**	**350**
Less than 500	**62**	**Less than 1000**	**420**
Less than 600	**118**	**Less than 1100**	**480**
Less than 700	**196**	**Less than 1200**	**500**

Answer:

In the question cumulative frequency distribution is given from where we have to construct the simple frequency distribution as follows:

Life time Class	No of tubes (f)
300-399	22
400-499	40
500-599	56
600-699	78
700-799	71
800-899	83
900-999	70
1000-1099	60
1100-1199	20

Detail computation of mean is given below:

Life time Class	No of tubes (f)	Mid Value m	$d' = \frac{m - 749.5}{100}$	fd'
300-399	22	349.5	-4	-88
400-499	40	449.5	-3	-120
500-599	56	549.5	-2	-112
600-699	78	649.5	-1	-78
700-799	71	749.5	0	0
800-899	83	849.5	1	83
900-999	70	949.5	2	140
1000-1099	60	1049.5	3	180
1100-1199	20	1149.5	4	80
Total	500			85

So mean = $\overline{X} = 749.5 + \frac{85}{500} * 100 = 766.5$

5. A factory pays workers on piece rate basis and also a bonus to each worker on the basis of individual output in each quarter. The rate of bonus payable is as follows:

Output in units	Bonus in ₹
70-74	**50**
75-79	**55**
80-84	**60**
85-89	**70**
90-94	**80**
95-99	**90**
100-104	**120**

The individual output of a batch of 50 workers is given below:

94	**83**	**78**	**76**	**88**	**86**	**93**	**80**	**91**	**82**
89	**97**	**92**	**84**	**92**	**80**	**85**	**83**	**98**	**103**
87	**88**	**88**	**81**	**95**	**86**	**99**	**81**	**87**	**90**
84	**97**	**80**	**75**	**93**	**101**	**82**	**82**	**89**	**72**
85	**83**	**75**	**72**	**83**	**98**	**77**	**87**	**71**	**80**

By suitable classification you are required to find

(i) Average bonus per worker for the quarter

(ii) Total quarterly bonus paid to the whole batch

(iii) Average output per worker

Answer:

In this problem from the output table we have to first find out the no of outputs (frequency) of respective class. Say first class is 70-74. We have to search from the output table how many outputs are falling in this class. It is 3—72, 72, 71. Similarly in the second class 75-79 we have 5 no of outputs – 78, 76, 75, 75, 77. This way if we proceed distribution of workers by outputs and bonus will be as follows:

Output in units	No of outputs	Bonus(in Rupees)
(Class)	(f)	(Y)
70-74	3	50
75-79	5	55
80-84	15	60
85-89	12	70
90-94	7	80
95-99	6	90
100-104	2	120

Detail computation table is as follows:

Output in units	No of outputs	Bonus(in Rupees)	Mid Value	$d' = \frac{m-87}{5}$	fd'	fy
			m			
(Class)	(f)	(Y)				
70-74	3	50	72	-3	-9	150
75-79	5	55	77	-2	-10	275
80-84	15	60	82	-1	-15	900
85-89	12	70	87	0	0	840
90-94	7	80	92	1	7	560
95-99	6	90	97	2	12	540
100-104	2	120	102	3	6	240
Total	50				-9	3505

(i) Average bonus per worker per quarter = $\frac{3505}{50} = 70.1$ ₹

(ii) Total quarterly bonus paid to the whole batch = ₹ 3505

(iii) Average output per worker = $87 - \frac{9}{50} * 5 = 86.1$

6. The following table gives details of weekly income and weekly expenditure on breakfast of a group of families:

Family	Total Income (₹)	Expenditure per head (₹)	No of members in family
A	**512.5**	**62.00**	**5**
B	**400.5**	**36.5**	**7**
C	**480.6**	**51.25**	**6**
D	**261.25**	**36.00**	**3**
E	**452.80**	**42.5**	**2**
F	**710.70**	**89.25**	**2**

Find the average income per head of the group and the average expenditure per family. Which families spend more/less than the average expenditure and by how much?

Answer:

The detail computation showing total income and total expenditure is as follows:

Family	Total Income	Expenditure/head	No of members/family	Total Expenditure
A	512.5	62	5	310
B	400.5	36.5	7	255.5
C	480.6	51.25	6	307.5
D	261.25	36	3	108
E	452.8	42.5	2	85
F	710.7	89.25	2	178.5
Total	2818.35		25	1244.5

Average income per head = $\frac{2818.35}{25} = 112.734$

Average expenditure per family = $\frac{1244.5}{6} = 207.42$

Families spending more:

A: $310 - 207.42 = 102.58$

B: $310 - 255.5 = 54.5$

C: $310 - 307.5 = 2.5$

Families spending less:

D: $310 - 108 = 202$

E: $310 - 85 = 225$

F: $310 - 178.5 = 131.5$

7. The following variations were found in the measurement of 120 parts produced by a machine:

No of parts	Variations of the standard (in mm)
1	10 & 5
3	5 & 10
17	0 & 5
27	-5 & 0
20	-10 & -5
12	-15 & -10
13	-20 & -15
15	-25 & -20
10	-30 & -25
2	-35 & -30

(i) Find the average variations

(ii) What proportion fell within a range of 5 mm either way of the standard?

(iii) If those which fall more than 10 mm apart are classed as bad, what percentage of the parts are bad?

(iv) Which stretch of 15 mm contains the greatest number of parts and which fraction of the total fall inside this stretch?

Answer:

Detail computation of the average is from the following table:

Variations	No of parts (f)	Mid value(m)	fm
10 to 15	1	12.5	12.5
5 to 10	3	7.5	22.5
0 to 5	17	2.5	42.5
-5 to 0	27	-2.5	-67.5
-10 to-5	20	-7.5	-150
-15 to -10	12	-12.5	-150
-20 to -15	13	-17.5	-227.5
-25 to -20	15	-22.5	-337.5
-30 to -25	10	-27.5	-275
-35 to - 30	2	-32.5	-65
Total	120		-1195

(i) Average variations = $-\frac{1195}{120} = -9.96$

(ii) No of parts from +5 to 0 & from 0 to -5 = 17+27 = 44. Proportion = $\frac{44}{120} * 100 = 36.67$

(iii) Bad classes are 10 to 15, -15 to -10, -20 to -15, -25 to -20, -30 to -25, and -35 to -30. The corresponding no of parts are 1, 12, 13, 15, 10, 2. Total = 53. Proportion = $\frac{53}{120} * 100 = 44.17$

(iv) A stretch from 5 to -10 is the 15 mm stretch which contains the greatest no of parts. Total no of parts is 0 to 5 = 17, -5 to 0 = 27, -10 to -5 = 20. Total = 64. Proportion = $\frac{64}{120} * 100 = 53.33$

8. The numbers 3.2, 5.8, 7.9 and 4.5 have frequencies x, (x+2), (x-3) and (x+6) respectively. If the arithmetic mean is 4.876 find the value of x.

Answer:

From the question we could write

$$4.876 = \frac{3.2x + 5.8(x + 2) + 7.9(x - 3) + 4.5(x + 6)}{x + (x + 2) + (x - 3) + (x + 6)}$$

Or $4.876 = \frac{3.2x+5.8x+7.9x+4.5x+5.8*2-7.9*3+4.5*6}{4x+2-3+6}$

Or $4.876 = \frac{21.4x + 14.9}{4x+5}$

Or $4.876(4x + 5) = 21.4x + 14.9$

Or $19.504x + 24.38 = 21.4x + 14.9$

Or $21.4x - 19.504x = 24.38 - 14.9$

Or $1.896x = 9.48$

Or $x = \frac{9.48}{1.896} = 5$

9. Calculate an arithmetic mean of the following distribution:

x	1	2	3	4-6	7-9	10-12	13-20	21-28	29-36
f	10	5	3	9	6	2	1	10	15

Answer:

In this problem there are three groups.

First group contains x: 1, 2, 3
Second group x: 4-6, 7-9, and 10-12
Third group x: 13-20, 21-28 and 29-36

The mean calculation of these three groups are as follows:

x	f	xf
1	10	10
2	5	10
3	3	9
Total	18	29
Mean	29/18 =	1.61

Class	Mid value (m)	Frequency (f)	fm
4-6	5	9	45
7-9	8	6	48
10-12	11	2	22
Total		17	115
Mean	115/17 =	6.76	

Class	Mid value (m)	Frequency (f)	fm
13-20	16.5	1	16.5
21-28	24.5	10	245
29-36	32.5	15	487.5
Total		26	749
Mean	749/26 =	28.81	

So mean f the combined group = $\frac{18*1.61+17*6.76+26*28.81}{18+17+26} = 14.64$

10. Comment on the performance of the students in three universities given below using simple and weighted average

Course of study	Mumbai University		Calcutta University		Madras University	
	% of pass	No of students (in 00's)	% of pass	No of students (in 00's)	% of pass	No of students (in 00's)
MA	**70**	**3**	**85**	**2**	**85**	**2**
M.Com	**83**	**4**	**79**	**3**	**86**	**3.5**
BA	**90**	**5**	**88**	**6**	**74**	**4.5**
B.Com	**74**	**2**	**76**	**7**	**75**	**2**
B. Sc	**85**	**3**	**80**	**3**	**73**	**7**
M. Sc	**66**	**3**	**60**	**7**	**75**	**2**

Answer:

The detail computation of simple average and weighted average of three universities are given below:

Mumbai			
Courses of Study	Pass (%) X	No of Students w	wx
MA	70	3	210
MCOM	83	4	332
BA	90	5	450
BCOm	74	2	148
BSc	85	3	255
MSc	66	3	198
Total	468	20	1593
Simple Mean	468/6 =	78	
Wtd Mean	1593/20 =	79.65	

Calcutta			
Courses of Study	Pass (%) X	No of Students w	wx
MA	85	2	170
MCOM	79	3	237
BA	88	6	528
BCOm	76	7	532
BSc	80	3	240
MSc	60	7	420
Total	468	28	2127
Simple Mean	468/6 =	78.00	
Wtd Mean	2127/28 =	75.96	

Madras			
Courses of Study	Pass (%) X	No of Students w	wx
MA	85	2	170
MCOM	86	3.5	301
BA	74	4.5	333
BCOm	75	2	150
BSc	73	7	511
MSc	75	2	150
Total	468	21	1615
Simple Mean	468/6 =	78.00	
Wtd mean	1615/21 =	76.90	

From the above tables we could conclude that we cannot distinguish performance of the universities on the basis of simple average as all are showing same result. But on the basis of wtd average performance of Mumbai University is the best followed by Madras & Calcutta.

11. A car owner buys petrol at ₹70.80, ₹71.2, ₹71.8 and ₹72.8 per litre for four successive years. What was the average cost of oil when he spends of ₹1000 each year or when he buys 1000 litres every year?

Answer:

Average cost of oil is to be expressed in ₹/litre.

Case I: when he spends ₹1000 each year. Computation detail will be:

Price	Litre purchased	Amount spent
70.8	14.124	1000
71.2	14.045	1000
71.8	13.928	1000
72.8	13.736	1000
Total	55.833	4000
Average price	4000/55.833 =	71.642

Case II: when he purchased 1000 litre every year. Computation detail will be:

Price	Litre purchased	Amount spent
70.8	1000	70800
71.2	1000	71200
71.8	1000	71800
72.8	1000	72800
Total	4000	286600
Average price	286600/4000 =	71.650

12. The following data pertaining to the number of members in a family has been collected from a sample survey in 60 families. You have to form a frequency distribution find the mean and the median size of the families and indicate which of the two methods is more appropriate in this case:

5	**11**	**4**	**6**	**3**	**10**	**5**	**7**	**9**	**6**
6	**2**	**3**	**7**	**8**	**6**	**4**	**3**	**6**	**5**
9	**5**	**6**	**4**	**7**	**1**	**5**	**8**	**6**	**2**
6	**4**	**8**	**7**	**5**	**12**	**4**	**7**	**10**	**6**
7	**8**	**3**	**6**	**7**	**5**	**5**	**8**	**6**	**4**
6	**11**	**5**	**2**	**6**	**9**	**7**	**3**	**7**	**5**

Answer:

From the above survey result we have to form a frequency distribution and hence the mean and median. Detail computations are as follows:

Number of Members (X)	Frequency f	Xf	c.f.	
1	1	1	1	
2	3	6	4	
3	5	15	9	
4	6	24	15	
5	10	50	25	
6	13	78	38	median Class
7	9	63	47	
8	5	40	52	
9	3	27	55	
10	2	20	57	
11	2	22	59	
12	1	12	60	
Total	60	358		
Mean	358/60 =	5.97		
Median	(N+1)/2 th term = 30.5th			
	term .			

The cumulative frequency just greater than 30.5 is 38 and the value of X corresponding to 38, the median class is 6. Hence the median size is 6 members per family which is an appropriate method because a fractional value given by mean does not indicate the average number of members in a family.

13. The following is the table which gives you the distribution of marks secured by some students in an examination

Marks	0-20	21-30	31-40	41-50	51-60	61-70	71-80
No of Students	40	48	125	79	40	36	32

Find (i) Median marks; (ii) the percentage of failure if the minimum for pass is 35 marks

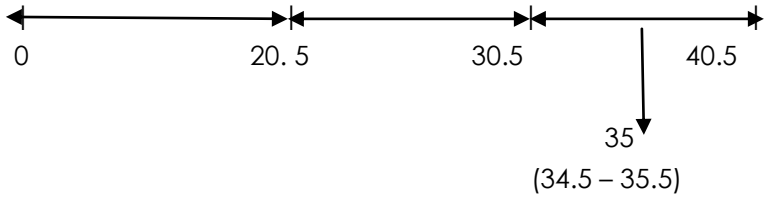
Answer:

Detail computation of the median is in the following table:

Marks	No of Student (f)	c.f.	
0-20	40	40	
21-30	48	88	
31-40	125	213	Median Class
41-50	79	292	
51-60	40	332	
61-70	36	368	
71-80	32	400	
Median	30.5+(200-88)*10/125 =		39.46

For computation of mean we have used the formula $M_d = l_1 + \frac{\frac{N}{2} - c}{f} * h$ where l_1 is the real lower limit of the median class, N is total number of students, c is cumulative frequency just previous to the median class, f is the simple frequency of the median class &h is width of the median class

ii) Minimum pass mark is 35. So students below 35 are failing. Now 35 is in the class 31-40. With continuous distribution the requirements is presented below:



So we have to calculate total number of students (cumulative frequency) up to 35---assuming 34.5 is rounded to the nearest whole number 35.
Let us assume c.f. less than 34.5 is F

So we have following distribution

Class	c.f
20.5 to 30.5	88
30.5 to 34.5	F
30.5 to 40.5	213

So we can write from 30.5 to 40.5 (a difference of 10) simple frequency is (213-88) = 125
From 30.5 to 34.5 difference is 4
So for a difference of 10 simple frequency is 125

Then for a difference of 4 simple frequency is $\frac{125}{10}$*4

And from the above diagram we can write $\frac{125}{10} * 4 = F - 88$

Therefore F= No of students up to 35 = $\frac{125}{10}$*4+88 = 138

The percentage of failure = $\frac{138}{400}$*100 = 34.5%

14. Amend the following table and locate the median from the amended table:

Size	Frequencies
10-15	**10**
16-17.5	**15**
17.5-20	**17**
20-30	**25**
30-35	**28**
35-40	**30**
40 & onward	**40**

Answer:

The series can be amended in different forms. One of these methods being its rewriting 10-20 as class interval. So the amended table is

Size	Frequencies
10-20	(10+15+17) = 42
20-30	25
30-40	(28+30) =58
40-onwards	40

Detail computation of median is given below:

Size	Frequency	c.f	
10-20	42	42	
20-30	25	67	
30-40	58	125	Median Class
40-onward	40	165	
Median Class = 165/2 = 82.5 th item			
Median =	30+(82.5-67)*10/58 =		32.67

For computation of mean we have used the formula $M_d = l_1 + \frac{\frac{N}{2} - c}{f} * h$ where l_1 is the real lower limit of the median class, N is total number of students, c is cumulative frequency just previous to the median class, f is the simple frequency of the median class & h is width of the median class.

15. Calculate the values of (i) median, (ii) the two quartiles for the data given below:

Income in ('000 ₹)	Number of persons
Under 1	15
1-2	90
2-3	81
3-5	115
5-10	64
10-25	25
25-50	6
50-100	2
100-1000	5

Answer:

Detail computation is given below

Income	Number of persons	C.f.	
Under 1	15	15	
1-2	90	105	1st Quartile class
2-3	81	186	
3-5	115	301	Median Class
5-10	64	365	2nd Quartile class
10-25	25	390	
25-50	6	396	
50-100	2	398	
100-1000	5	403	
Median class =	407/2 = 203.5th item		
Median =	3+(203.5-186)/115*2 =		3.30
First Quartile class = 407/4 = 101.75th item			
Q1 = 1+(101.75-15)/90*1 =			1.96
Second Quartile class = 407*3/4 = 305.25th item			
Q2= 5+(305.25-301)/64*5=			5.33

16. Following frequency distribution is with regard to weight in grams of mangoes of a given variety. If mangoes of weight less than 443 grams be considered unsuitable for foreign market, what is the percentage of the total yield suitable for it? Assume the given frequency distribution to be typical of the variety:

Weight in grams	Frequency
410-419	**14**
420-429	**20**
430-439	**40**
440-449	**54**
450-459	**47**
460-469	**18**
470-479	**7**

Answer:

Cumulative frequency distribution is shown below:

Weight In grams	Frequency	c.f
410-419	14	14
420- 429	20	34
430-439	40	74
440-449	54	128
450-459	47	175
460-469	18	193
470-479	7	200

So we have to calculate total yield (cumulative frequency) up to 443. It is in the class 440-449
Let us assume c.f. less than 443 is F

So we have following distribution

Class	c.f
430 to 439	74
440 to 443	F
440 to 449	128

So we can write from 439.5 to 449.5 (a difference of 10) simple frequency is (128-74) = 54
From 439.5 to 442.5 difference is 3

So for a difference of 10 simple frequency is 54
Then for a difference of 4 simple frequency is $\frac{54}{10}$*3
And from the above diagram we can write $\frac{54}{10} * 3 = \mathrm{F} - 74$

Therefore, F= $\frac{54}{10}$*3+74 = 90.2 ≅ 90

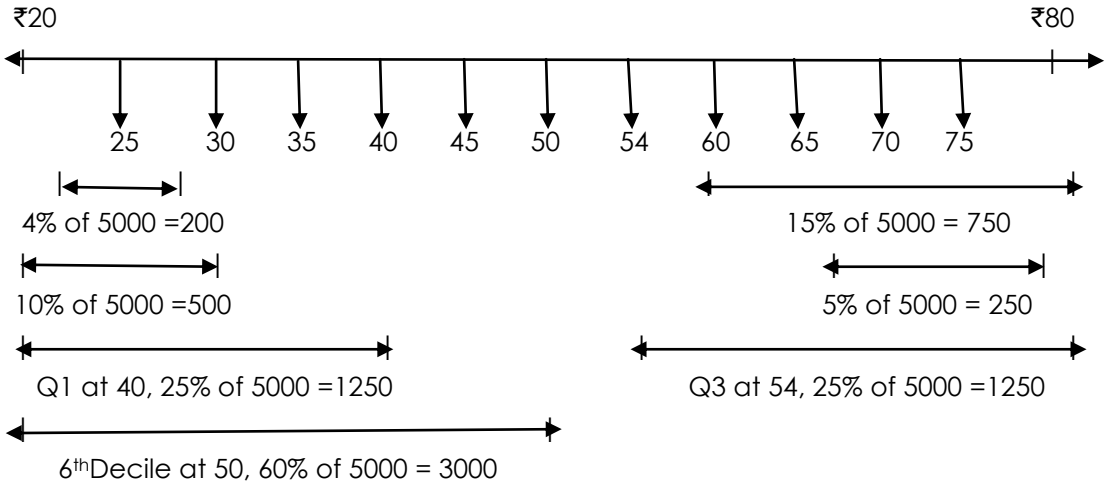
The percentage of unsuitable mangoes = $\frac{90}{400}$*100 = 45%

Hence percentage of mangoes suitable for foreign market = 100 – 45 = 55%

17. For a group of 5000 workers, the weekly wages vary from ₹20 to ₹80. The wage of 4 percent of the workers are under ₹25 and those of 10 percent are under ₹30. 15 percent of the workers earn ₹60 and over and 5 percent of them get ₹70 and over. The quartile wages are ₹40 and ₹55 and the sixth decile is ₹50. Put this information in the form of a frequency table.

Answer:

This type of problem can be best understood by drawing the following type of diagrams



From the above diagram we can write

20 to 25 = 200

20 to 30 = 500. So 25 to 30 = 500 – 200 = 300. **So 20 to 30 = 500**

So 30 to 40 = 1250 – 500 = 750

So 40 to 50 = 3000 – 1250 = 1750

50 to 80 = 5000 -3000 = 2000

54 to 80 = 1250

50 to 54 = 2000 -1250 = 750

54 to 80 = 1250

60 to 80 = 750

So 54 to 60 = 1250 – 750 = 500. **& 50 to 60 = 750+500 = 1250**

60 to 80 = 750

70 to 80 = 250

So 60 to 70 = 750 -250 = 500

Therefore frequency distribution of workers:

Weekly Wages (₹)	No of workers (f)
20-30	500
30-40	750
40-50	1750
50-60	1250
60-70	500
70-80	250

18. The following are the marks obtained by 50 students in statistics.

Marks	Number of Students
10 marks or less	**4**
20 marks or less	**10**
30 marks or less	**30**
40 marks or less	**40**
50 marks or less	**47**
60 marks or less	**50**

Compute (i) Range of marks obtained by middle 80% of the students; (ii) The median

Answer:

From the cumulative distribution given in the question the frequency distribution is given below Middle 80% students lie between 1st decile and 9th decile. So we have to calculate $d_9 - d_1$. The detail computation is shown below:

Marks	Frequency (f)	c.f	
0-10	4	4	
10-20	6	10	1st decile
20-30	20	30	Median
30-40	10	40	
40-50	7	47	9th decile
50-60	3	50	
1st decile=	50/10 =	5th item	
1st decile=	10+(5-4)/6*10	11.67	
9th decile =	9*50/10 =	45th item	
9th decile =	40+ (45-40)/7*10	47.14	
Median Class	50/2 =	25th item	
Median =	20+(25-10)/20*10	27.5	

So range of marks obtained by the middle 80% of the students = $47.14 - 11.67 = 35.47$

19. The following table gives the distribution of monthly income of 600 families in a certain city:

Monthly Income (₹)	No of Families
Below 75	60
75-150	170
150-225	200
225-300	60
300-375	50
375-450	40
450 & over	20

(i) Compute the median income and obtain the limits of income of central 50% of the observed frequency

(ii) If a tax is to be collected from those families having income exceeding ₹350 what percentage of families will be asked to pay tax?

Answer:

From the given question frequency distribution, c.f. and other detail computation are given below:

Monthly Income	No of families	c.f.	
Below 75	60	60	
75- 150	170	230	Q1 class
150-225	200	430	Median class
225-300	60	490	Q3 class
300-375	50	540	
375-450	40	580	
450 & over	20	600	
1st Quartile class	600/4 =	150 th item	
Q1 = 75+(150-60)/170*75 =		114.706	
2nd Quartile class 3*600/4 =		450 th item	
Q3 = 225+(450-430)/60*75 =		250	
Median class =	600/2 =	300 th item	
Median = 150+(300-230)/200*75 =		176.25	

Limits of income of central 50% of the observed frequency correspond to Q1 and Q3 and limits are 114.706 & 250.

No of families less than 350 by simple interpolation = 490 + 50 * (540-490)/75 = 523(approx.)

So no of tax paying families = 600 -523 = 77.

%age of families paying tax = (77/600)*100 = 12.83%

20. Modal marks for a group of 94 students is 54. Ten students got marks between 0-20, 30 students got marks between 40-60 & 14 students got marks between 80-100. Maximum marks in the test were 100. Find out the number of students getting between 20-40 and 60-80

Answer:

Let number of students between 20-40 and 60-80 are x and y respectively. So the following frequency distribution could be constructed

Class Interval	Frequency	
0-20	10	
20-40	x	
40-60	30	Modal class
60-80	y	
80-100	14	

So by question we can write

$54 = 40 + \frac{30-x}{(30-x)+(30-y)} * 20$ ..... (1) [As mode of the distribution 54]

& 10+x+30+y+14 =94 or x + y = 40..... (2) [As total no of students is 94]

From (1) we can write $54 - 40 = \frac{30-x}{60-(x+y)} * 20$

Or, $14 = \frac{30-x}{60-40} * 20$

Or, 14 = 30 – x

Or, x = 30 – 14 = 16

So y = 40 – 16 = 24

21. You are given the following incomplete frequency distribution. It is known that the total frequency is 1000 and that the median is 413.11. Estimate by calculation the missing frequencies and find the value of the mode.

Value (X)	Frequency (f)
300-325	**5**
325-350	**17**
350-375	**80**
375-400	**?**
400-425	**326**
425-450	**?**
450-475	**88**
475-500	**9**

Answer:

Let the missing frequencies are x and y and accordingly the cumulative frequency distribution is given below:

Value	Frequency	c.f	
300-325	5	5	
325-350	17	22	
350-375	80	102	
375-400	x	102+x	
400-425	326	428+x	Median class
425-450	y	428+x+y	
450-475	88	516+x+y	
475-500	9	525+x+y	

Median is 413.11

So we can write

$413.11 = 400 + \frac{500-102-x}{326} * 25$..... (1) [From median formula]

$\&525 + x + y = 1000$..... (2)

From (2), we could write $x + y = 475$

From (1) we can write $413.11 - 400 = \frac{398-x}{326} * 25$

Or, $13.11 * 326 = (398 - x) * 25$

Or, $\frac{13.11*326}{25} = 398 - x$

Or, $x = 398 - \frac{13.11*326}{25}$

Or, $x = 227$

So $y = 475 - 227 = 248$

Modal class is 400-425

So Mode = $400 + \frac{326-227}{(326-227)+(326-248)} * 25$

Or, Mode = 414

22. A frequency distribution of profits earned by different firms are given below:

Profits (in'000 rupees)	No of firms
20-40	**6**
40-60	**9**
60-80	**11**
80-100	**14**
100-120	**20**
120-140	**15**
140-160	**10**
160-180	**8**
180-200	**7**

The mean, median and mode of the above distributions were calculated as 110,110 and 110.91repectively but while calculating them the frequencies of the first and last classes were interchanged by mistake. Find the corrected measures of central tendency.

Answer:

Mode remains the same as it is not affected by interchange of the extreme frequencies.
Mean & Median is affected.

With incorrect data given the computation of mean table is as follows:

Profits (in'000 rupees)	mid value(x)	No of firms(f)	fx
20-40	30	6	180
40-60	50	9	450
60-80	70	11	770
80-100	90	14	1260
100-120	110	20	2200
120-140	130	15	1950
140-160	150	10	1500
160-180	170	8	1360
180-200	190	7	1330
Total		100	11000

The corrected $\Sigma fx = 11000 - 180 - 1330 + 30 * 7 + 190 * 6 = 10840$
Σfremain the same
So correct mean = $\frac{10840}{100} = 108.4$

With incorrect data for computation of median the cumulative frequency table will be as follows:

Profits (in'000 rupees)	No of firms(f)	c.f	
20-40	6	6	
40-60	9	15	
60-80	11	26	
80-100	14	40	
100-120	20	60	Median Class
120-140	15	75	
140-160	10	85	
160-180	8	93	
180-200	7	100	

So corrected Median is $= 100 + \frac{50-(40+7-6)}{20} * 20 = 100 + 50 - 41 = 109$

23. You are given two variables A and B. using quartile deviation state which is more variable.

A							
Mid-point	**15**	**20**	**25**	**30**	**35**	**40**	**45**
Frequency	**15**	**33**	**56**	**103**	**40**	**32**	**10**

B								
Mid-point	**100**	**150**	**200**	**200**	**300**	**350**	**400**	**450**
Frequency	**340**	**492**	**890**	**1420**	**620**	**360**	**187**	**140**

Answer:

The detail computation of Quartiles for the two series are as follows:

	Series A		
Class Interval	Frequency	c.f	
12.5-17.5	15	15	
17.5-22.5	33	48	
22.5-27.5	56	104	Q1 class
27.5-32.5	103	207	
32.5-37.5	40	247	Q2 class
37.5-42.5	32	279	
42.5-47.5	10	289	
Q1 class =	289/4 =	72.25	item
Q2 class =	289*3/4	216.75	item
Q1	22.5+(72.25-48)*5/56	24.665	
Q2	32.5+(216.75-207)*5/40	33.719	

	Series B		
Class Interval	Frequency	c.f	
75-125	340	340	
125-175	492	832	
175-225	890	1722	Q1 class
225-275	1420	3142	
275-325	620	3762	Q2 class
325-375	360	4122	
375-425	187	4309	
425-475	140	4449	
Q1 class =	4449/4=	1112.25	item
Q2 class =	4449*3/4=	3336.75	item
Q1	175+(1112.25-832)*50/890	190.744	
Q2	275+(3336.75-3142)*50/620	290.706	

So Quartile deviation of Series A = $\frac{33.719-24.665}{2} = 4.53$

So Quartile deviation of Series B = $\frac{290.706-190.744}{2} = 49.98$

Therefore Series B is more variable

24. Calculate Mean deviation from median from the following data:

Sales (in '000 ₹)	1-3	3-5	5-7	7-9	9-11	11-13	13-15	15-17
Number of shops	6	53	85	56	21	26	4	4

Answer:

The detail computation is shown below in the two tables:

Sales(in'000 rupees)	Number of shops (f)	c.f	
1-3	6	6	
3-5	53	59	
5-7	85	144	medain class
7-9	56	200	
9-11	21	221	
11-13	26	247	
13-15	4	251	
15-17	4	255	
Median Class	255/2	127.5	item
Median (Md)	5+(127.5-59)*2/85	6.612	

Sales(in'000 rupees)	Number of shops (f)	mid values(m)	d= |m-Md|	fd
1-3	6	2	4.612	27.67
3-5	53	4	2.612	138.42
5-7	85	6	0.612	52.00
7-9	56	8	1.388	77.74
9-11	21	10	3.388	71.15
11-13	26	12	5.388	140.09
13-15	4	14	7.388	29.55
15-17	4	16	9.388	37.55
Total	255			574.19
Mean Deviation=	574.19/255	2.252		

25. For two groups of observations the following data is available:

Group I	Group II
$\Sigma(X-5)=3$	$\Sigma(X-8)=-11$
$\Sigma(X-5)^2=43$	$\Sigma(X-8)^2=76$
$n=18$	$n=17$

Find the mean and the standard deviation of the 35 observations obtained by combining the two groups.

Answer:

Group A:

$\Sigma(X-5)=3$

So for 18 items we could write $\Sigma X-18*5=3.$ So $\Sigma X=93$

Similarly

$\Sigma(X-5)^2=43$

So for 18 items we could write $\Sigma X^2-10\Sigma X+\Sigma 25=43$

Or, $\Sigma X^2-10*93+18*25=43$

Or, $\Sigma X^2=43+930-450=523$

Group B:

$\Sigma(X-8)=-11$

So for 17 items we could write $\Sigma X-17*8=-11.$ So $\Sigma X=125$

Similarly

$\Sigma(X-8)^2=76$

So for 17items we could write $\Sigma X^2-16\Sigma X+\Sigma 64=76$

Or, $\Sigma X^2-16*125+17*64=76$

Or, $\Sigma X^2=76+2000-1088=988$

Therefore we can write for combined group

$$\Sigma X=93+125=218, n=18+17=35 \;\&\; \Sigma X^2=523+988=1511$$

So combined mean = $\frac{218}{35}=6.23$

Combined s.d. = $\sqrt{\frac{\Sigma X^2}{35}-(\frac{\Sigma X}{35})^2}=\sqrt{\frac{1511}{35}-(\frac{218}{35})^2}$ = 2.092

26. Calculate standard deviation from the following data regarding marks obtained by students in an accountancy test:

Marks	1	2	3	4	5	6	7	8	9
No of students	32	41	57	98	123	83	46	17	3

What will be the value of standard deviation if the marks obtained by each of the students are increased by one?

Answer:

We have found out s.d. by $\sqrt{\frac{\Sigma fx^2}{\Sigma f} - (\frac{\Sigma fx}{\Sigma f})^2}$, where x is the marks and $f$is the corresponding frequency

Marks (x)	No of Students(f)	fx	x^2	fx^2
1	32	32	1	32
2	41	82	4	164
3	57	171	9	513
4	98	392	16	1568
5	123	615	25	3075
6	83	498	36	2988
7	46	322	49	2254
8	17	136	64	1088
9	3	27	81	243
Total	500	2275		11925
Var	11925/500-(2275/500)^2			3.1475
s.d.	SQRT(Var)			1.774

Since standard deviation is independent of the change of origin the results obtained above would not be affected if the marks obtained by each of the students are increased by one.

27. A manufacturer of collars supplies you the following data regarding the neck circumferences of the students of the Delhi University

Neck Circumference (in inches)	12.0	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0
Frequency	5	20	30	43	60	56	37	16	3

Calculate the standard deviations and advise the manufacturer as to the largest and the smallest size of collars, he should make in order to meet the needs of most of his customers. Use formula ($\bar{X} \pm \sigma$)

Answer:

Here we have found out s.d. as $\sqrt{\frac{\Sigma fd^2}{\Sigma f} - (\frac{\Sigma fd}{\Sigma f})^2}$ where $d = x - A$, where A is an arbitrary value. Here we have taken $A = 14$, f is the frequency

The detail computation is as follows:

Neck Circumfrence(inch) (x)	frequency (f)	d= x-14	fd	fd^2
12	5	-2	-10	20
12.5	20	-1.5	-30	45
13	30	-1	-30	30
13.5	43	-0.5	-21.5	10.75
14	60	0	0	0
14.5	56	0.5	28	14
15	37	1	37	37
15.5	16	1.5	24	36
16	3	2	6	12
Total	270		3.5	204.75
S.d.	SQRT((204.75/270-(3.5/270)^2)			0.8707

The mean of the distribution = $A + \frac{\Sigma fd}{\Sigma f} = 14 + \frac{3.5}{270} = 14.013$

Using the criterion $\overline{X} \pm \sigma = 14.013 \pm 0.8707 = 14.8837 \text{ or } 13.1423$
The largest size of the collar 14.8837& the smallest size of the collar 13.1423

28. A charitable organization decided to give old age pension to people over sixty years of age. The scales of pension were fixed as follows:

Age group 60-65	**₹2,000 per month**
Age group 65-70	**₹2,500 per month**
Age group 70-75	**₹3,000 per month**
Age group 75-80	**₹3,500 per month**
Age group 80-85	**₹4,000 per month**

The ages of 25 persons who secured the pension right are as given below:

74	**62**	**84**	**72**	**61**	**83**	**72**	**81**	**64**	**71**	**63**	**61**	**60**
67	**74**	**64**	**79**	**73**	**75**	**76**	**69**	**68**	**78**	**66**	**67**	

Calculate the monthly average pension payable per person and the standard deviation.

Answer:

The frequency distribution of age groups is as follows

Age Group	frequency
60-65	7
65-70	5
70-75	6
75-80	4
80-85	3

The details computation of average and s.d. are as follows:

Age Group	frequency(f)	pension (x)	d'=(x-3000)/1000	fd'	fd'^2
60-65	7	2000	-1	-7	7
65-70	5	2500	-0.5	-2.5	1.25
70-75	6	3000	0	0	0
75-80	4	3500	0.5	2	1
80-85	3	4000	1	3	3
Total	25			-4.5	12.25
Average	3000-4.5/25*1000		2820		
S.d	SQRT(12.25/25-(-4.5/25)2)*1000			676.5	

Herewe have used the following formula

Mean = $A + \frac{\Sigma fd'}{\Sigma f} * h$

s.d.=$\sqrt{\frac{\Sigma fd'^2}{\Sigma f} - (\frac{\Sigma fd'}{\Sigma f})^2}$

where$h = 1000, A = 3000, f = \text{frequency}$

29. Calculate arithmetic mean and standard deviation from the following data:

Value	>800	>700	>600	>500	>400	>300	>200	>100
Frequency	14	44	96	175	381	527	615	660

Answer:

The above information let us arrange in the following frequency distribution:

We have made each class interval in steps of 100. By question more than 800, there are 14 items. Whereas more than700 there are 44 items. So from 700 to 800 there will be 44-14 =30 items.

More than	Class intrval	c.f.	s.f (f)
100	100-200	660	45
200	200-300	615	88
300	300-400	527	146
400	400-500	381	206
500	500-600	175	79
600	600-700	96	52
700	700-800	44	30
800	800-900	14	14

Now the detail computation is given below:

More than	Class intrval	c.f.	s.f (f)	mid value(m)	d'= (m-450)/100	fd'	fd'^2
100	100-200	660	45	150	-3	-135	405
200	200-300	615	88	250	-2	-176	352
300	300-400	527	146	350	-1	-146	146
400	400-500	381	206	450	0	0	0
500	500-600	175	79	550	1	79	79
600	600-700	96	52	650	2	104	208
700	700-800	44	30	750	3	90	270
800	800-900	14	14	850	4	56	224
Total			660			-128	1684
Mean	450-(128/660)*100				430.606		
s.d.	SQRT(1684/660 - (-128/660)^2)*100				158.553		

30. Two automatic filling machines A and B are used to fill tea in 500 grams cartons. A random sample of 100 cartons on each machine showed the following:

Tea Contents (grams)	Machine A	Machine B
485-490	**12**	**10**
490-495	**18**	**15**
495-500	**20**	**24**
500-505	**22**	**20**
505-510	**24**	**18**
510-515	**4**	**13**

Comment on the performance of the two machines on the basis of average filling and dispersion

Answer:

The detail computation of mean, s.d and coefficient of variation of machine A and B are given below.

		A			
Tea	Machine	mid value	d'= (x-497.5)/5		
contents(gm)	(f)	x		fd'	fd'^2
485-490	12	487.5	-2	-24	48
490-495	18	492.5	-1	-18	18
495-500	20	497.5	0	0	0
500-505	22	502.5	1	22	22
505-510	24	507.5	2	48	96
510-515	4	512.5	3	12	36
Total	100			40	220
Mean		497.5+40/100*5		499.5	
s.d.	SQRT(220/100-(40/100)^2)*5			7.141	
Coeeficient					
of variation=	(s.d/mean)*100			1.43	

		B			
Tea	Machine	mid value	d'= (x-497.5)/5		
contents(gm)	(f)	x		fd'	fd'^2
485-490	10	487.5	-2	-20	40
490-495	15	492.5	-1	-15	15
495-500	24	497.5	0	0	0
500-505	20	502.5	1	20	20
505-510	18	507.5	2	36	72
510-515	13	512.5	3	39	117
Total	100			60	264
Mean		497.5+60/100*5		500.5	
s.d.	SQRT(264/100-(60/100)^2)*5			7.550	
Coeeficient					
of variation=	(s.d/mean)*100			1.51	

Performance of Machine B is more variable as its coefficient of variation is more than that of A.

31. The following distribution was obtained by a change of origin and scale:

dx	-4	-3	-2	-1	0	1	2	3	4
f	4	8	14	18	20	14	10	6	6

Write down the frequency distributions of X if it is given that mean and variance are 59.5 and 413 respectively.

Answer:

Calculation for Am, & s.d. are from the following table:

dx	f	fdx	fdx^2
-4	4	-16	64
-3	8	-24	72
-2	14	-28	56
-1	18	-18	18
0	20	0	0
1	14	14	14
2	10	20	40
3	6	18	54
4	6	24	96
Total	100	-10	414

Now by question we can write

Mean = 59.5 = $A + \frac{-10}{100} * h$…….. (1)

variance = 413 = $[\frac{414}{100} - \left(\frac{-10}{100}\right)^2]h^2$…….. (2)

From (2) we can write =$413 = [(4.14 - 0.01)] * h^2$

Or, $\left(\frac{413}{4.13}\right) = h^2$

Or, $h = 10$

Putting this value in (1) we get

$$59.5 = A - 0.1 * 10$$

Or, $A = 60.5$

Now we could write $\frac{x_1+x_2}{2} = 60.5$ &$x_2 - x_1$=10, where x_1&$x_2$are class boundaries of the class whose mid value is 60.5

So we can write $x_2 + x_1 = 121$

Or, we can write $2x_2 = 131, \text{or } x_2 = 65.5$

So $x_1 = 65.5 - 10 = 55.5$

Therefore the class for which dx = 0 is 55.5 – 65.5

Class previous to this will be then 45.5 – 55.5 & class next to this will be 65.5 – 75.5

Hence the original frequency distribution will be as follows:

dx	Mid values	Class Interval	frequency
-4	20.5	15.5-25.5	4
-3	30.5	25.5-35.5	8
-2	40.5	35.5-45.5	14
-1	50.5	45.5-55.5	18
0	60.5	55.5-65.5	20
1	70.5	65.5-75.5	14
2	80.5	75.5-85.5	10
3	90.5	85.5-95.5	6
4	100.5	95.5-105.5	6

32. The analysis of the results of a budget survey of 150 families gave an average monthly expenditure of ₹120 on food items with a standard deviation of ₹15. After the analysis was completed it was noted that figure recorded for one house hold was wrongly taken as ₹15 instead of ₹105. Determine the correct value of the average expenditure and its standard deviation.

Answer:

Average expenditure = ₹120. No of families = 150
Total Expenditure = 120*150 = ₹18000
Actual Expenditure after correction = 18000-15+105= 18090
So the correct average monthly expenditure = $\frac{18090}{150}$ = ₹120.6

By question we can write

$15^2 = \frac{\sum x^2}{n} - (\frac{\sum x}{n})^2$

Or, $225 = \frac{\sum x^2}{n} - (120)^2$

Or, $225 + 14400 = \frac{\sum x^2}{150}$

Or, $150*14625 = \sum X^2$

Or, $\sum X^2 = 2193750$

Corrected $\sum X^2 - 15 * 15 + 105 * 105 = 2204550$

So Corrected variance = $\frac{2204550}{150} - (120.6)^2 = 152.64$

So corrected s.d. = 12.35

33. For a frequency distribution of marks in statistics of 200 candidates (grouped in intervals of 0-5, 5-10 etc.) the mean and standard deviation were found to be 40 and 15. Later it was discovered that the score 43 was misread as 53 in obtaining the frequency distribution. Find the corrected mean and standard deviation corresponding to the corrected frequency distribution

Answer:

We are given $N = 200, \overline{X} = 40, \sigma = 15$
Uncorrected $\sum X = N * \overline{X} = 40 * 200 = 8000$
Now $\sigma^2 = \frac{1}{N}\sum x^2 - (\overline{x})^2$
Or $\sum x^2 = N\{\sigma^2 + (\overline{x})^2\}$
So uncorrected $\sum x^2 = 200\{15^2 + 1600\} = 365000$

Also incorrect score 53 falls in the group of 50-55, the mid value of which is 52.5. Similarly correct score 43 falls in the group of 40-45, the mid value of which is 42.5
So corrected $\sum X = 8000 - 52.5 + 42.5 = 7990$
So corrected mean = $\frac{7990}{200} = 39.95$
Corrected $\sum x^2 = 365000 - 52.5 * 52.5 + 42.5 * 42.5 = 364050$
So corrected $\sigma^2 = \frac{1}{N}\sum x^2 - (\overline{x})^2 = \frac{364050}{200} - (39.95)^2$
SO corrected $\sigma = 14.97$

34. For a group containing 100 observations the arithmetic mean and standard deviation are 8 and $\sqrt{10.5}$. For 50 observations selected from these 100 observations the mean and the standard deviation are 10 and 2 respectively. Find the arithmetic mean and standard deviation of the other half.

Answer:

Formula for mean and s.d of combined group are
$(n_1 + n_2)\overline{X}_{12} = n_1\overline{X}_1 + n_2\overline{X}_2$.....(1)
And $(n_1 + n_2) * {\sigma_{12}}^2 = (n_1 * {\sigma_1}^2 + n_2 * {\sigma_2}^2) + (n_1 * {d_1}^2 + n_2 * {d_2}^2)$.....(2)

Substituting the values in (1) we get
100*8 = 50*10+50*$\overline{X}_2$
So $\overline{X}_2 = 6$
$d_1 = \overline{X}_{12} - \overline{X}_1 = 8 - 10 = -2$
$d_{12} = \overline{X}_{12} - \overline{X}_2 = 8 - 6 = 2$

From (2) we can write
$$100 * 10.5 = \{50 * 2^2 + 50{\sigma_2}^2\} + \{50 * (-2)^2 + 50 * 2^2\}$$
Solving this we are getting $\sigma_2 = 3$

Here the mean and standard deviation of the remaining 50 observations are 6 and 3 respectively.

35. The following is the record f number of bricks laid each day for 10 days by two brick-layers A and B. Calculate the coefficient of variation in each case and discuss the relative consistency of the two brick layers:

A	**700**	**675**	**725**	**625**	**650**	**700**	**650**	**700**	**600**	**650**
B	**550**	**600**	**575**	**550**	**650**	**600**	**550**	**525**	**625**	**600**

If the figures for A were in every case 10 more and those of B in every case 20 more than the figures given above, how would the answer be affected.

Answer:

The detail computation for two series are given below:

A				B				
X	d'= (X-700)/25	d'^2		X	d'= (X-625)/25	d'^2		
700	0	0		550	-3	9		
675	-1	1		600	-1	1		
725	1	1		575	-2	4		
625	-3	9		550	-3	9		
650	-2	4		650	1	1		
700	0	0		600	-1	1		
650	-2	4		550	-3	9		
700	0	0		525	-4	16		
600	-4	16		625	0	0		
650	-2	4		600	-1	1		
Total	-13	39		Total	-17	51		
Mean=	700+(-13)/10*25	667.5			625+(-17)/10*25		582.5	
s.d	SQRT(39/10-(-13/10)^2)*25		37.16517		SQRT(51/10-(-17/10)2)*25			37.16517
c.v	(37.16517/667.5)*100		5.567816		(37.16517/582.5)*100			6.380287

As the coefficient of variation for brick layer A is less than that of B, Brick layer A is more consistent
If the figures for A in every case were 10 more and that of B 20 more, the arithmetic mean of A will increase by 10 and that of B by 20. But the s.d. of both will remain unchanged.
So AM of A will be 667.5+10 = 677.5 bricks per
AM of B will be 582.5+20 = 602.5 bricks per day
So new coefficient of variation of A = (37.16/677.5)*100 = 5.49%
New coefficient of variation of B = (37.17/602.5)*100 = 6.17%
Brick layer A is more consistent.

36. Indicate which brand you will choose and why:

Brand	Mean	Standard Deviation
A	16000	2000
B	20000	4000

Also indicate what percentage of brand B might be expected to run more than 24000 km.

Answer:

If more reliance is desired the brand A shall be preferred to brand B, because of lower coefficient of variation as shown below:

$$\text{C.V}_\text{A} = \frac{2000}{16000} * 100 = 12.5\%$$

$$\text{C.V}_\text{B} = \frac{4000}{20000} * 100 = 20\%$$

The $\overline{X} \pm \sigma$ gives a range of 24000 to 16000 in the case of brand B.

III. State the following statements are true or false

1. **If one or two data items in a sample or population have either extremely large or extremely small values then arithmetic mean is the appropriate average to represent the data.**
2. **Median can be computed for ordinal level data whereas Mode can be computed for all levels of data – nominal, ordinal, interval and ratio.**
3. **Like Mean Mode has the disadvantage of being affected by extremely high or low values.**
4. **For many sets of data, there is no mode because no value appears more than once.**
5. **The sum of squares of deviations of a set of observations is the minimum when deviations are taken from the arithmetic average.**
6. **Arithmetic average of a data set gives weight to all items in the set in direct proportion to their size.**
7. **Arithmetic average can be determined easily by inspection or by graphically**
8. **Arithmetic average of a data set is called a fictitious average since it may not correspond with a single item in the series.**
9. **For an extremely asymmetrical (skewed) distributions usually arithmetic mean is the most suitable measure of location.**
10. **The geometric mean of a series containing N observations is the (N-1)th root of the product of the values.**
11. **Arithmetic mean has a bias for lower values whereas geometric mean has a bias for higher values.**
12. **To study the behavior of prices of shares in a stock market study of averages is more important**
13. **In a frequency distribution a given fraction or proportion of the data lie at or below a fractile**
14. **Standard deviation is an absolute measure of dispersion whereas Coefficient of variation is a relative measure of dispersion.**

15. **Absolute measures of dispersion are used for comparing dispersion of two or more distributions only when distributions are given in same units.**
16. **Mean and standard deviation have the same units as the original data but the skewness has no units, it's a pure number.**
17. **If skewness is positive, the data are positively skewed or skewed right, meaning that the right tail of the distribution is longer than the left**

Answers:

1/(F), 2/(T), 3/(F),4/(T), 5/(T), 6/(T), 7/(F), 8/(T), 9/(F), 10/(F), 11/(F), 12/(F),13/(T), 14/(T), 15/(T), 16/(T),17/(T)

Study Note – 5

Correlation and Regression

Learning Objective:

After studying this chapter, the student will be able to:

- ***The concept of correlation between two variables and quantitative measurement of correlation,***
- ***The concept of regression, and its application in estimation of a variable from known set of data,***

1. Choose the correct Answer:

(i) Karl Pearson's coefficient of correlation between two variables x and y is 0.52, their covariance is 7.8. If the variance of x is 16, Then the standard deviation of the variable y is:

(a).4.5 (b).3.5 (c).4.75 (d).3.75

Answer: (d).

Variance of x = 16. Then S.D. of x = $\sigma_x = \sqrt{16} = 4$.

We Know that correlation coefficient of the variables x and y = $r_{xy} = \frac{cov\ (x,y)}{\sigma_x \times \sigma_y}$

=> 0.52= $\frac{7.8}{4\times\sigma_y}$ => 4× σ_y (0.52) = 7.8 => $\sigma_y = \frac{7.8}{4\times 0.52}$= 3.75

∴ S.D of y = σ_y= 3.75.

(ii) If b_{xy} = - 0.4 and b_{yx} = - 0.9, then r_{xy} is

(a).- 0.6 (b).0.6 (c).0.48 (d).None of these

Answer: (a).

Since $r_{xy}^2 = b_{xy} \times b_{yx}$. We get, r_{xy}^2 = (- 0.4)× (−0.9) => r_{xy}^2 = 0.36

=> r_{xy} = ± 0.6

Since both b_{xy} *and* b_{yx} are negative, we get, r_{xy} is also negative.

∴ r_{xy} = - 0.6.

(iii) If σ_x=10, σ_y=12, b_{xy} = - 0.8, then the value of r is:
(a).0.96 (b). - 0.96 (c). - 0.48 (d). - 0.15

Answer: (b)

We have, b_{xy} = r × $\frac{\sigma_x}{\sigma_y}$ => - 0.8 = r × $\frac{10}{12}$ =>r = - $\frac{9.6}{10}$ = - 0.96

(iv) If cov(x,y) = 12 and σ_y= 5, r between x and y be 0.6, then the value of σ_x is -
(a).2 (b).6 (c).4 (d).5

Answer: (c)

We have, r = $\frac{cov(x,y)}{\sigma_x.\sigma_y}$ =>0.6= $\frac{12}{\sigma_x \times 5}$ => σ_x=$\frac{12}{5\times0.6}$ =$\frac{120}{5\times6}$ = 4.

(v) If $\sum D^2$=33 and N = 10 where D denotes the difference between the ranks of the two series of observations and N is the total number of observation, then the value of the rank correlation Coefficient is :
(a). - 0.6 (b).0.6 (c). - 0.8 (d).0.8.

Answer: (d).

Rank Correlation coefficient = R = 1- $\frac{6\sum D^2}{N\,(N^2-1)}$
∴ R = 1- $\frac{6\times33}{10\,(10^2-1)}$ = 1- $\frac{6\times33}{10\times99}$ = 1-0.2=0.8.

(vi) If the two regression coefficient are b_{xy} = - 0.4 and b_{yx} = - 0.9, then the correlation co efficient (r) with be -
(a).0.6 (b).0.65 (c).- 0.6 (d).- 0.65

Answer: (c).

Here both b_{xy} and b_{yx} are negative and therefore r, the Correlation Coefficient will also be negative.

Now r^2 = b_{xy} × b_{yx} = $(-0.4) \times (-0.9) = 0.36$
∴ $r = \pm\sqrt{0.36} == \pm0.6$, Since r is negative, r = - 0.6.
∴ $The\ required\ correlation\ coefficient\ is - 0.6.$

(vii) The equations of regression lines are 4x - 5y + 33 = 0 and 20x – 9y = 107. Then the mean values of x and y i.e.$\bar{x}$ and $\bar{y}$ are –
(a). $\bar{x} = 13, \bar{y} = 17$ (b). $\bar{x} = 17, \bar{y} = 13$ (c) $\bar{x} = 15, \bar{y} = 12$ (d). $\bar{x} = 19, \bar{y} = 11$

Answer: (a).

The regression lines intersect at $(\bar{x}, \bar{y})$. Thus the regression lines pass through $(\bar{x}, \bar{y})$. So solving the regression equations we get $\bar{x}, \bar{y}$.
$\therefore 4x - 5y = -33 \ldots\ldots (1)$ and 20x – $9y = 107 \ldots\ldots (2)$

Multiplying equation (1) by 9 and equation (2) by 5 and then subtracting, we get,
64 x = 832 =>x = $\frac{832}{64} = 13$

From (1), we get , $4\times 13 - 5y = -33$ => -5y = - 33 – 52 = - 85 =>y = $\frac{-85}{-64} = 17$
$\therefore \bar{x} = x = 13, \bar{y} = y = 17.$
$\therefore Mean\ value\ of\ x = \bar{x} = 13\ and\ that\ of\ y = \bar{y} = 17.$

(viii) The correlation coefficient of bivariate x and y is r = 0.6; variance of x and y are respectively 2.25 and 4.00; $\bar{x}$= 10, $\bar{y} = 20$. The regression equqtion of y on x is –
(a). y = 0.65$x + 12$ (b). y = 0.8$x - 12$ (c). y = 0.8$x + 15$ (d). y = 0.8$x + 12$

Answer: (d).

σ_x^2 = variance of x = 2.25; σ_y^2 = variance of y = 4.00
r = the correlation coefficient of bivirate variables x, y = 0.6.
$\therefore \sigma_x = \sqrt{2.25} = 1.5;\ \sigma_y = \sqrt{4.00} = 2$
$\therefore b_{yx} = r \times \frac{\sigma_y}{\sigma_x} = 0.6 \times \frac{2}{1.5} = 0.8.$
The regression equation of y on x is y - $\bar{y} = b_{yx}\ (x - \bar{x})$ =>y – 20 = 0.8 (x - 10)
=>y = 0.8$x + 12$.

(ix) From the following data: $\sum x = 21,\ \sum y = 20,\ \sum x^2 = 91, \sum xy = 74$, The value of n = 7, the regression equation is:
(a). $y = 0.35\ x + 2.3$ (b).x = 0.5 y – 2.1 (c).y = 0.498 x +1.366 (d). x = 0.5 y + 2.5

Answer: (c)

Here the value of $\sum y^2$ is not given. So the regression equation of x on y is not possible. We have to find the regression equation of y on x.

The regression equation of y on x is, y - $\bar{y} = b_{yx}\ (x - \bar{x})$.....(1)

Here, $\bar{x} = \frac{\sum x}{n} = \frac{21}{7} = 3$; $\bar{y} = \frac{\sum y}{n} = \frac{20}{7} = 2.86$. Now, $b_{yx} = \frac{Cov\ (x,y)}{\sigma_x^2} = \frac{\frac{1}{n}\sum xy - \left(\frac{\sum x}{n}\right)\left(\frac{\sum y}{n}\right)}{\left(\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\right)^2}$

$= \frac{\frac{1}{7}\times 74 - 3\times 2.86}{\frac{91}{7} - (3)^2} = \frac{10.57 - 8.58}{4} = \frac{1.99}{4} = 0.498$

∴The regression equation of y on x is => $y - 2.86 = 0.498\ (x - 3)$

=> y = 0.498x + 1.366.

(x) For some bivarite data, the following results were obtained : The mean value of x =53.2, the mean value of y = 27.9, the regression coefficient of y on x = -1.5 , the regression coefficient of x on y = - 0.2. Then the most probable value of y when x = 60 is-
(a).17.7 (b).18.5 (c).19.5 (d).17.5.

Answer: (a)

Since the value of x is given, we have to find the value of y, we use the regression equation of y on x. This equation is $y - \bar{y} = b_{yx}\ (x - \bar{x})$ => y – 27.9 = - 1.5 (x - 53.2)

=>y = -1.5 x + 27.9 +79.8 => y = -1.5 x + 107.7

Putting x = 60, we get, y = -1.5 × 60 + 107.7 = $-90 + 107.7 = 17.7$

2. Miscellaneous Question:

(i) If x and y be two variables such that Cov (x,y) =12 and the variances of x and y are respectively 25 and 16, then what is r_{xy}?

Answer:

S.D. of x is σ_x and that of y is σ_y.

$\sigma_x = \sqrt{25} = 5\ and\ \sigma_y = \sqrt{16} = 4.$

Hence, $r_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{12}{5\times 4} = \frac{3}{5} = 0.6$

(ii) If $\sigma_x = 10, \sigma_y = 12, b_{xy} = -0.8$, then find r_{xy} .

Answer:

We have $b_{xy} = r_{xy} \times \frac{\sigma_x}{\sigma_y}$, => -0.8 = $r_{xy} \times \frac{10}{12}$ => $r_{xy} = \frac{12}{10} \times (-0.8)$ = - 0.96

(iii) The coefficient of correlation between two variables x and y is 0.28 and their covariance is 7.6. If the variance of x is 9, then find the S.D. of y.

Answer:

We have, Cov(x, y) = 7.6, $\sigma_x^2 = 9 \Rightarrow \sigma_x = 3$ and $r_{xy} = 0.28$.

Now, $r_{xy} = 0.28 = \frac{cov(x,y)}{\sigma_x \sigma_y} \Rightarrow 0.28 = \frac{7.6}{3 \times \sigma_y} \Rightarrow \sigma_y = \frac{7.6}{3 \times 0.28} = 9.0476 = 9.05.$

The required S.D. of y is 9.05.

(iv) The correlation-coefficient between x and y is 0.5. Calculate the correlation-coefficient between u and v where $u = 3x - 11, v = 4y + 7$.

Answer:

Since the correlation-coefficient is independent upon the change of origin and the change of scale, we get correlation-coefficient between u and v $r_{uv} = r_{xy} = 0.5$.

So, the required correlation-coefficient = 0.5.

(v) If $r_{xy} = 0.6$, then find r_{uv} where u = 3x + 5, v = - 4y + 3.

Answer:

Here $r_{xy} = 0.6$ and u = 3x + 5, v = - 4y + 3

Therefore, $r_{uv} = \frac{3 \times (-4)}{|3||-4|} \times r_{xy} = -r_{xy} = -.06 \left[Here\ \sigma_u = 3\sigma_x, \sigma_v = |-4|\sigma_y\right]$

Hence the correlation-coefficient between u and v is – 0.6.

(vi) Find the Spearman's Rank Correlation Coefficient where $\sum D^2 = 45$, where D is the difference between the ranks of the pairs in the variables x and y, and N = 12.

Answer:

Spearman's Rank Correlation Coefficient = R = $1 - \frac{6 \sum D^2}{N(N^2-1)} = 1 - \frac{6 \times 45}{12 \times (144-1)}$

$= 1 - \frac{45}{286} = \frac{241}{286} = 0.8427.$

(vii) If the Spearman's Rank Correlation Coefficient is 0.88 when $\sum D^2 = 20$, where D is the difference between the ranks, then find N, the number of observations in pairs of the variables.

Answer:

We have, R = Rank Correlation Coefficient,

$=> 0.88 = 1 - \frac{6\sum D^2}{N(N^2-1)} = 1 - \frac{6\times 20}{N^3-1} => \frac{120}{N^3-1} = 1 - 0.88 = \frac{12}{100}$

$=> N^3 - N - 1000 = 0$. This is satisfied if $N = 10.035 \approx 10$

Hence the required value of N = 10.

(viii) Find the values of the correlation coefficient (r_{xy}) of the two variables x and y when $b_{xy} = -0.8$ and $b_{yx} = -0.45$.

Answer:

We know that $r^2 = b_{xy} \times b_{yx}$ $\quad => r = \pm\sqrt{b_{xy} \times b_{yx}}$

Since here both b_{xy} and b_{yx} are negative, r_{xy} is also negative.

Hence $r = -\sqrt{b_{xy} \times b_{yx}} = -\sqrt{-0.8 \times -0.45} = -\sqrt{0.36}$ = - 0.6.

The value of the required correlation coefficient is – 0.6.

(ix) Find the regression equation of y on x from the following data: $\bar{x} = 10, \bar{y} = 15$ and $b_{yx} = 2.50$.

Answer:

Here we have $\bar{x} = 10, \bar{y} = 15$ and $b_{yx} = 2.50$.

The regression equation of y on x is $y - \bar{y} = b_{yx}(x - \bar{x})$ $=> y - 15 = 2.50(x - 10)$

$y = 2.50x - 25 + 15 => y = 2.50x - 10$.

The required regression equation of y on x is $y = 2.50x - 10$.

(x) Find the regression equation of x on y from the following data:
$\bar{x} = 4, \bar{y} = 5$ and $b_{xy} = 0.65$.

Answer:

Here we have $\bar{x} = 4, \bar{y} = 5 \; and \; b_{xy} = 0.65$.
The regression equation of x on y is $x - \bar{x} = b_{xy}(y - \bar{y}) \; => x - 4 = 0.65(y - 5)$
$x = 0.65y - 3.25 + 4 => x = 0.65y + 0.75$.
The required regression equation of x on y is $x = 0.65y + 0.75$.

3. State whether the following statements are TRUE or FALSE:

(i) If $r_{xy} = 0.6, cov(x, y) = 12 \; and \; \sigma_x = 5$, then the value of σ_y is 5.

Answer: False.

We have $r_{xy} = \frac{cov(x,y)}{\sigma_x \times \sigma_y} \; => 0.6 = \frac{12}{5 \times \sigma_y} \; => \sigma_y = \frac{12}{5 \times 0.6} = 4$.
The required S.D. of y $= \sigma_y = 4$.

(ii) If $cov(x, y) = \; -40, \sigma_x = 8, \sigma_y = 15$, then the correlation coefficient is – 0.33.

Answer: True

r_{xy} =The correlation coefficient between two variables $= \frac{cov(x,y)}{\sigma_x \times \sigma_y} = \frac{-40}{8 \times 15} = -0.33$.
The given statement is true.

(iii) The correlation coefficient between two variables x and y is 0.75. Then the correlation coefficient between two variables u and v is 0.75, where u = 3x – 15 and v = 4y – 28.

Answer: True

We have, u = 3x – 15 and v = 4y – 28.
$=> \sigma_u = 3\sigma_x \; ; \sigma_v = 4\sigma_y \; and \; cov(u, v) = 12cov(x, y)$.
Since the magnitude of correlation coefficient is independent upon the change of scale and the change of origin, the correlation coefficient of u, v is the same that of x, y; because $r_{uv} = \frac{cov(u,v)}{\sigma_u \times \sigma_v} = \frac{3 \times 4 \times cov(x,y)}{3 \times 4 \times \sigma_x \times \sigma_y} = \frac{cov(x,y)}{\sigma_x \times \sigma_y} = r_{xy} = 0.75$.

(iv) If $r_{xy} = 0.8$ then the value of r_{uv} where u = 2x + 3 and v = - 5y + 10 is also 0.8.

Answer: False

We have, u = 2x + 3 and v = - 5y + 10
$\Rightarrow \sigma_u = 2\sigma_x ; \sigma_v = |-5|\sigma_y = 5\sigma_y \; and \; cov(u,v) = 2 \times (-5)cov(x,y).$
Since the magnitude of correlation coefficient is independent upon the change of scale and the change of origin, we have, $r_{uv} = \frac{cov(u,v)}{\sigma_u \times \sigma_v} = \frac{-10 \times cov(x,y)}{2 \times 5 \times \sigma_x \times \sigma_y} = -\frac{cov(x,y)}{\sigma_x \times \sigma_y} = -r_{xy} = -0.8.$

(v) The correlation coefficient (r_{xy}) between two variables x and y when cov(x, y) = - 2.05, var(x) = 2.1609 and var(y) = 5.1076 is – 0.617.

Answer: True.

We have, $r_{xy} = \frac{cov(x,y)}{\sigma_x \times \sigma_y} = \frac{-2.05}{\sqrt{2.1609} \times \sqrt{5.1076}} = -\frac{2.05}{1.47 \times 2.26} = -\frac{2.05}{3.3222} = -0.6170609$
$\approx -0.617.$

(vi) If the sum of squares of the differences in ranks of Physics and Chemistry marks of 15 students is 27 and then the rank correlation coefficient is 0.7925.

Answer: False.

We have, $\sum D^2 = 27$, N= 15, where D is the differences of the ranks in Physics and Chemistry marks of 15 students.
So, the rank Correlation coefficient= $1 - \frac{6\sum D^2}{N(N^2-1)} = 1 - \frac{6 \times 27}{15(225-1)} = 1 - \frac{162}{3360} = 1 - 0.0482$
$= 0.9518.$

(vii) When the correlation coefficient between two variables x and y is $r_{xy} = 0$, then cov(x,y) = 0.

Answer: True.

We have $r_{xy} = \frac{cov(x,y)}{\sigma_x \times \sigma_y} \Rightarrow 0 = \frac{cov(x,y)}{\sigma_x \times \sigma_y} \Rightarrow cov(x.y) = 0.$
So the result is true.

(viii) When the variables are not independent, the correlation coefficient may be zero.

Answer: True.

Let x and y be two variables and $y = x^2$ i.e, the variables are not independent. We take

x	-3	-2	-1	0	1	2	3
y	9	4	1	0	1	4	9

Here $\sum x = 0, \sum y = 28 \; and \; \sum xy = -27 - 8 - 1 + 0 + 1 + 8 + 27 = 0$

So, $\bar{x} = \frac{\sum x}{n} = 0 \; and \; \bar{y} = \frac{\sum y}{n} = 4$. Thus $cov(x, y) = \frac{1}{n}\sum xy - \bar{x} \times \bar{y} = 0 - 0 = 0$

Hence $r_{xy} = \frac{cov(x,y)}{\sigma_x \times \sigma_y} = \frac{0}{\sigma_x \times \sigma_y} = 0.$

4. Answer the following questions:

(i) For any two variables x and y, var(ax + by) = a^2var(x) + b^2var(y) + 2ab×cov(x,y). Where a and b are any two constant.

Answer:

Let v = ax + by. So, $\bar{v} = \frac{\sum(ax+by)}{n} = \frac{a\sum x}{n} + \frac{b\sum y}{n} = a\bar{x} + b\bar{y}.$

Now $v - \bar{v} = ax + by - (a\bar{x} + b\bar{y}) = a(x - \bar{x}) + b(y - \bar{y}).$

So, $var(v) = var(ax + by) = \frac{1}{n}(v - \bar{v})^2 = \frac{1}{n}\sum\{a(x - \bar{x}) + b(y - \bar{y})\}^2$

$= \frac{1}{n}\sum\{a^2(x - \bar{x})^2 + b^2(y - \bar{y})^2 + 2ab \times (x - \bar{x})(y - \bar{y})\}$

$= a^2\frac{1}{n}\sum(x - \bar{x})^2 + b^2\frac{1}{n}\sum(y - \bar{y})^2 + 2ab\frac{1}{n}\sum(x - \bar{x})(y - \bar{y}).$

= a^2var(x) + b^2var(y) + 2ab×cov(x,y).

(ii) Find the coefficient of correlation from the following data:

X	**1**	**2**	**3**	**4**	**5**	**6**	**7**
y	**6**	**8**	**11**	**9**	**12**	**10**	**14**

Answer: 0.8457

Here $\bar{x} = \frac{\sum x}{7} = 4 \; and \; \bar{y} = \frac{\sum y}{7} = 10.$

Table: Calculation for obtaining coefficient of correlation

x	y	u = $x - \bar{x}$	v = $y - \bar{y}$	u^2	v^2	uv
1	6	- 3	- 4	9	16	12
2	8	- 2	- 2	4	4	4
3	11	- 1	1	1	1	-1
4	9	0	- 1	0	1	0
5	12	1	2	1	4	2
6	10	2	0	4	0	0
7	14	3	4	9	16	12
$\sum x = 28$	$\sum y = 70$	$\sum u = 0$	$\sum v = 0$	$\sum u^2 = 28$	$\sum v^2 = 42$	$\sum uv = 29$

Hence, the required correlation coefficient $= r_{xy} = r_{uv} = \frac{\frac{1}{n}\sum uv - \left(\frac{\sum u}{n}\right)\left(\frac{\sum v}{n}\right)}{\sqrt{\frac{\sum u^2}{n} - \left(\frac{\sum u}{n}\right)^2} \times \sqrt{\frac{\sum v^2}{n} - \left(\frac{\sum v}{n}\right)^2}}$

$= \frac{\frac{1}{7}\times 29 - 0\times 0}{\sqrt{\frac{28}{7} - \left(\frac{0}{7}\right)^2} \times \sqrt{\frac{42}{7} - \left(\frac{0}{7}\right)^2}} = \frac{29}{\sqrt{28}\times\sqrt{42}} = \frac{29}{\sqrt{1176}} = \frac{29}{34.293} = 0.8457.$

The required correlation coefficient = 0.8457.

(iii) In order to find the correlation coefficient between two variables X and Y from 12 pairs of observations, the following calculations were made: $\sum x = 30, \sum y = 5, \sum x^2 = 670, \sum y^2 = 285 \; and \; \sum xy = 334$, on subsequent verification it was found that the pair (x = 11, y = 4) was copied wrongly, the correct value being (x = 10, y = 14). Find the correct value of correlation coefficient.

Answer: 0.775

Correct values are $\sum x = 30 - 11 + 10 = 29; \sum y = 5 - 4 + 14 = 15;$
$\sum x^2 = 670 - 121 + 100 = 649; \sum y^2 = 285 - 16 + 196 = 465$
$\sum xy = 334 - 44 + 140 = 430 \; and \; n = 12.$

Thus correct value of correlation coefficient $= \frac{\frac{1}{n}\sum xy - \left(\frac{\sum x}{n}\right)\left(\frac{\sum y}{n}\right)}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \times \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}$

$= \frac{\frac{1}{12}\times 430 - \frac{29}{12}\times\frac{15}{12}}{\sqrt{\frac{649}{12} - \left(\frac{29}{12}\right)^2} \times \sqrt{\frac{465}{12} - \left(\frac{15}{12}\right)^2}} = \frac{35.8333 - 3.02083}{\sqrt{54.0833 - 5.8404}\times\sqrt{38.75 - 1.5625}} = \frac{32.8124}{\sqrt{48.2429}\times\sqrt{37.1875}}$

$= \frac{32.8124}{42.356} = 0.775.$

The correct value of correlation coefficient is 0.775.

(iv) The rankings of 10 individuals at the start and on the finish of a course of training are as follows:

Individuals	A	B	C	D	E	F	G	H	I	J
Rank Before	1	6	3	9	5	2	7	10	8	4
Rank After	6	8	3	7	2	1	5	9	4	10

Calculate Spearman's Coefficients of correlation.

Answer: 0.3939

Table: Calculation for obtaining Rank correlation coefficients.

Individuals	Rank Before (R_1)	Rank After (R_2)	Difference in Ranks ($D = R_1 - R_2$)	D^2
A	1	6	-5	25
B	6	8	-2	4
C	3	3	0	0
D	9	7	2	4
E	5	2	3	9
F	2	1	1	1
G	7	5	2	4
H	10	9	1	1
I	8	4	4	16
J	4	10	-6	36

Here $N = 10, \sum D = 0 \; and \; \sum D^2 = 100.$

Spearman's Rank correlation coefficients = R= $1 - \frac{6\sum D^2}{N(N^2-1)} = 1 - \frac{6\times100}{10(100-1)} = 1 - \frac{600}{990}$

$= 0.3939.$

Hence, the Rank correlation coefficients = 0.3939.

(v) Find the rank correlation co efficient for the following data of marks obtained by 10 students in Mathematics and Statistics:

Student (Roll Nos.)	1	2	3	4	5	6	7	8	9	10
Marks in Mathematics	80	38	95	30	74	84	91	60	66	40
Marks in statistics	85	50	92	58	70	65	88	56	52	46

Answer: 0.8182.

Table: Calculation for obtaining Rank Correlation Coefficient:

Student (Roll Nos.)	Marks in Mathematics	Marks in statistics	Rank in Mathematics (R_1)	Rank in Statistics (R_2)	D= $R_1 - R_2$	D^2
1	80	85	4	3	1	1
2	38	50	9	9	0	0
3	95	92	1	1	0	0
4	30	58	10	6	4	16
5	74	70	5	4	1	1
6	84	65	3	5	-2	4
7	91	88	2	2	0	0
8	60	56	7	7	0	0
9	66	52	6	8	-2	4
10	40	46	8	10	-2	4

Here N=10, $\sum D = 0$ and $\sum D^2 = 30$

Spearman's Rank Correlation coefficient = R = $1 - \frac{6\sum D^2}{N(N^2-1)}$

$= 1 - \frac{6\times 30}{10(10^2-1)} = 1 - \frac{180}{10\times 99} = 1 - \frac{2}{11} = \frac{9}{11} = 0.8182.$

$\therefore The\ required\ ranks\ correlation\ coefficient = 0.8182$

(vi) In the following table are recorded data showing the lest scores made by 10 salesmen on an intelligence test and their weekly sales:

Salesmen	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Test Score	**50**	**70**	**50**	**60**	**80**	**50**	**90**	**50**	**60**	**60**
Sales ('000 ₹)	**25**	**60**	**45**	**50**	**45**	**20**	**55**	**30**	**45**	**30**

Calculate the rank correlation coefficient between intelligence and efficiency in salesmanship.

Answer: 0.70.

Table : Calculation for obtaining rank correlation Coefficient.

Salesmen	Intelligence		Sales ('000 ₹)		D= $R_1 - R_2$	D^2
	Test Score	Rank (R_1)	Amount	Rank (R_2)		
1	50	8.5	25	9	-0.5	0.25

2	70	3	60	1	2	4
3	50	8.5	45	5	3.5	12.25
4	60	5	50	3	2	4
5	80	2	45	5	-3	9
6	50	8.5	20	10	-1.5	2.25
7	90	1	55	2	-1	1
8	50	8.5	30	7.5	1	1
9	60	5	45	5	0	0
10	60	5	30	7.5	-2.5	6.25

Here N=10 $\sum D = 0 \quad \sum D^2 = 40$

This is tied problem.

$\therefore$ $Rank\ Correlation\ coefficient\ for\ tied\ ranks$ = R = $1 - \frac{6\left(\sum D^2 + \sum\frac{t^3-t}{12}\right)}{N(N^2-1)}$ ..... (1)

(t = Number of ties.)

Now $\sum\frac{(t^3-t)}{12} = \frac{3^3-3}{12} + \frac{4^3-4}{12} + \frac{3^3-3}{12} + \frac{2^3-2}{12}$ = 2+ 5+ 2+ 0.5 = 9.5

From (1), we get, R = $1 - \frac{6\,(40+9.5)}{10(99)} = 1 - \frac{6\times 49.5}{990} = 1 - \frac{297}{990} = \frac{693}{990} = 0.70$

The required rank Correlation = 0.70 (Ans.)

(vii) If u = 2x - 3 and v = $\frac{1}{3}y + 1.5$, find the value of regression coefficient b_{uv} and b_{vu} when b_{xy}= 0.5 and b_{yx}= 1.4.

Answer: b_{uv}= 3 and b_{vu} = 0.2333.

We have, u = 2x –3 and v = $\frac{1}{3}y + 1.5$

We know that var (ax+b) = a^2 var (x) and cov $[ax + b\,, cy + d] = ac\ cov\ (x, y)$

$\therefore$ σ_u^2= var (u) = 2^2 var (x) = $4\sigma_x^2$; σ_v^2= var (v) = $\left(\frac{1}{3}\right)^2$ var (y) = $\frac{1}{9}$ var (y) = $\frac{1}{9}\,\sigma_y^2$

and $cov\ (u, v) = 2 \times \frac{1}{3}\ cov\ (x, y) = \frac{2}{3}$ cov (x,y)

$\therefore$ $b_{uv} = \frac{cov\ (u,v)}{\sigma_v^2} = \frac{\frac{2}{3}cov\ (x,y)}{\frac{1}{9}\sigma_y^2} = \frac{2}{3} \times \frac{9}{1}\frac{cov\ (x,y)}{\sigma y^2} = 6\times b_{xy} = 6 \times 0.5 = 3$

and b_{vu}= $\frac{cov\ (u,v)}{\sigma_u^2} = \frac{\frac{2}{3}cov\ (x,y)}{4\sigma_x^2} = \frac{1}{6}\frac{cov\ (x,y)}{\sigma_x^2} = \frac{1}{6}\ b_{yx} = \frac{1}{6} \times 1.4$ = 0.2333.

The required value of b_{uv}= 3 and b_{vu} = 0.2333.

(viii) Find the two linear regression equations from the following data:

X	1	2	3	4	5
Y	3	2	5	4	6

Answer: x = 0.8y -0.2 *and* $y = 0.8x + 1.6$.

Solution:

Table: Calculation for obtaining regression equations:

x	y	u = x - $\bar{x}$ = x - 3	v = y - $\bar{y}$ = y - 4	u^2	v^2	uv
1	3	-2	-1	4	1	2
2	2	-1	-2	1	4	2
3	5	0	1	0	1	0
4	4	1	0	1	0	0
5	6	2	2	4	4	4

Here, $\sum x = 15$; $\sum y = 20$; $\sum u = 0$; $\sum v = 0$; $\sum u^2 = 10$; $\sum v^2 = 10$; $\sum uv = 8$

$\therefore \bar{x} = \frac{\sum x}{n} = \frac{15}{5} = 3$; $\bar{y} = \frac{\sum y}{n} = \frac{20}{5} = 4$; $b_{xy} = b_{uv} = \frac{\sum uv}{\sum v^2} = \frac{8}{10} = 0.8$

and $b_{yx} = b_{vu} = \frac{\sum uv}{\sum u^2} = \frac{8}{10} = 0.8$

$\therefore$ *The regression equation of x on y is* $x - \bar{x} = b_{xy}(y - \bar{y})$ =>x − 3 = 0.8 (y - 4)

=> x = 0.8y + 3 -3.2 =>x = 0.8y -0.2

The regression equation of y on x is $y - \bar{y} = b_{yx}(x - \bar{x})$ =>y− $4 = 0.8(x - 3)$

=>y = 0.8x + 1.6.

The regression equation of x on y and y on x are respectively

x = 0.8y -0.2 *and* $y = 0.8x + 1.6$

(ix) Two lines of regression are given by 8x – 10y + 66 = 0 & 40x – 18y = 214 and variance of x = 9. Then find the (i). averages values of x and y; (ii). Correlation coefficient r and (iii). S.D. of y. Also estimate the value of x when y = 4.

Answer: $\bar{x}$=13, $\bar{y}$ =17; 0.6; 4; 7.15.

(i) We know that the regression lines intersect at $(\bar{x}, \bar{y})$. So solving the regression equations 8x – 10y + 66 = 0 & 40x – 18y = 214 we get the average value of x and y i.e., $\bar{x}$ and $\bar{y}$. Let, 8x – 10y = - 66..... (1) and 40x – 18y = 214..... (2)

Multiplying (1) by 9 and (2) by 5, we get, 72x – 90y = - 594..... (3)
and 200x – 90y = 1070..... (4)

Subtracting (3) from (4) , we get, 128x = 1664 => $x = \frac{1668}{128} = 13$

$Putting\ the\ value\ of\ x\ in\ (1), we\ get\ 8 \times 13 - 10y + 66 = 0$ =>10y = 170 => y = 17.
∴ $\bar{x}$= x=13, $\bar{y}$= y= 17 and hence, the average values of x and y are respectively 13 & 17.

(ii) $let\ 8x - 10y + 66 = 0$ be the regression equation of $x\ on\ y\ .Now, 8x = 10\ y - 66$
=> x= $\frac{10}{8}y - \frac{66}{8}$ => $b_{xy} = \frac{10}{8}$

∴ $40x - 18y = 214$ be the regression equation of y on x. So, $18y$= 40x- 214
=> y = $\frac{40}{18}x - \frac{214}{18}$ =>$b_{yx} = \frac{40}{18}$
∴ $r_{xy}^2 = b_{xy} \times b_{yx} = \frac{10}{8} \times \frac{40}{18} = \frac{50}{18} > 1$

That is impossible. So we take the other way.
∴ $8x - 10y + 66 = 0\ is\ the\ regression\ equation\ of\ y\ on\ x\ and\ 40x - 18y = 214\ is\ the\ regression\ equation\ of\ x\ on\ y.$
∴ $10y = 8x + 66$ => y = $\frac{8}{10}x + 6.6$ => $b_{yx} = \frac{8}{10}$
and 40x = 18y +214 => x = $\frac{18}{40}y + \frac{214}{40}$ => $b_{xy} = \frac{18}{40}$
∴ $r_{xy}^2 = b_{xy} \times b_{yx} = \frac{8}{10} \times \frac{18}{40} = \frac{9}{25} < 1$
∴ r_{xy}= $\pm \frac{3}{5}$ = 0.6; Since both b_{xy} and b_{yx} are positive, r_{xy} is also positive.
∴ $The\ required\ correlation\ co\ efficient = 0.6.$

(iii) Again $b_{yx} = \frac{8}{10}$ => $r_{xy} . \frac{\sigma_y}{\sigma_x} = \frac{8}{10}$ => $\frac{3}{5} \times \frac{\sigma_y}{3} = \frac{8}{10}$ =>$\sigma_y = \frac{8 \times 5}{10} = 4$ $(Since\ \sigma_x^2 = 9)$
∴ $S.D\ of\ y\ is\ 4.$

When y=4 given, we have to find x. We use, the regression equation of x on y .
∴ $x = \frac{18}{40}y + \frac{214}{40} = \frac{18}{40} \times 4 + 5.35 = \frac{18}{10}$+ 5.35= 1.8 + 5.35= 7.15.
∴ $The\ required\ value\ of\ x\ is\ 7.15.$

(x) Regression equation of two correlated variables x and y are 5X – 6Y + 90 = 0 and 15X – 8Y -130 = 0. Find which equation is meant for y on x and which one is for x on y. Find also means and correlation coefficient of the two series.

Answer:

Regression equation of X on Y is 15X – 8Y -130 = 0 and the regression equation of Y on X is 5X – 6Y + 90 = 0; $\bar{X} = 30\ and\ \bar{Y} = 40$; $\frac{2}{3}$.

We know that the point of intersection of two lines of regression equation is $(\bar{X}, \bar{Y})$.

So solving the regression lines 5X – 6Y + 90 = 0 ……. (i) and 15 X – 8Y -130 = 0 ……(ii) we get form $(i) \times 4 - (ii) \times 3$ that – 25X = - 750, => X = 30.
Substituting X = 30 in (i), Y = 40.

Hence, the mean values of X and Y are respectively $\bar{X} = 30\ and\ \bar{Y} = 40$.

Let us assume that 5X – 6Y + 90 = 0 is the regression equation of X on Y and 15X – 8Y -130 = 0 is the regression equation of Y on X.

Thus, 5X – 6Y + 90 = 0 $=> X = \frac{6}{5}Y - 18 => b_{XY} = \frac{6}{5}$
and 15X – 8Y -130 = 0 $=> Y = \frac{15}{8}X - \frac{130}{8} => b_{YX} = \frac{15}{8}$

In that case $r_{XY}^2 = b_{XY} \times b_{YX} = \frac{6}{5} \times \frac{15}{8} = 2.25 > 1$, which is not possible. So our assumption is wrong. We take the other way.

So, 5X – 6Y + 90 = 0 is the regression equation of Y on X $=> Y = \frac{5}{6}X + 15 => b_{YX} = \frac{5}{6}$
& 15X – 8Y -130 = 0 is the regression equation of X on Y $=> X = \frac{8}{15}Y - \frac{130}{15} => b_{XY} = \frac{8}{15}$.

Hence, $r_{XY}^2 = b_{XY} \times b_{YX} = \frac{8}{15} \times \frac{5}{6} = \frac{4}{9} < 1$. So, our assumption is correct in this case.

Therefore, $r_{XY} = \pm\frac{2}{3}$. Since both b_{XY} and b_{YX} are positive. So, $r_{xy} = \frac{2}{3}$

Thus, the regression equation of X on Y is 15X – 8Y -130 = 0 and the regression equation of Y on X is 5X – 6Y + 90 = 0

The correlation coefficient = $r_{xy} = \frac{2}{3}$.

Study Note – 6

Probability

Learning Objective:

After studying this chapter, the student will be able to:

- ***Find the theoretical probability of a single event occurring, including probabilities of 0 and 1.***
- ***Use the basic counting principle to determine the number of possible outcomes. Determine whether given situations are mutually exclusive and find the theoretical probability.***
- ***Determine whether events are independent.***
- ***Find the permutation of an ordered arrangement of a set of objects, including duplication of objects, using a calculator.***
- ***Solve probability problems using combinations and binomial combinations.***

I. Objective Type Question:

1. From the definition, prove that P $\left(\overline{A}\right)$ = 1-P (A).

Answer:

Let n be total number of sample points of the sample space of the random experiment and m be the total number of sample points in favour of the event A. Then

$P(A) = \frac{m}{n}$.

Again , (n – m) sample points will go in favour of the event $\overline{A}$.

$\therefore P\ \overline{A} = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - P(A)$.

2. From The definition prove that O ≤ P (A) ≤ 1 for any event A.

Answer:

Let n = total number of sample points are in favour of the random experiment E.
m = total number of sample points are in favour of the event A.

$\therefore P(A) = \frac{m}{n}$.

Then $O \leq m \leq n$

Or $\frac{o}{n} \leq \frac{m}{n} \leq \frac{n}{n}$

Or, $O \leq P(A) \leq 1$.

3. The odds in favour of the event A is 4: 3. What is probability of the event A.

Answer:

Since the odds in favour of the event A is 4: 3, if 4 outcomes go in favour of the event A, then 3 outcomes will go against the event A.

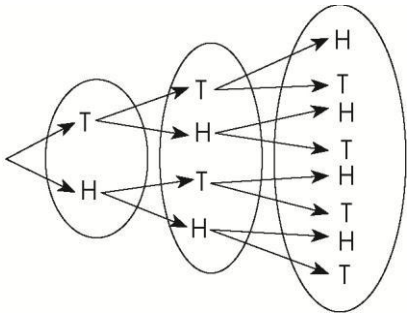
$\therefore P(A) = \frac{m}{n} = \frac{4}{4+3} = \frac{4}{7}$.

4. An unbiased coin is tossed 3 times in succession. Construct the sample its sample space.

Answer:

When a coin is tossed, then the sample space is head and tail. We denote the head by H and the tail by T.
$\therefore$ For the tossing of a coin, the sample space is {H, T}.

Now, we draw the tree diagram for tossing a coin 3 times in succession.



$\therefore$ The required sample space of the random experiment of tossing a coin 3 times in succession is {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

5. For any two events A and B, prove that $P(A \cap B) \geq P(A) + P(B) - 1$.

Answer:

For any two events A and B, We have, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
But we known that any event the probability ≤ 1.
$\therefore P(A \cup B) \leq 1$.
Or, $P(A) + P(B) - P(A \cap B) \leq 1$.
Or, $P(A) + P(B) - 1 \leq P(A \cap B)$
$\therefore P(A \cap B) \geq P(A) + P(B) - 1$.

6. If A and B are two independent events, then prove that $\bar{A}$ and $\bar{B}$ are also independent (Where $\bar{A}$ is the complement event of A).

Answer:

For any two event A and B we have, $\overline{A \cap B} = \bar{A} \cap \bar{B}$ ($\because$ by De Morgan's Law)

$\therefore P\left(\bar{A} \cap \bar{B}\right) = P\left(\overline{A \cup B}\right) = 1 - P(A \cup B)$

$= 1 - [P(A) + P(B) - P(A \cap B)]$

$= 1 - P(A) - P(B) + P(A \cap B)$

$= P\left(\bar{A}\right) - P(B) + P(A).\ P(B) \quad [\because P(A \cap B) = P(A + B)P(A - B)$

$= P\left(\bar{A}\right) - P(B)\ [1 - P(A)]$

$= P\left(\bar{A}\right) - P(B).P(\bar{A})$

$= P\left(\bar{A}\right)\ [1 - P(B)]$

$= P\left(\bar{A}\right).\ P(\bar{B})]$

$\therefore \bar{A}$ & $\bar{B}$ are also independent

7. For any two independent events A and B, P (A ∪ B) = 1 - P($\bar{A}$) P($\bar{B}$).

Answer:

Since A and B are independent, we have $P(A \cap B) = P(A) \times P(B)$

We have $P(A \cup B) = 1 - P(\overline{A \cup B})$

$= 1 - P(\bar{A} \cap \bar{B})$ (By De Morgan law)

$= 1 - P\left(\bar{A}\right).\ P(\bar{B})]$

($\because$ A and B are independent, A and B are independent and $P\left(\bar{A} \cap \bar{B}\right) = P\left(\bar{A}\right).P(\bar{B})$.

8. If $P(A \cap B^c) = 1/3$ and $P(A \cup B) = 2/3$, then find P (B).

Answer:

We have by De Morgan's Law.

$(A \cup B)^c = A^c \cap B^c$

$\therefore P(A \cup B)^c = P(A^c \cap B^c)$

Or, $P(A^c \cap B^c) = P(A \cup B)^c = 1 - P(A \cup B) = 1 - \frac{2}{3} = \frac{1}{3}.$

Now $P(B^c) = P(A \cap B^c) + P(A^c \cap B^c)$

$= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

$\therefore P(B) = 1 - P(B^C) = 1 - \frac{2}{3} = \frac{1}{3}$

9. If P (A) = 1/3, P (B) = 2/7 and P (A ∪ B) = ½ then find P (A^c ∩ B).

Answer:

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

Or, $P(A \cap B) = P(A) + P(B) - P(A \cup B)$

$= \frac{1}{3} + \frac{2}{7} - \frac{1}{2} = \frac{14+12-21}{42} = \frac{26-21}{42} = \frac{5}{42}$

$P(A^C \cap B) = P(B) - P(A \cap B) = \frac{2}{7} - \frac{5}{42} = \frac{12-5}{42} = \frac{7}{42} = \frac{1}{6}.$

$\therefore P(A^C \cap B) = \frac{1}{6}.$

10. The odds against student X solving a problem in statistics is 8 to 6 and odds in favour of student Y solving the same problem is 14 to 16. What is the chance that the problem will be solved if they both try independent by of each other?

Answer:

Let A denote the event that the student X will solve the problem and B denote the event that the student Y will solve the problem then

$P(A) = \frac{6}{8+6} = \frac{6}{14} = \frac{3}{7}$ and $P(B) = \frac{14}{14+16} = \frac{14}{30} = \frac{7}{15}$

The required Prb. = $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\frac{3}{7} + \frac{7}{15} - P(A) \cdot P(B)$ (Since A and B are independent)

$= \frac{3}{7} + \frac{7}{15} - \frac{3}{7} \times \frac{7}{15}$

$= \frac{45+49-21}{105} = \frac{73}{105}$

$\therefore$ The required probability of solving the problem = $\frac{73}{105}$.

II. True & False Problem: (State whether the problem is true false.)

1. If the odds against an event is 2 : 3, Then the probability of the event is 3/5.

Answer:

The statement is True.

Because since the odds against the event is 2: 3. Then the probability of the event is $\frac{3}{2+3} = \frac{3}{5}$. So the gives statement is true.

2. If a random experiment has three possible events A, B and C, Which are mutually exclusive, It is given that $P(A) = \frac{1}{2}$, $P(B) = \frac{3}{2}$, $P(C) = \frac{1}{4}$. The probabilities of the events A, B, C given above are possible.

Answer:

The given statement of the probabilities are not possible and hence the given statements of probabilities are not correct as P (A) + P(B) + P (C) = $\frac{1}{2}+\frac{3}{4}+\frac{1}{4} = \frac{3}{2} > 1$ Which is not true.

3. If two events A and b are independent, then $\overline{A}$ and B are also independent.

Answer:

The given statement is true. Because if A and B are independent then $\overline{A}$ and B are also in dependent.

4. If a die is their own, then the probability of getting an even number or a multiple of 3 is $\frac{5}{6}$.

Answer:

The given statement is false.

Because when a die is thrown then the sample space is S = {1,2,3,4,5,6}

Let A and b denote the event that an even number and a multiple of 3 be obtained respectively.

Then P (A) $= \frac{3}{6} = \frac{1}{2}$; P (B) = $= \frac{2}{6} = \frac{1}{3}$; P (A ∩ B) = $= \frac{1}{6}$;

Since A = { 2, 4, 6} B = {3, 6}, A ∩ B = {6}.

The required prob. = P (A ∪ B) = P (A) + P (B) – P(A ∩ B)

$$= \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{3+2-1}{6} = \frac{4}{6} = \frac{2}{3}$$

So the statement is false.

5. If an unbiased coin be tossed twice, A and B denote the events of getting "Head in the first throw and "Head" in second there, then the events A and B are independent.

Answer:

The given statement is true.

Because, the sample space here is = S = { HH, HT, TH, TT} A = {HH, HT}, B= {HH, TH}, A ∩ B = {HH}

$$\therefore P(A) = \frac{2}{4} = \frac{1}{2}; P(B) = \frac{2}{4} = \frac{1}{2}; P(A \cap B) = \frac{1}{4}.$$

P (A ∩ B) = ¼ and P (A) × P(B) = $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

∴ P (A ∩ B) = P (A) . P (B).

So A and B are independent.

∴ The given statement true.

6. If two events A and b are independent where P (A) > O and P (B) > O, then A and B are mutually exclusive.

Answer:

The given statement is false.

Because, if A and B be two independent events where P (A) > O, P (B) > O, P (A ∩ B) = P(A). P(B).

Since P (A) > O and P (B) > O, ∴ P(A ∩ B) = P (A), P (B) > O

i.e., P (A ∩ B) > O and so A ∩ B ≠ Ø.

∴ A and b cannot be mutually exclusive. So the statement is false.

7. If 3 balls are drawn at random form a bag containing 6 balls and 4 red balls, then the probability that 2 balls are blue and one ball is red is $\frac{2}{3}$.

Answer:

The given statement is false.

Because if 3 balls are drawn from the bag containing 6 blue and 4 red balls, then total no. of sample points in the sample space = $^{10}C_3 = \frac{10\times 9\times 8}{3\times 2\times 1} = 120.$

Let A denote the event of drawing 2 blue balls and one red ball = $^{6}C_2 \times {}^{4}C_1 = \frac{6\times 5}{2\times 1}\times 4 = 60.$

$\therefore$ The requires probability = P (A) = $\frac{60}{120} = \frac{1}{2}$. So the statement is false.

8. The probability that in a game of bridge hand of 13 cares will contain all the 4 aces is $\frac{11}{4165}$.

Answer:

The given statement is true.
Because total number of sample point in the sample space of the random experiment = $^{5}C_{13} = \frac{\lfloor 5}{\lfloor 13\times \lfloor 37}$

Let A denote the event that a bride hand will contain u the 4 aces and 9 other cards (the pack of 52 cards contains 4 aces and 52 – 4 = 48 other cards).

Then total no. of sample points in favour of the event A = $^{4}C_4 \times 48cq = 1\times \frac{\lfloor 48}{\lfloor 9 \lfloor 39}$.

$\therefore$ P (A) = $\frac{\lfloor 48}{\lfloor 9 \lfloor 39} \times \frac{\lfloor 13\times \lfloor 39}{\lfloor 52} = \frac{13\times 12\times 10\times 11}{50\times 49\times 52\times 51} = \frac{11}{4165}$.

$\therefore$ The statement is true.

9. The probability that A speaks the truth is 0.4 and that B speak the truth is 0.7. Then the probability that they will contradict eacess other of the same fact is 0.60.

Answer:

The given statement is false.
Because let X and Y denote the events that A speaks the truth and B speaks the truth.

Then P (X) = 0.4, P (Y) = 0.7
Here X and Y are independent.
The required probability = $P(X \cap Y) + P(X^c \cap Y)$

$= P(X) \times P(Y^c) + P(X^c) \times P(Y)$

$= P(X)\,[1 - P(Y)] + [1 - P(X)]\,P(Y)$

$= 0.4 \times [1 - 0.7] + [1 - 0.4] \times 0.7$

$= 0.4 \times 0.3 + 0.6 \times 0.7 = 0.12 + 0.42 = 0.54$

So the statement is false.

10. The probability that the teacher in statistics will take a surprise test on any day of a week is $\frac{2}{5}$. A student absents for 2 days in a week. The probability that the student will miss at least one test is $\frac{16}{25}$.

Answer:

The statement is true.

The prob. That the student will not miss a test on ant of two days. $= \left(1-\frac{2}{5}\right)\left(1-\frac{2}{5}\right) = \frac{3}{5}\times\frac{3}{5} = \frac{09}{25}$.

$\therefore$ The required probability $= \left(1-\frac{09}{25}\right) = \frac{16}{25}$.

$\therefore$ The given statement is true.

III. Multiple Chaise Question :

1. Bag contains 3 white and 5 black balls. The probability that a ball drawn at random will be black is:

(a) $\frac{5}{8}$ (b) $\frac{3}{8}$ (c) 1 (d) None of these.

Answer: (a)

Total no. of balls in the bag = 3+5 = 8.1 ball can be drawn out of these 8 balls in ${}^{8}c_{1} = 8$ ways. The bag contain 5 black balls, among them one ball can be drawn in ${}^{5}c_{1} = 5$ ways. So total number of favorable case let A denote the event that a black ball is drawn. Then P (A) = $\frac{5}{8}$. So (a) is the correct answer.

2. 3 coins are tossed. Then the probability that they all fall heads is:

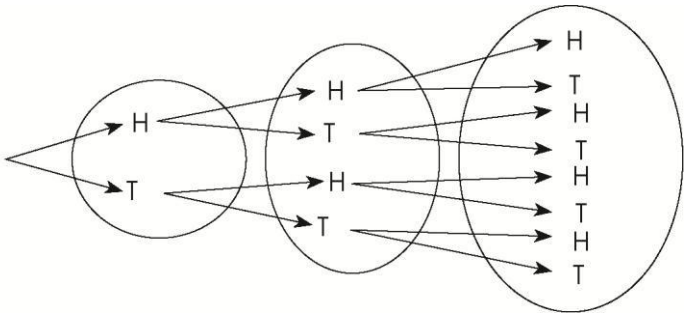
(a) $\frac{3}{8}$ **(b)** $\frac{1}{4}$ **(c)** $\frac{1}{2}$ **(d)** $\frac{1}{8}$.

Answer: (d)

When 3 coins are tussle then the sample space is { HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
So total number of sample points in the sample space of the random experiment that 3 coins are tossed is 8.

Let A denote the event that all falls head. Then observing the sample space, only one sample point goes in favour of the event A.

∴ The probability P (A) = $\frac{1}{8}$. So (d) is the correct answer.



3. The probability of drawing either a spade or an ace from a pack of 52 cards is:

(a) $\frac{3}{13}$ **(b)** $\frac{4}{13}$ **(c)** $\frac{6}{13}$ **(d) None of these.**

Answer: (b)

A pack of cards contains 52 cards, Let the random experiment be of drawing a card from the pack. Then total no. of sample points in the sample space be $52_{c1} = 52$.

Let A denote the event of drawing either a spaced or an ace from the pack. Then there are 13 spacers and 3 other aces which are not spaced. So total no. of sample points in favour of the event A is $16_{C1} = 16$.

$\therefore P(a) = \frac{16}{52} = \frac{4}{13}$.

$\therefore$ The required probability is $\frac{4}{13}$. So (b) is the correct answer.

4. Given P (A) = $\frac{1}{4}$, P (B) = $\frac{1}{5}$ and P (A $\cup$ B) = $\frac{1}{3}$, then P (A $\cap$ B) and P ($\bar{A} \cup \bar{B}$) are:

(a) $\frac{7}{60}; \frac{53}{60}$ (b) $\frac{7}{50}; \frac{52}{61}$ (c) $\frac{8}{49}; \frac{49}{60}$ (d) None of these.

Answer: (a)

Now $P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{1}{4} + \frac{1}{5} - \frac{1}{3} = \frac{15+12-20}{4\times5\times3} = \frac{7}{60}$

$P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B})$ (By De Morgan's law)

$= 1 - P(A \cap B) = 1 - \frac{7}{60} = \frac{60-7}{60} = \frac{53}{60}$.

$\therefore P(A \cap B) = \frac{70}{60}$, $P(\bar{A} \cup \bar{B}) = \frac{53}{60}$.

$\therefore$ (a) is the correct answer.

5. Given P (A) = $\frac{2}{5}$, P (B) = $\frac{3}{5}$, and P (A $\cup$ B) = $\frac{4}{5}$. Then A, B are:

(a) Mutually exclusive (b) Mutually independent (c) Both (a) and (b) (d) None of these.

Answer: (d)

We have , $P(A \cap B) = P(A) + P(B) - P(A) \cap B)$

$= \frac{2}{5} + \frac{3}{5} - \frac{4}{5} = \frac{2+3-4}{5} = \frac{1}{5}$.

So A and B are not mutually exclusive.

$P(A) \cdot P(B) = \frac{2}{5} \times \frac{3}{5} = \frac{6}{5}$ and $P(A \cap B) = \frac{1}{5}$

$\therefore P(A) \cdot P(B) \neq P(A \cap B)$.

So and B are not mutually exclusive.

$\therefore$ A and B are dependent.

So (d) is the correct answer.

6. A pack of blade contains 10 blades of which 3 and defective. If 2 blades are selected at random and removed from the pack in succession, the probability that all two blades are defective is:

(a) $\frac{2}{9}$ (b) $\frac{2}{45}$ (c) $\frac{1}{15}$ (d) None of these.

Answer: (c)

Let A denote the vent that the first blade is defective and B denote the event that the second blade is defective.

Now A ∩ B is the event that all two blades are defective.

$\therefore$ P (A) = $\frac{3}{10}$, P (B /A) = $\frac{2}{9}$.

Now P (B/A) = $\frac{P(A \cap B)}{P(A)}$ or, P (A ∩ B) = P (A). P

$\therefore$ P (A ∩ B) = $\frac{3}{10} \times \frac{2}{9} = \frac{1}{15}$.

$\therefore$ The probability that all two blades are defective is $\frac{1}{15}$. So (c) is the correct answer.

7. 3 cards are drawn at random from a pack of 52 cards. The probability of getting 2 aces is:

(a) $\frac{3}{11050}$ (b) $\frac{72}{5525}$ (c) $\frac{24}{425}$ (d) None of these.

Answer: (b).

If 3 cards be drawn from a pack of 52 cards in 52

$$= \frac{\lfloor 52}{\lfloor 3 \lfloor 49} = \frac{52 \times 51 \times 50}{3 \times 2} = 22100 \text{ ways.}$$

Let A denote the event of getting 2 aces. Since the pack of card contains only 4 aces and 48 other cards which are not aces. We have to draw 2 aces from 4 aces and 1 card from remaining 48 cards.

$\therefore$ P (A) = $\frac{{}^4C_2 \times {}^{48}C_1}{22100} = \frac{6 \times 48}{22100} = \frac{72}{5525}$.

$\therefore$ The required probability = $\frac{72}{5525}$. $\therefore$ (b) is the correct answer.

8. Two letters are drawn at random from the word "HOME". The probability that one is a vowel is:

(a) $\frac{1}{6}$ (b) $\frac{1}{2}$ (c) $\frac{5}{6}$ (d) None of these.

Answer: (c).

Here the random experiment is to draw two letters from the word "HOME".
∴ The sample space is { (H, O), (H, M), (H, E), (O, H), (O, E), (M, E)}
So total number of sample point in the sample space = 6.
Let A denote the event that "none of the close letter is a vowel".
∴ The event a contains only one point (H, M),

$\therefore P(A) = \frac{1}{6}$

So the required probability that at least one chosen letter is a vowel

$= 1\ P(A) = 1 - \frac{1}{6} = \frac{5}{6}$

∴ (c) is the correct answer.

9. The probability of drawing either a spade or an ace from a pack of 52 cards is :

(a) $\frac{4}{13}$ **(b)** $\frac{1}{52}$ **(c)** $\frac{3}{52}$ **(d) None of these.**

Answer: (a)

Let A denote the event of drawing a spade and B denote the event of drawing an Ace.
Then A ∪ B = The event of getting either a spade or an Ace.
A ∩ B = the event of getting a card which is spade and ace.

There are 13 spades and 4 acces in the pack of 52 cards. Again 1 card is spade and ace out of 52 cards.

$\therefore P(A) = \frac{13_{C1}}{52_{C1}} = \frac{13}{52} = \frac{1}{4}$

$P(B) = \frac{4_{C1}}{52_{C1}} = \frac{4}{52} = \frac{1}{13}$

$P(A \cap B) = \frac{1}{52}$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{4} + \frac{1}{13} - \frac{1}{52} = \frac{13+4-1}{52} = \frac{16}{52} = \frac{4}{13}$.

∴ The required probability = $P(A \cup B) = \frac{4}{13}$.

∴ (a) is the correct answer.

10. 40% of the students in a class are girls. If 60% and 70% of boys and girls respectively of the class pass a certain test, Then the probability that a randomly selected student from the class will have passed the test is:

(a) 0.36 (b) 0.28 (c) 0.64 (d) None of these.

Answer: (c)

In the class the event that a student chosen passed the test. Let the total number of student in the class = 100.

∴ Number girls student who passed the test = 40 × 70/100 = 28

Number of boys student who passed the test = 60 × 60/100 = 36

∴ Total number of student who passed the test = 28 + 36 = 64.

$\therefore P(A) = \frac{64}{100} = 0.64.$

∴ The required probability = 0.64.

∴ (c) is the correct answer.

11. The three events A, B and C are mutually exclusive and exhaustive and exhaustive. If P (B) = $\frac{3}{2}$ P (A) and P (C) = $\frac{1}{3}$ P (B), then P(C) =

(a) $\frac{1}{3}$ (b) $\frac{1}{6}$ (c) $\frac{1}{2}$ (d) None of these.

Answer: (b)

Let P (C) = $\frac{1}{x}$ ∴ P (B) = 3P (C) = 3x

P (B) = $\frac{3}{2}$ P(A) or, P(A) = $\frac{2}{3}$ P(B) = $\frac{2}{3} \times 3x = 2x.$

Again, A, B, C are exhaustive events.

∴ P (A) + P(B) + P(C) = 1 as P (A ∪ B ∪ C) = 1 and A, B, C are pair wise mutually exhaustive.

Or , 2x + 3x + x = 1 or, 6x = 1 or, x = 1/6 .

∴ P(C) = $\frac{1}{6}$ ∴ (b) is the correct answer.

IV. Short Answer type Question:

1. In an examination 30% of the students have failed in mathematics, 20% of the students have failed in chemistry and 10% have failed in both mathematics and chemistry. A student is selected at random.

(i) What is the probability that the student has fail in mathematics if it is known that he has failed in chemistry?

(ii) What is the probability that the student has failed either in mathematics or in chemistry?

Answer:

Let M and C denote the events that a student at random. Who fails in Mathematics' and 'fails in chemistry" respectively.

Then $P(M) = \frac{30}{100} = 0.3$; $P(C) = \frac{20}{100} = 0.2$

$P(M \cap C) = \frac{10}{100} = 0.1$.

(i) The required probability that the student selected random has failed in Mathematics and it is known that he has failed in chemistry $= P(M/C) = \frac{P(M \cap C)}{P(C)} = \frac{0.1}{0.2} = \frac{1}{2} = 0.5.$

(ii) The required probability that the student selected at random has failed either in mathematics or in chemistry

$= P(M \cup C) = P(M) + P(C) - P(M \cap C)$

$= 0.3 + 0.2 - 0.1 = 0.4.$

2. A problem in statistics is given to 3 students A, B, C, whose chances of solving it are $\frac{1}{2}, \frac{1}{3}$ and $\frac{1}{4}$ respectively. If they try it independently, What is the probability that the problem will solved?

Answer:

Let X, Y, Z denote the events that the problem will be solved by the students A, B and C respectively. Then

$P(X) = \frac{1}{2}$; $P(Y) = \frac{1}{3}$ and $P(Z) = \frac{1}{4}$

$P(X^c)$ = the probability that the problem will not be solved by A = $1 - P(X) = 1 - \frac{1}{2} = \frac{1}{2}$. Semi since, they try $(Y^C) = 1 - P(Y) = 1 - \frac{2}{3} = \frac{2}{3}; P(Z^C) = 1 - (Z) = 1 - \frac{1}{4} = \frac{3}{4}$. to solve the problem independently,

then the probability that the problem will be solved by none of A, B and C = P (X^c) .P(Y^c). P (Z^c)

$= \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} = \frac{1}{4}.$

∴ The required probability that the problem will be solved by at least any one of A, B, C

$= 1 - \frac{1}{4} = \frac{3}{4}.$

∴ The required probability $= \frac{3}{4}.$

3. Two sets of candidates are competing for the positions on the Board of Directors of a company. The probabilities that the first and second sets will win are 0.6 and 0.4 respectively. If the first set wins, the probability of introducing a new product is 0.8 and the corresponds probability, if the second set wins is 0.3. What is the probability that the new produce will be introduced?

Answer:

Let A, B denote the set of events that the 1st set wins and 2nd set wins respectively. Let C denote the event that the new product will be introduced. We have to find P (C).

Now P (C) = P (A ∩ C) + P (B ∩ C) ……..(1)

But $P(C/A) = \frac{P(C \cap A)}{P(A)}$ Or, $P(C \cap A) = P(A \cap C) = P(A).P(C/A)$

$P(C/B) = \frac{P(C \cap B)}{P(B)}$ Or, $P(C \cap B) = P(B \cap C) = P(B).P(C/B)$

∴ From (1), we get P (C) = P(A) . P (C/A) + P (B). P (C/B). ……..(2)

It is given that P (A) = 0.6; P (B) = 0.4

P(C/A) = 0.8; P (C/B) = 0.3

Substituting in (2), we get, P (C) = 0.6 × 0.8 + 0.4 × 0.3 = 0.48 + 0.12 = 0.60

∴ The required probability = P (C) = 0.60.

4. There are 3 men aged 60, 65, and 70 years. The probability to live 5 years more is 0.8 for a 60 year old, 0 for a 65 – year old and 0.3 for a 70 year old persons. Find the probability that at least two of 3 – persons will remain alive 5 years hence.

Answer:

Let a be the event that the man aged 6 years will live 5 years more.

Let B and C denote the events that two men aged 65 and 70 years respectively will live 5 years more.

By the given conditions, we get,
P (A) = 0.8; P (B) = 0.6; P (C) = 0.3.
$\therefore$ P (A^c) = 1 – P (A) = 1 – 0.8 = 0.2
P (B^c) = 1 – P (B) = 1 – 0.6 = 0.4
P (C^c) = 1 – P (C) = 1 – 0.3 = 0.7.

Here A, B, C are independent events. So A^c, B^c, C^c are also mutually independent events.
We have to find at least two of the 3 persons will remain alive 5 years hence.

$\therefore$ The required probability = P ($A \cap B \cap C^c$) + P ($A \cap B^c \cap C$) + P ($A^c \cap B \cap C$) + P ($A \cap B \cap C$).
= P (A). P (B) . P (C^c) + P (A). P (B^c) . P (C) + P (A^c). P (B) . P (C) + P (A). P (B) . P (C)
= 0.8 × 0.6 × 0.7 + 0.8 × 0.4 × 0.3 + 0.2 × 0.6 × 0.3 + 0.8 × 0.6 × 0.3
= 0.336 + 0.096 + 0.036 + 0.144 = 0.612.

5. In a bolt factory, machines M_1, M_2, M_3 manufacture respectively 25, 35 and 40 percent of the total output, of their output 5, 4 and 2 percent respectively are defective bolts. One bolt is drawn at random from the product and is found to be defective. What is the probability that it is manufactured from the machine M_2?

Answer:

Let B_1, B_2, B_3 be the events that the bolt drawn at random be manufactured by machines M_1, M_2, and M_3 respectively. Let A be the events those the bolt drawn is defective.

Then by the given condition,

$$P(B_1) = \frac{25}{100} = \frac{1}{4};\ P(B_2) = \frac{35}{100} = \frac{7}{20};\ P(B_3) = \frac{40}{100} = \frac{1}{50}.$$

$$P(B_2 / A) = \frac{P(B_2).P(A/B_2)}{P(B_1)\times P(A/B_1) + P(B_2)\times P(A/B_2) + P(B_3)\times P(A/B_3)}$$

$$= \frac{\frac{7}{20}\times\frac{1}{25}}{\frac{1}{4}\times\frac{1}{20} + \frac{7}{20}\times\frac{1}{25} + \frac{2}{5}\times\frac{1}{50}} = \frac{\frac{7}{500}}{\frac{25+28+16}{2000}} = \frac{7}{500}\times\frac{2000}{69} = \frac{28}{69}.$$

6. Boxes 1 and 2 contain respectively 4 white, 3 red and 3 blue balls, 5 white, 4 red and 3 blue balls. If one ball is drawn at random from each box, what is the probability that both the balls are of the same color?

Answer:

Sine one ball is drawn from each box and both the draw balls are of the same color, the balls may be both white or both red r both blue.

Let, A, B, C denote the events that both the balls are while, red and blue respectively.
The first box contains 10 balls and the second box contains 12 balls.

Let A1 and A2 denote the events that the drawing of 1 white ball from box 1 and 1 white ball from box 2 respectively.

$$\therefore P(A_1) = \frac{4C_1}{10_{C1}} = \frac{4}{10} = \frac{2}{5}; P(A_2) = \frac{5C_1}{12_{C1}} = \frac{4}{10} = \frac{5}{12}$$

$$\therefore P(A) = P(A_1) \times P(A_2) = \frac{2}{5} \times \frac{5}{12} = \frac{1}{6}$$

Let B_1, and B_2, denote the events that the drawing of 1 red ball from 1 and 1 red ball from box 2 respectively.

Then $P(B_1) = \frac{3_{C_1}}{10_{C1}} = \frac{3}{10}$; $P(B_2) = \frac{4C_1}{12_{C1}} = \frac{4}{12} = \frac{1}{3}$.

$\therefore$ P (B) = P (B_1) × P (B_2) = $= \frac{3}{10} \times \frac{1}{3} = \frac{1}{10}$.

Let C_1 and C_2 denote respectively the events that the drawing of one blue ball from the box 1 and one blue ball from the box 2. Then

$$P(C_1) = \frac{3_{c1}}{10_{C1}} = \frac{3}{10} \quad ; P(C_2) = \frac{3C_1}{12_{C1}} = \frac{3}{12} = \frac{1}{4}.$$

$\therefore$ P (C) = P (C_1) × P (C_2) = $\frac{3}{10} \times \frac{1}{4} = \frac{3}{40}$

Since A, B and C are 3 mutually exclusive events, the required probability = $\frac{1}{6} + \frac{1}{10} = \frac{3}{40} = \frac{20+12+9}{120} = \frac{41}{120}$.

7. A candidate is selected for interview for 3 posts. For the post there are 3 candidates, for the second post there are 4 candidates an for the third post there are 2 candidates. What is the chance his getting at least one post?

Answer:

Let A, B C denote the events that the candidate will get the first post, second post and third post respectively. Then $P(A)=\frac{1}{3}$; $P(B)=\frac{1}{4}$; $P(C)=\frac{1}{2}$.

$P(A^C)$ = The candidate will not get the first post

$= 1 - P(A) = 1 - \frac{1}{3} = \frac{2}{3}$.

$P(B^C)$ = The candidate will not get the second post

$= 1 - P(B) = 1 - \frac{1}{4} = \frac{3}{4}$.

$P(C^C)$ = The candidate will not get the second post

$= 1 - P(C) = 1 - \frac{1}{2} = \frac{1}{2}$.

Hence A, B, C are independent. So A^C, B^C, C^C are also mutually independent.

$\therefore$ The prob. That the candidate does not get any post = $P(A^C)$, $P(B^C)$, $P(C^C)$

$$= \frac{2}{3} \times \frac{3}{4} \times \frac{1}{2} = \frac{1}{4}.$$

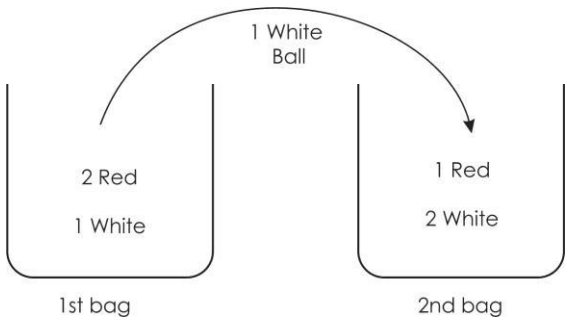
$\therefore$ The required probability = $1 - \frac{1}{4} = \frac{3}{4}$.

8. There are two bags. The first bag contains 2 red and 1 white balls where as the second bag has only, red and 2 white balls. One ball is taken out at random from the first bag and is put into the second bag. Then a ball is chosen at random from the second bag. What is the probability that this last drawn ball is red?

Answer:

Since the 1st bag contains balls, the ball transferred from the 1st bag to the second box any be either white or red

Case 1. One white ball is transferred from the first bag to the second bag.



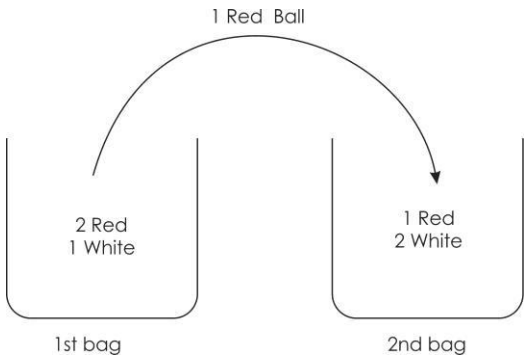
∴ Prob. of transferring 1 white ball from the first bag to the second bag = $\frac{^1C_1}{^3C_1} = \frac{1}{3}$.

After 1 white ball is transferred, then the number of white balls in the 2nd bag is 2 +1 = 3 and the total number of balls in the 2nd bag = 3+1=4.

∴ The prob. of drawing 1 white ball from the 2nd bag = $\frac{^1C_1}{^4C_1} = \frac{1}{4}$.

∴ The prob. of first transferring 1 white ball and drawing 1 red ball from the 2nd bag = $\frac{1}{4} \times \frac{1}{4} = \frac{1}{12}$.

Case 2. One red ball is transferred from the first bag to the second bag.



∴ Prob. of transferring 1 red ball from the first bag to the second bag = $\frac{^2C_1}{^3C_1} = \frac{2}{3}$.

After 1 red ball is transferred, then the number of red balls in the 2nd bag is = 2 and the total number of balls in the 2nd bag = 2 +2 =4.

∴ The prob. of drawing 1 red ball from the 2nd bag = $\frac{^2C_1}{^4C_1} = \frac{2}{4} = \frac{1}{2}$.

∴ The prob. of first transferring 1 red ball and drawing 1 red ball from the 2nd bag = $\frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$.

∴ The required probability = $\frac{1}{12} + \frac{1}{3} = \frac{5}{12}$.

9. Two boxes contain respectively 4 white and 2 black balls; 1 white and 3 black balls. One ball is transfer red from the first box into the second and then one ball is drawn from the latter. It turns to be black. What is the probability that the transferred ball was white?

Answer:

Let B_1 and B_2 denote the event that the transfers ball was white and the transferred ball was black.

Let A = the event that the drawn ball from the 2nd box is black.

We have to find P (B_1 /A)

Now P (B_1) = $\frac{4}{6} = \frac{2}{3}$; $P(B_2) = \frac{2}{6} = \frac{1}{3}$.

Also P (A/B1) = probability that the ball drawn from the 2nd box is black, assuming that the transference ball was white = $\frac{3}{5}$ (Since after transfer the white, ball from the 1st box to 2nd box, the 2nd box contain 2 white and 3 black balls

P(A/B2) = Probability that the ball drawn from the 2nd box is black assuming that the transformed ball was black = $\frac{4}{5}$.

(Since after transfer the black ball from the 1st box to the 2nd box, the 2nd box contains 1 white and 4 black ball).

Using boys' theorem, we get,

$$P(B_1 / A) = \frac{P(B_1) \times P(A / B_1)}{P(B_1) \times P(A / B_1) + P(B_2) \times P(A / B_2)}$$

$$= \frac{\frac{2}{3} \times \frac{3}{5}}{\frac{2}{3} \times \frac{3}{5} + \frac{1}{3} \times \frac{4}{5}} = \frac{\frac{6}{15}}{\frac{6}{15} + \frac{4}{15}} = \frac{6}{15} \times \frac{15}{10} = \frac{3}{5}.$$

10. 3 Identical boxes I, II, III contain respectively 4 white, 3 red balls; 3 white , 7 red balls; 2 white, 3 red, balls. A box is chosen at random and a ball is drawn from it. If the ball is found to be white, what is the probability that the Box was selected?

Answer:

Let B_1, B_2, B_3 denote the events that the Box - I to be chosen, Box- II be chosen, Box – III be chosen respectively. Then

$P(B_1) = \frac{1}{3}$; $P(B_2) = \frac{1}{3}$; $P(B_3) = \frac{1}{3}$.

Let A denote the event that the drawn ball is white. Then we have to find $P(B_2/A)$.

Now P (A/B_1) = the probability of getting a white ball, assuming that the Box – I was selected = $\frac{4}{7}$.

Similarly $P(A/B_2) = \frac{3}{10}$; $P(A/B_3) = \frac{2}{5}$.

U sing Boys' theorem, $P(B_2/A) = \frac{P(B_2)\times P(A/B_2)}{P(B_1)\times P(A/B_1)+P(B_2)\times P(A/B_2)+P(B_3)\times P(A/B_3)}$

$$= \frac{\frac{1}{3}\times\frac{3}{10}}{\frac{1}{3}\times\frac{4}{7}+\frac{1}{3}\times\frac{3}{10}+\frac{1}{3}\times\frac{2}{5}} = \frac{\frac{3}{10}}{\frac{4}{7}+\frac{3}{10}+\frac{2}{5}}$$

$$= \frac{\frac{3}{10}}{\frac{40+21+28}{70}} = \frac{3}{10}\times\frac{70}{89} = \frac{21}{89}.$$

11. The personnel department of a company has records which show the following analysis of its 200 engineers:

Age (years)	Bachelor's degree only	Master's degree	Total
Under 30	**90**	**10**	**100**
30 to 40	**20**	**30**	**50**
Over 40	**40**	**10**	**50**
Total	**150**	**50**	**200**

If one engineer is selected at random from the company find:

(a) The probability he has only bachelor's degree;

(b) The probability he has a master's degree given that he is over 40;

(c) The probability he is under 30 given that he has only a bachelor's degree.

Answer:

Let A = An engineer has a bachelor's degree only;

B = An engineer has a master's degree only;

C = An engineer is under 30 years of age;

D = An engineer is are over 40 years of age.

(a) $P(A) = \frac{150}{200} = \frac{3}{4} = 0.75$.

∴ The probability an engineer has only bachelor's degree is 0.75.

(b) $P(B/D) = \frac{P(B\cap D)}{P(D)} = \frac{\frac{10}{200}}{\frac{50}{200}} = \frac{10}{200}\times\frac{200}{50} = \frac{1}{5} = 0.20$.

∴ The probability that an engineer has a master's degree given that he is over 40 years = 0.20.

(c) $P(C/A) = \frac{P(C \cap A)}{P(A)} = \frac{\frac{90}{200}}{\frac{150}{200}} = \frac{90}{200} \times \frac{200}{150} = \frac{90}{150} = \frac{3}{5} = 0.60..$

$\therefore$ The probability that an engineer is under 30 years given than he has only a bachelor's degree = 0.60.

THE INSTITUTE OF COST ACCOUNTANTS OF INDIA

(Statutory body under an Act of Parliament)

Website : www.icmai.in
Email: studies@icmai.in
Toll Free: 1800 345 0092 / 1800 110 910

Follow us on

HEADQUARTERS
CMA Bhawan
12, Sudder Street, Kolkata - 700 016
Ph: +91-33-2252-1031/34/35/1602/
1492/1619/7373/7143
Fax: +91-33-2252-7993/1026/1723

DELHI OFFICE
CMA Bhawan
3, Institutional Area, Lodhi Road
New Delhi - 110003
Ph: + 91-11-24666100/24666124/24666122
Fax: +91-11-43583642