# SYLLABUS

| | TOPIC |
|---|---|
| 1. | **Population, Sample and Data Condensation :** Definition and scope of statistics, concept of population and simple with Illustration, Raw data, attributes and variables, classification, frequency distribution, Cumulative frequency distribution. |
| 2. | **Measures of Central Tendency :** Concept of central Tendency, requirements of a good measures of central tendency, Arithmetic mean, Median, Mode, Harmonic Mean, Geometric mean for grouped and ungrouped data. |
| 3. | **Measures of Dispersion :** Concept of dispersion, Absolute and relative measure of dispersion, range variance, Standard deviation, Coefficient of variation. |
| 4. | **Permutations and Combinations :** Permutations of 'n' dissimilar objects taken 'r' at a time (with or without repetitions). $^nP_r = n!/(n-r)!$(without proof). Combinations of 'r' objects taken from 'n' objects. $^nC_r = n!/(r!(n-r)!)$ (without proof). Simple examples, Applications. |
| 5. | **Sample space, Events and Probability :** Experiments and random experiments, Ideas of deterministic and non-deterministic experiments; Definition of sample space, discrete sample space, events; Types of events, Union and intersections of two or more events, mutually exclusive events, Complementary event, Exhaustive event; Simple examples. Classical definition of probability, Addition theorem of probability without Proof (upto three events are expected). Definition of conditional probability Definition of independence of two events, simple numerical problems. |
| 6. | **Statistical Quality Control :** Introduction, control limits, specification limits, tolerance limits, process and product control; Control charts for X and R; Control charts for number of defective {n-p chart} ,control charts for number of defects {c - chart} |

# POPULATION, SAMPLE AND DATA CONDENSATION

1. ◆ ***Write a short note on discrete and continuous distribution.*** **(2023-24)**
   ◆ ***Write short note on the Discrete and continuous distribution.*** **(2017)**

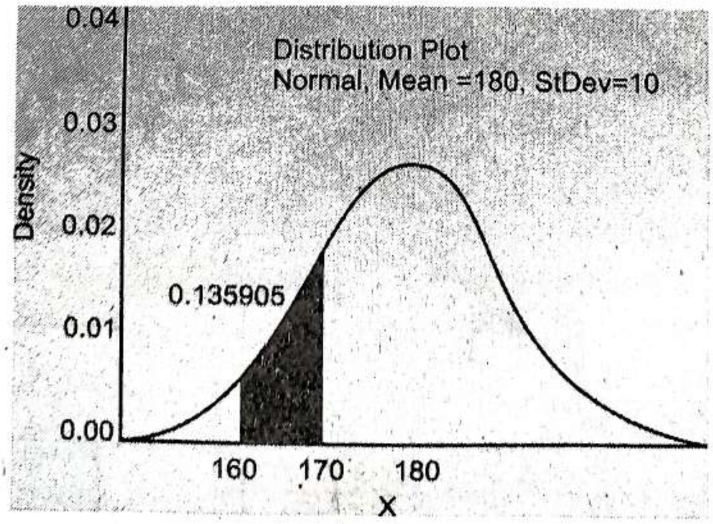
### *Continuous Distribution*

A continuous distribution describes the probabilities of the possible values of a continuous random variable. A continuous random variable is a random variable with a set of possible values (known as the range) that is infinite and uncountable.

Probabilities of continuous random variables ($X$) are defined as the area under the curve. Thus, only ranges of values can have a nonzero probability. The probability that a continuous random variable equals some value is always zero.

### *Example of the Distribution of Weights*

The continuous normal distribution can describe the distribution of weight of adult males. For example, you can calculate the probability that a man weighs between 160 and 170 pounds.




**(Figure : Distribution Plot of the Weight of Adult Males)**

The shaded region under the curve in this example represents the range from 160 and 170 pounds. The area of this range is 0.136; therefore, the probability that a randomly selected man weighs between 160 and 170 pounds is 13.6%. The entire area under the curve equals 1.0.

However, the probability that X is exactly equal to some value is always zero because the area under the curve at a single point, which has no width, is zero. For example, the probability that a man weighs exactly 190 pounds to infinite precision is zero. You could calculate a nonzero probability that a man weighs more than 190 pounds, or less than 190 pounds, or between 189.9 and 190.1 pounds, but the probability that he weighs exactly 190 pounds is zero.

### *Discrete Distribution*

A discrete distribution describes the probability of occurrence of each value of a discrete random variable. A discrete random variable is a random variable that has countable values, such as a list of non-negative integers.

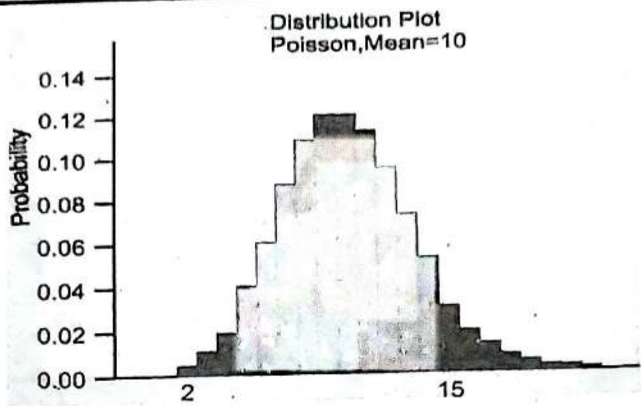
With a discrete probability distribution, each possible value of the discrete random variable can be associated with a non-zero probability. Thus, a discrete probability distribution is often presented in tabular form.

### *Example of the Number of Customer Complaints*

With a discrete distribution, unlike with a continuous distribution, you can calculate the probability that $X$ is exactly equal to some value. For example, you can use the discrete Poisson distribution to describe the number of customer complaints within a day. Suppose the average number of complaints per day is 10 and you want to know the probability of receiving 5, 10 and 15 customer complaints in a day.

| x | P (X = x) |
|---|---|
| 5 | 0.037833 |
| 10 | 0.125110 |
| 15 | 0.034718 |

You can also view a discrete distribution on a distribution plot to see the probabilities between ranges




**(Figure : Distribution Plot of the Number of Customer Complaints)**

The shaded bars in this example represent the number of occurrences when the daily customer complaint is 15 or more. The height of the bars sums to 0.08346; therefore, the probability that the number of calls per day is 15 or more is 8.35.

**2.** ***The average marks of 100 students were found to be 50. Later, it was discovered that a score of 64 was misread as 84. Find the corrected mean of the 100 students.*** **(2023-24)**

***Steps to Find the Corrected Mean***

$X_1$ = the total marks before correction.

$X_2$ = the total marks after correction.

$n$ = the number of students (which is 100 in this case).

Before correction, the total marks were

$$x_1 = 100 \times 50 = 5000$$

After correction, the misread score of 84 should be corrected to 64, so the difference is 84−64 = 20.

Therefore, $X_1 = X_2 - 20$.

So, $X_2 = 5000 - 20$

$= 4980.$

Now, to find the corrected mean, we divide $X_2$ by the number of students $n$: Corrected mean = $X_2 / n$

$= 4980/100$

= 49.8

Therefore, the corrected mean of the 100 students is 49.8.

**3. *Find 'less than' and 'more than' cumulative frequencies and draw 'Ogives' from the following data :***

| Weight (in kg) | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 |
|---|---|---|---|---|---|---|---|
| Frequency | 3 | 5 | 12 | 18 | 14 | 6 | 2 |

**(2018, 2023-24)**

Before proceeding to construct the ogives, we first need to convert the given inclusive series into an exclusive series using the following formula :

$$\text{Value of Adjustment} = \frac{\text{Value of the Lower limit of One Class} - \text{Value of Upper limit of the Preceeding Class}}{2}$$

The value of adjustment as calculated is then added to the upper limit of each class and subtracted from the lower limit of each class. Here,

$$\text{Value of Adjustment} = \frac{35-34}{2} = 0.5$$

Therefore, we add 0.5 to the upper limit and subtract 0.5 from the lower limit of each class.

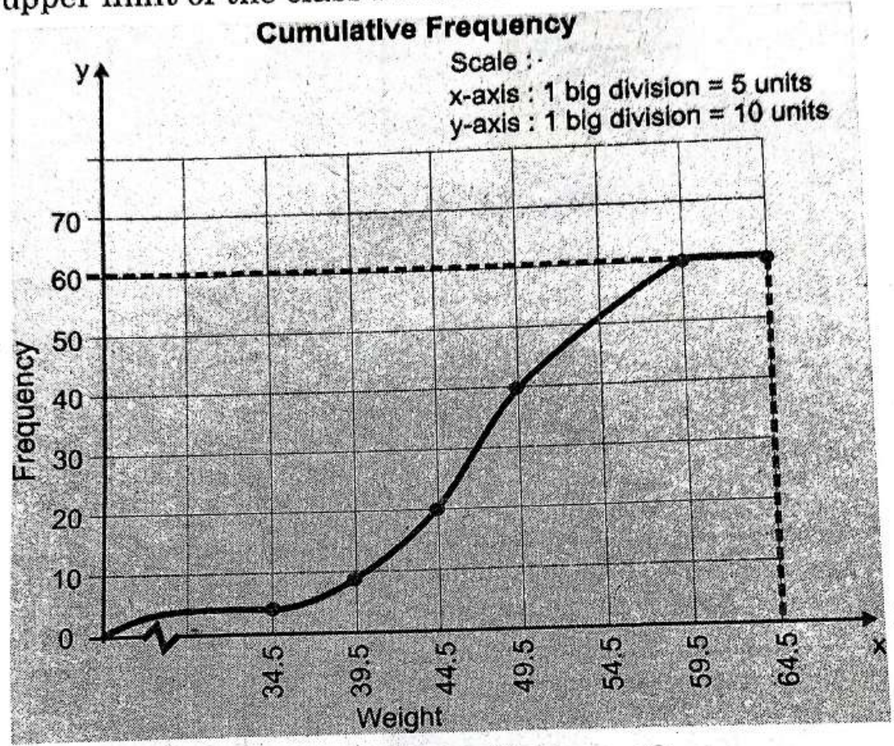
| Weight | Frequency |
|---|---|
| 29.5-34.5 | 3 |
| 34.5-39.5 | 5 |
| 39.5-44.5 | 12 |
| 44.5-49.5 | 18 |
| 49.5-54.5 | 14 |
| 54.5-59.5 | 6 |
| 59.5-64.5 | 2 |

Now for constructing a less than ogive, we convert the frequency distribution into a less than cumulative frequency distribution as follows :

| Weight | Cumulative Frequency |
|---|---|
| Less than 34.5 | 3 |
| Less than 39.5 | 3 + 5 = 8 |
| Less than 44.5 | 8 + 12 = 20 |
| Less than 49.5 | 20 + 18 = 38 |

| | |
|---|---|
| Less than 54.5 | 38 + 14 = 52 |
| Less than 59.5 | 52 + 6 =58 |
| Less than 64.5 | 58 + 2 = 60 |

We now plot the cumulative frequencies against the upper limit of the class intervals.



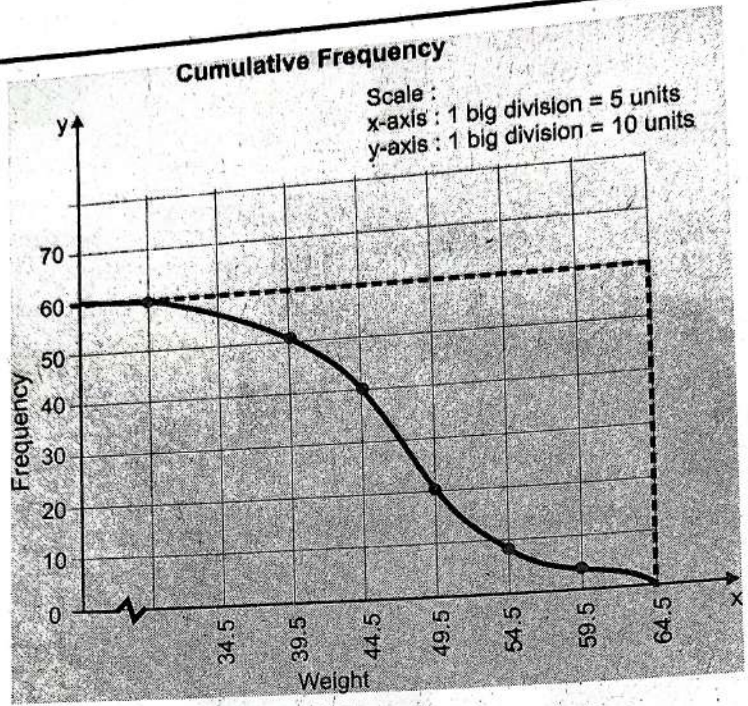

**(Figure)**

The curve obtained on joining the points so plotted is known as the less than ogive.

For constructing a more than ogive, we convert the frequency distribution into a more than cumulative frequency distribution as follows :

| Weight | Cumulative Frequency |
|---|---|
| More than 0 | 60 |
| More than 34.5 | 60 – 3 = 57 |
| More than 39.5 | 57 – 5 = 52 |
| More than 44.5 | 52 – 12 = 40 |
| More than 49.5 | 40 – 18 = 22 |
| More than 54.5 | 22 – 14 = 8 |
| More than 59.5 | 8 – 6 = 2 |
| More than 64.5 | 2 – 2 = 0 |

We now plot the cumulative frequencies against the lower limit of the class intervals. The curve obtained on joining the points so plotted is known as the more than ogive.



(Figure)

> **4.** ***What do mean by distrust of Statistics?***
> **(2018,2023-24)**

***Distrust of Statistics***

By definition, distrust means a lack of confidence or belief. Further, the science of statistics is always subject to doubt and suspicion because of its misuse by unscrupulous elements for their selfish motives. The common beliefs about statistics are :

**(1)** One ounce of truth can produce tonnes of statistics
**(2)** Statistics can prove anything
**(3)** It can prove nothing
**(4)** Figures don't lie; cheater make figures up
**(5)** There are three types of lies – lies, damn lies, and statistics
**(6)** Numbers, though accurate, are open to manipulation by selfish people to conceal the truth and present a distorted picture of the facts.

Therefore, it is important to understand that statistics is a tool, which if misused can cause a disaster. Statistics neither approves nor disapproves anything. Hence, you must take utmost care and precaution while interpreting statistical data in all manifestations.

**5. *Discuss the importance of classification in statistics.*** **(2021, 2022-23)**

Statistics classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known as statistics classification

A classification is an ordered set of related categories used to group data according to its similarities .

A classification is a useful tool for anyone developing statistical surveys. It is a framework which both simplifies the topic being studied and makes it easy to categories all data or responses received.

A standard classification will usually meet a number of requirements which are outlined below.

(1) Exhaustive categories

(2) Conceptually sound

(3) Operationally feasible

(4) Internationally comparable

**6. *Prepare "Less than" and "More than" cumulative frequency distribution. Also draw ogives from the following data :***

| Weigh (in kg ) | Frequency |
|---|---|
| 30-35 | 12 |
| 34-40 | 18 |
| 40-45 | 27 |
| 45-50 | 20 |
| 50-55 | 17 |
| 55-60 | 6 |
| 60-65 | 5 |

**(2022-23)**

To prepare the "Less than" and "More than" cumulative frequency distribution and draw ogives from the given data on weight in kilograms, we need to follow these steps :

**(1) Calculate the Cumulative Frequencies :** Cumulative frequency is the sum of all frequencies up to a particular data point.

For the given data, we have :

Cumulative frequency for 30-35 kg = 12

Cumulative frequency for 34-40 kg = 12 + 18 = 30

Cumulative frequency for 40-45 kg = 30 + 27 = 57

Cumulative frequency for 45-50 kg = 57 + 20 = 77

Cumulative frequency for 50-55 kg = 77 + 17 = 94

Cumulative frequency for 55-60 kg = 94 + 6 = 100

Cumulative frequency for 60-65 kg = 100 + 5 = 105

**(2) Prepare the "Less than" Cumulative Frequency Distribution :** In the "Less than" cumulative frequency distribution, we calculate the sum of frequencies up to and including each data point.

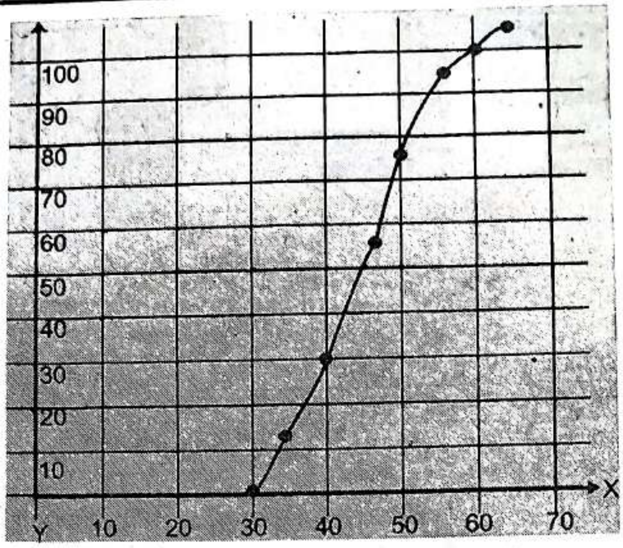
| Class | f | <upper | c.f |
|---|---|---|---|
| 25 - 30 | - | 30 | 0 |
| 30 - 35 | 12 | 35 | 12 |
| 35 - 40 | 18 | 40 | 30 |
| 40 - 45 | 27 | 45 | 57 |
| 45 - 50 | 20 | 50 | 77 |
| 50 - 55 | 17 | 55 | 94 |
| 55 - 60 | 6 | 60 | 100 |
| 60 - 65 | 5 | 65 | 105 |

**(3) Prepare the "More than" Cumulative Frequency Distribution :** In the "More than" cumulative frequency distribution, we calculate the sum of frequencies starting from each data point.

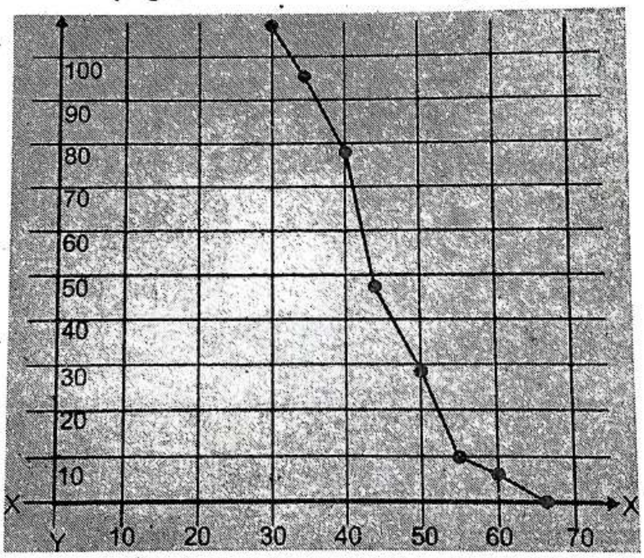
| Class | Frequency *f* | > Lower | cf |
|---|---|---|---|
| 30 - 35 | 12 | 30 | 105 |
| 35 - 40 | 18 | 35 | 93 |
| 40 - 45 | 27 | 40 | 75 |
| 45 - 50 | 20 | 45 | 48 |
| 50 - 55 | 17 | 50 | 28 |
| 55 - 60 | 6 | 55 | 11 |
| 60 - 65 | 5 | 60 | 5 |
| 65 - 70 | - | 65 | 0 |

**(4) Plotting Ogives :** An ogive is a graph that represents the cumulative frequency distribution. We plot the cumulative frequencies on the y-axis and the corresponding data points on the x-axis.

To draw ogives, we start by plotting the "Less than" cumulative frequency distribution :




(Figure : Less than C.F Ogive)




(Figure : More than C.F Ogive)

**7.** ◆ ***Distinguish between primary and secondary data.*** **(2014, 2018)**

◆ ***Write about difference between primary and secondary data.*** **(2016, 2022-23)**

The difference between primary and secondary data is as follows.

| S. No | Points | Primary Data | Secondary Data |
|---|---|---|---|
| (1) | Meaning | Data collected by researcher himself | Data collected by other person |

| | | | |
|---|---|---|---|
| (2) | **Originality** | Original or unique information | Not original or unique information. |
| (3) | **Adjustment** | Doesn't need adjustment | Need adjustment |
| (4) | **Sources** | Surveys, observations, experiments, interview | Internal Records, Govt. published records etc. |
| (5) | **Methods** | Observations, experiments, interview | Desk Research method, searching online etc. |
| (6) | **Reliability** | More reliable | Less reliable |
| (7) | **Time consumed** | More time consuming | Less time consuming |
| (8) | **Need of Investigator** | Needs team of trained investigator | Doesn't need team of trained investigator |
| (9) | **Cost Effectiveness** | Costly | Economical |
| (10) | **Collected when** | Secondary data is inadequate | Scarcity of resources like time, money and manpower. |
| (11) | **Capability** | More capable to solve a problem | Less capable to solve a problem |
| (12) | **Suitability** | Most suitable to achieve objective | Less suitable to achieve objective |
| (13) | **Bias** | Possibility of biasness exist | Less prone to biasness |
| (14) | **Collected by** | Researcher or his agents | Persons other than who collect primary data |
| (15) | **Precautions to use** | Not necessary | Quite necessary |

8. ◆ ***What do you mean by cumulative frequency distribution?*** **(2021)**

◆ ***Write a short note on Ordinary and cumulative Frequency Distribution.*** **(2017)**

◆ ***What do you mean by Cumulative frequency distribution?*** **(2017)**

## *Frequency Distribution*

The premise of data in the form of frequency distribution describes the basic pattern which the data assumes in the mass. Frequency distribution gives a better picture of the pattern of data if the number of items is large enough.

From a frequency array, it is not possible to compare characteristics of different groups. Hence for this, the classes are established to make the series of data more compact and understandable. Class limits can sometimes arbitrarily or by Sturge's formula be delimited. The width of a class that is the difference between the upper and the lower limit of the class is termed class interval. Once the classes are formed, the frequencies for these classes from raw data are expedited with the help of tally marks, little slanting vertical strokes. A bunch of four tally marks is crossed by the fifth to make the counting simpler.

**Example :** In a survey, the age of 52 women at marriage by eight parishes was reported as given below.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 24 | 25 | 27 | 26 | 22 | 23 | 24 | 25 | 24 | 25 | 24 | 23 | 26 |
| 28 | 24 | 25 | 23 | 24 | 25 | 25 | 24 | 25 | 25 | 22 | 27 | 28 |
| 27 | 26 | 25 | 24 | 25 | 28 | 26 | 25 | 27 | 25 | 24 | 27 | 24 |
| 25 | 25 | 24 | 25 | 24 | 26 | 27 | 25 | 27 | 26 | 25 | 28 | 26 |

The data can be presented in the form of frequency distribution with the help of tally marks.

| Age in years (i) | Tally marks (ii) | No. of woman (iii) |
|---|---|---|
| 22 | \|\| | 2 |
| 23 | \|\|\| | 3 |
| 24 | \|\|\|\| \|\|\|\| \|\| | 12 |
| 25 | \|\|\|\| \|\|\|\| \|\|\|\| \|\| | 17 |
| 26 | \|\|\|\| \|\| | 7 |
| 27 | \|\|\|\| \|\| | 7 |
| 28 | \|\|\|\| | 4 |
| 29 | | 0 |

The distribution constituted by columns (i) and (ii) in the above table is known as frequency distribution. It gives the number of women according to their age at marriage i.e. two women were married at the age of 22 years, three at the age of 23 years, twelve at the age of 24 years and so on.

## *Cumulative Frequency Distribution*

It can be formed on "less than" or "more than" basis.

**Example :** The birth weights (*kilogram*) of 30 children were recorded as follows :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2.0 | 2.1 | 2.3 | 3.0 | 3.1 | 2.7 | 2.8 | 3.5 | 3.1 | 3.7 |
| 4.0 | 2.3 | 3.5 | 4.2 | 3.7 | 3.2 | 2.7 | 2.5 | 2.7 | 3.8 |
| 3.1 | 3.0 | 2.6 | 2.8 | 2.9 | 3.5 | 4.1 | 3.9 | 2.8 | 2.2 |

Frequency distribution can be formed in the manner described so far, using various class intervals. The width of the classes and the number of classes will be found out by Sturge's formula. The range of data is 2.0 to 4.2 i.e.,

$L = 4.2, S = 2.$

The class interval

$$i = \frac{4.2-2.0}{1+3.322\log_{10}30} = \frac{2.2}{1+3.322\times1.4771} = \frac{2.2}{5.91} = 0.37 = 0.4$$

and

Hence six classes with a width of 0.4 kg are to be taken in the frequency distribution. The distribution with the help of tally marks is,

| Classes (weight in kg) | Tally marks | No. of children (Frequency) |
|---|---|---|
| 2.0-2.4 | IIII | 5 |
| 2.4-2.8 | IIII | 5 |
| 2.8-3.2 | IIII IIII | 9 |
| 3.2-3.6 | IIII | 4 |
| 3.6-4.0 | IIII | 4 |
| 4.0-4.4 | III | 3 |

**Note :**

**(1)** The lower limit of a class is included in that class.

**(2)** It is not necessary to choose the smallest value as the lower limit of the lowest class or the largest value as upper limit of the highest class. One may choose as 1.0-1.4, 1.4-1.8 and so on.

**9.** ***Find less than and more than cumulative frequencies and draw 'Ogives' from the following data :*** **(2021)**

| Class | Frequency |
|---|---|
| 0-9 | 5 |
| 10-19 | 15 |
| 20-29 | 18 |
| 30-39 | 30 |
| 40-49 | 15 |
| 50-59 | 10 |
| 60-69 | 5 |
| 70-79 | 2 |

| Mark Class | Less than (c.f) | Mark Class | more than (c.f) |
|---|---|---|---|
| Less than 9.5 | 5 = 5 | More than -0.5 | 100 – 0 = 100 |
| Less than 19.5 | 5 + 15 = 20 | More than 9.5 | 100 – 5 = 95 |
| Less than 29.5 | 5 + 15 + 18 = 38 | More than 19.5 | 100 – 5 – 15 = 80 |
| Less than 39.5 | 5 + 15 + 18 + 30 = 68 | More than 29.5 | 100 – 5 – 15 – 18 = 62 |
| Less than 49.5 | 5 + 15 + 18 + 30 + 15 = 83 | More than 39.5 | 100 – 5 – 15 – 18 – 30 = 32 |
| Less than 59.5 | 5 + 15 + 18 + 30+ 15 + 10 = 93 | More than 49.5 | 100 – 5 – 15 –18 – 30 – 15 = 17 |
| Less than 69.5 | 5 + 15 + 18 + 30 + 15 + 10 + 5 = 98 | More than 59.5 | 100 – 5 – 15 – 18 – 30 – 15 – 10 = 7 |
| Less than 79.5 | 5 + 15 + 18 + 30 + 15 + 10 + 5 + 2 = 100 | More than 69.5 | 100 – 5 – 15 – 18 – 30 – 15 – 10 – 5 = 2 |

**10. *Write a short note on scope of Statistics.* (2019)**

***Scope of Statistics***

Statistics is of immense use in the following cases :

**(1) In Planning :** Planning is necessary for efficient workmanship and in formulating future policies. Statistics provides the valued interpretation of facts and figures relevant to planning. Planning depends on forecasting the future. Statistics provides the necessary tools of estimation and forecasting. So statistics is indispensable in planning.

**(2) In Business :** Statistical knowledge is very helpful to the businessman. He formulates different plans and policies using statistics. It helps him in forecasting the future trends and tendencies. To estimate the market fluctuations, changes in the demand conditions etc. statistical techniques are often used. Hence for becoming a successful businessman, ideas in statistics are essential.

**(3) In Economics :** Statistics are the 'straws out of which Economists have to make bricks'. Statistics indispensable in economic studies. Statistical data and their analysis are used to solve a variety of economic problems such as in consumption, production, distribution of income and assets, poverty, unemployment etc. These problems are described, compared and correlated by using statistics. Use of statistics in Economics has led to the formulation of many economic laws like Engel's law of consumption, Pareto's Laws of distribution of income etc. Similarly statistical tools like index numbers are sued as economic barometers.

(4) **In Administration :** In ancient times statistics was used as the science of statecraft. It was used to collect data relating to manpower, crimes, income etc. for formulating different policies. In modem times, its role has become manifold. It has become indispensable for the administrators. It is being widely used formulating and forecasting different plans and policies of the state administration.

(5) **In Business Management :** Business manager take decisions in the face of uncertainty. Statistical tools like collection classification, analysis and interpretation of data are essential in business-management. The success of modem business primarily depends on accurate forecasting of the future demand and mark trends. For this statistics is essential. Hence statistics is widely used in business management.

(6) **In Research Activities :** One cannot think of undertaking any research activities without using statistics. Primarily, statistical techniques are used for collecting information in any research. Besides, statistical methods are used for analysis and interpretation of research findings. Thus there is hardly any branch of study where statistics is not being used. It is used in all spheres of human activities.

**11.** ***Find less than and more than cumulative frequencies and draw 'Ogives' from the following data :*** **(2019)**

| Class Interval | Frequency |
|---|---|
| 100 – 110 | 2 |
| 110 – 120 | 3 |
| 120 – 130 | 7 |
| 130 – 140 | 11 |
| 140 – 150 | 15 |
| 150 – 160 | 7 |
| 160 – 170 | 2 |
| 170 – 180 | 3 |

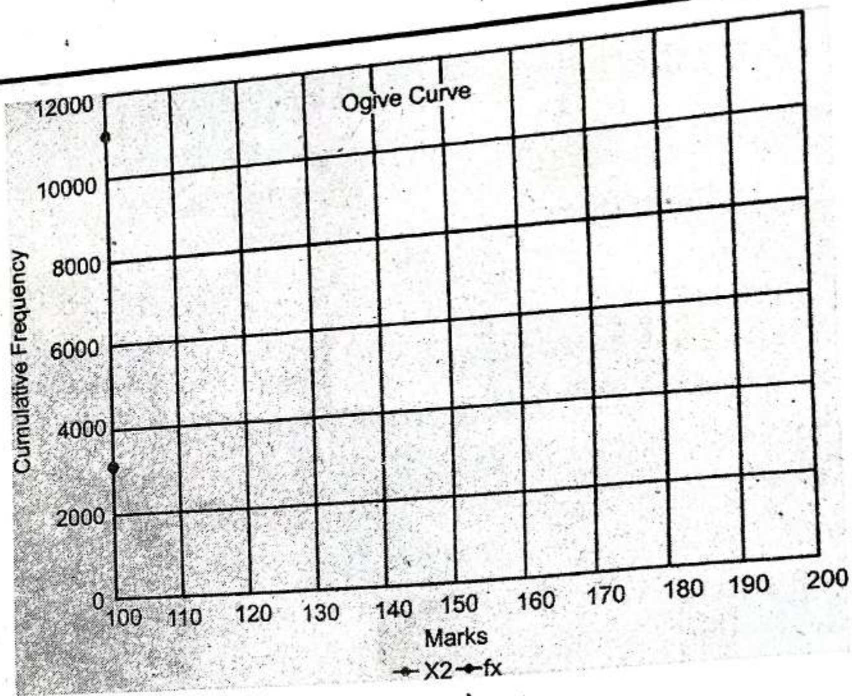
## Cumulative Frequency Curve

A cumulative frequency curve or ogive curve, is a type of frequency polygon that shows cumulative frequencies. An ogive graph plots cumulative frequency on the y-axis and class boundaries along the x-axis.

Since the cumulative frequencies can either be 'less than' or 'more than' type, there are two type of ogives called 'less than' type and 'more than' type ogive. The value of median and other partition values can be located from the ogives. The technique of drawing frequency curves and cumulative frequency curves is more or less the same. The only difference is that in case of simple frequency curves the frequency is plotted against the mid point of a class interval whereas in case of a cumulative frequency curve it is plotted at the upper or limit of a class interval depending upon the manner in which the series has been cumulated.

**Drawing the Cumulative Frequency Curve from the Data :**

| Marks | No. of Students | Less than cf | More than cf |
|---|---|---|---|
| 100-110 | 2 | 2 | 50 |
| 110-120 | 3 | 5 | 48 |
| 120-130 | 7 | 12 | 45 |
| 130-140 | 11 | 23 | 38 |
| 140-150 | 15 | 38 | 27 |
| 150-160 | 7 | 45 | 12 |
| 160-170 | 2 | 47 | 5 |
| 170-180 | 3 | 50 | 3 |




(Figure)

**12. *Write the objects of classification.*** **(2016, 2019)**

***Objects of Classification***

The principle objectives of classifying data are :

(1) To condense the mass of data in such a manner that similarities and dissimilarities can be readily apprehended. Millions of figures can thus be arranged in a few classes having common features.

(2) To facilitate comparison.

(3) To pinpoint the most significant features of the data at a glance.

(4) To give prominence to the important information gathered while dropping out the unnecessary elements.

(5) To enable a statistical treatment of the material collected.

**13. *What are main limitations of Statistics?*** **(2017)**

***Main Limitations of Statistics***

(1) Statistics does not deal with individual measurements. Since statistics deals with aggregates of facts, it cannot be used to study the changes that have taken place in individual cases. For example, the wages earned by a single industry worker at any time, taken by itself is not a statistical datum. But

(2)

(3)

(4)

(5)

14. D f q

the wages of workers of that industry can be used statistically. Similarly the marks obtained by John of your class or the height of Beena (also of your class) are not the subject matter of statistical study. But the average marks or the average height of your class has statistical relevance.

(2) Statistics cannot be used to study qualitative phenomenon like morality, intelligence, beauty etc. as these cannot be quantified. However, it may be possible to analyze such problems statistically by expressing them numerically. For example we may study the intelligence of boys on the basis of the marks obtained by them in an examination.

(3) Statistical results are true only on an average : The conclusions obtained statistically are not universal truths. They are true only under certain conditions. This is because statistics as a science is less exact as compared to the natural science.

(4) Statistical data, being approximations, are mathematically incorrect. Therefore, they can be used only if mathematical accuracy is not needed.

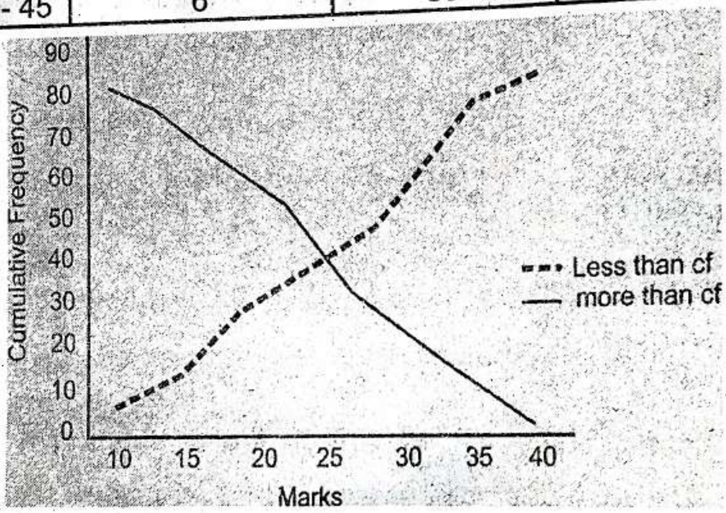
(5) Statistics, being dependent on figures, can be manipulated and therefore can be used only when the authenticity of the figures has been proved beyond doubt.

**14. *Draw the cumulative frequency curve from the following data and find out median and quartiles :*** **(2017)**

| Marks | No. of students |
|---|---|
| 10-15 | 5 |
| 15-20 | 8 |
| 20-25 | 15 |
| 25-30 | 20 |
| 30-35 | 16 |
| 35-40 | 10 |
| 40-45 | 6 |

## Drawing the Cumulative Frequency Curve from the Data :

| Marks | No. of Students | Less than cf | More than cf |
|---|---|---|---|
| 10 - 15 | 5 | 5 | 80 |
| 15 - 20 | 8 | 13 | 75 |
| 20 - 25 | 15 | 28 | 67 |
| 25 - 30 | 20 | 48 | 52 |
| 30 - 35 | 16 | 64 | 32 |
| 35 - 40 | 10 | 74 | 16 |
| 40 - 45 | 6 | 80 | 6 |




**(Figure : Cumulative Frequency Curve)**

**Median and Quartiles from the following Data :**

| Marks | No. of Students | Less than cf |
|---|---|---|
| 10 - 15 | 5 | 5 |
| 15 - 20 | 8 | 13 |
| 20 - 25 | 15 | 28 |
| 25 - 30 | 20 | 48 |
| 30 - 35 | 16 | 64 |
| 35 - 40 | 10 | 74 |
| 40 - 45 | 6 | 80 |

**Median :**

$$Md = L + \left[ \frac{\left( \frac{N}{2} - c.f \right)}{f} \times h \right]$$

Where : $L$ = Lower limit of Median Class

$c.f.$ = Cumulative frequency of the class preceding Median Class.

$f$ = Frequency of Median Class.

$h$ = Width of Median Class.

$N$ = Sum of the frequency or Total number of items in the series.

**Median Class :** Following is the method of finding the Median Class :

**(1)** Calculate cumulative frequency ($c.f.$)

**(2)** Look for $c.f.$ "just greater" than $N/2$

**(3)** The corresponding Class is Median Class.

| Marks | No. of Students | Less than cf |
|---|---|---|
| 10 - 15 | 5 | 5 |
| 15 - 20 | 8 | 13 |
| 20 - 25 | 15 | 28 |
| 25 - 30 | 20 | 48 (Median Class) |
| 30 - 35 | 16 | 64 |
| 35 - 40 | 10 | 74 |
| 40 - 45 | 6 | 80 |
| | N = 80 | |

**Note :** The above data is presented in form of Unequal Class Intervals but the formula of Median holds good for both equal as well as unequal class intervals

**Median Class :**

$N/2 = 80/2 = 40$

$c.f.$ just greater than 40 is 48

Hence, 25 – 30 is Median Class.

Therefore,

$L = 25$

$c.f = 28$

$f = 20$

$h = 25 - 30 = 5$

Substituting the values in the formula :

$$Md = 25 + \left[\frac{(40-28)}{20} \times 5\right]$$

$$Md = 25 + \left[\frac{12}{20} \times 5\right]$$

$$Md = 25 + [3]$$

$Md = 28$ Marks

Step to find $r^{th}$ Quartile Class :

**(1)** Find $rN/4$

**(2)** Look for $c.f.$ just greater than $rN/4$

(3) Corresponding Class is $r^{th}$ Quartile Class

| Marks | No. of Students | Less than cf |
|---|---|---|
| 10 - 15 | 5 | 5 |
| 15 - 20 | 8 | 13 |
| 20 - 25 | 15 | 28 ($Q_1$ Class) |
| 25 - 30 | 20 | 48 ($Q_2$ Class) |
| 30 - 35 | 16 | 64 ($Q_3$ Class) |
| 35 - 40 | 10 | 74 |
| 40 - 45 | 6 | 80 |
| | N = 80 | |

**First Quartile :** $= Q_1 = L + \left(\frac{\frac{N}{4} - c.f}{f}\right) \times h$

$Q_1$ class corresponding to $c.f$ just greater then $\frac{N}{4}$

$$\frac{N}{4} = \frac{80}{4} = 20$$

$c.f$ just greater than 20 is 28, hence $(20 - 25)$ is $Q_1$ class

Hence :

$$L = 20,\ c.f. = 13,\ f = 15,\ h = 5 \text{ and } \frac{N}{4} = 20$$

Substituting the values in the formula and calculating we get :

$$Q_1 = 20 + \left(\frac{20-13}{15}\right) \times 5$$

$$Q_1 = 20 + \left(\frac{7}{3}\right)$$

$$Q_1 = 20 + (2.33) = 22.33$$

$$Q_1 = 22.33$$

**Second Quartile :** $Q_2 = L + \left[\frac{\left(\frac{2N}{4} - c.f\right)}{f}\right] \times h$

$Q_2$ class = class corresponding to $c.f$ just greater than

$\frac{2N}{4}$

$$\frac{2N}{4} = \frac{2(80)}{4} = 40$$

*c.f.* just greater than 40 is 48

Hence, 25 – 30 is $Q_2$ Class.

Therefore,

$$L = 25$$
$$c.f. = 28$$
$$f = 20$$
$$h = 25 - 30 = 5$$

Substituting the values in the formula :

$$Q_2 = 25 + \left[\frac{(40-28)}{20} \times 5\right]$$

$$Q_2 = 25 + \left[\frac{12}{20} \times 5\right]$$

$$Q_2 = 25 + [3]$$

$$Q_2 = 28 \text{ Marks}$$

(**Note :** Median and Second Quartile are same as II Quartile divides the series into two parts and Median also does the same)

**Third Quartile :** $Q_3 = L + \left[\frac{\left(\frac{3N}{4} - c.f\right)}{f}\right] \times h$

$Q_3$ class = class corresponding to $c.f$ just greater than $\frac{3N}{4}$

$$\frac{3N}{4} = \frac{3(80)}{4} = 60$$

$c.f$ just greater than 60 is 64 , hence (30 - 35) is $Q_3$ class

Hence :

$$L = 30,\ c.f = 48, f = 16, h = 5 \text{ and } \frac{3N}{4} = 60$$

Substituting the values in the formula and calculating we get :

$$Q_3 = 30 + \left[\frac{60-48}{16}\right] \times 5$$

$$Q_3 = 30 + \left(\frac{12}{16}\right) \times 5$$

$$Q_3 = 30 + 3.75 = 33.75$$

$$Q_3 = 33.75$$

Hence :

Median : $Md = 28$ Marks

Quartiles :

$Q_1 = 22.33$ Marks

$Q_2 = 28$ Marks

$Q_3 = 33.75$ Marks

**15.** ◆ ***Write a short note on Chronological classification.*** **(2017)**

◆ ***Write a short note on Geographical classification.*** **(2017)**

◆ ***Explain Classification.***

According to L.R. Connor, classification is the process of arranging things (either actually or notionally) in groups or classes according to their resemblance and affinities, and gives expression to the unity of attributes that may subsist amongst the diversity of individuals. Professor A.R. Ilersie has stated that the statistician's first task is to reduce and simplify the details into a form so that the salient features may be brought out, while still facilitating the interpretation of the assembled data. This procedure is known as classifying and tabulating the data.

In short, a table is a systematic arrangement of data in rows and/or columns. As a matter of fact, the kind of classification or tabulation mostly depends on the type of information required for study and the type of further statistical treatment to be undertaken. Though, there are no hard and fast rules, still some norms can be given for an ideal classification and tabulation of data. Norms for an ideal classification are :

(1) The classes should be complete and non – overlapping. It means that each observation or unit must belong to a unique class. For instance, if we classify people according to their marital status, generally we classify them as married and

unmarried. But there are many who do not belong to either of these groups namely, divorcees, widows or widowers. Since the number of such people is very small as compared to the number of married and unmarried people, they can be classified as 'others'.

(2) Clarity of classes is another important property. It means that classes should be such that one can place a unit or an observation in a class without confusion.

(3) One should use standardized classes so that the comparison of results can be possible from time to time. For example, to know the educational development, we should use classification like illiterate, literate, primary, secondary, graduate, post – graduate and technical. Further, the unit of each class should be the same. The classification of data is generally done on geographical, chronological, qualitative or quantitative basis on the following lines :

**(a)** In geographical classification, data are arranged according to places, areas or regions.

**(b)** In chronological classification, data are arranged according to time i.e. weekly, monthly, quarterly, half yearly, annually, quinquennially, etc.

**(c)** In qualitative classification, the data are arranged according to attributes like sex, marital status, educational standard, stage or intensity of disease etc. It is not essential that a table should have only one attribute. If the data can be classified according to any number of attributes. If the data can be classified into two classes only, it is said to be classified according to dichotomy. If the number of classes is more than two, it is said to be a manifold classification.

**(d)** Quantitative classification means arranging data according to certain characteristic that has been measured e.g. according to height, weight or income of persons, vitamin content in a substance etc. In this type of classification, certain classes are formed and the units belonging to these classes are attached to them. One problem that arises is of determining the class intervals. This class interval will depend

ELEMENTS OF

upon the number of classes which will be arbitrarily decided keeping in view the quantum of data, measurement of characteristic and type of information required. A numerical formula as suggested by H.A. Sturges may be used for determining approximately the class interval and number of classes. The formula is,

$$i = \frac{L - S}{1 + 3.322 \log_{10} n}$$

Where $1 + 3.322 \log_{10} n = k$, number of classes

$i$ = class interval
$L$ = Largest observation
$S$ = smallest observation
$n$ = total number of observations.

**Time Series :** Another type of classification is the time series in which data or the derived values from data for each time period are arranged chronologically.

**16. *Explain the rule, types and limitation of presentation of data.*** **(2016)**

### *Data Presentation*

It is the method by which people summarize, organize and communicate information using a variety of tools, such as diagrams, distribution charts, histograms and graphs.

### *Shewart's Rule for Presentation of Data*

(1) **Rule One :** Data should always be presented in such a way that preserves the evidence in the data for all the predictions that might be made from these data.

(2) **Rule Two :** Whenever an average, range, or histogram is used to summarize data, the summary should not mislead the user into taking any action that the user would not take if the data were presented in a time series.

### *Types of Data Presentation*

The main methods of presenting numerical data are through :

(1) **Graphs :** Graphs use visual elements to make large numbers and complex information more comprehensible. They efficiently display large

(2)

(3)

*Li*
(1)

(2)

amounts of data and help in identifying and interpreting patterns in the data. The common types of graphs are the bar graph, pie charts, line graphs and scatter plots. Each type has a different function in displaying data, such as for showing comparison, analyzing proportional distributions, identifying changes over time and clarifying the connection between two numerical measurements.

(2) **Tables :** Tables are recommended for relatively smaller amounts of values that belong to a single category. Whereas graphs provide a quick overview of data, tables can emphasize these values. Incorporating data in texts even exceeds such level of emphasis, especially when there are only two values to present. As long it succeeds in communicating the results and interpretations to the audience, the researcher can use any of these methods of presenting data.

(3) **Thematic Analysis :** Qualitative data, or data that cannot translate into quantifiable measurements, requires thematic analysis to report patterns appearing in a theme or category. When dealing with qualitative data, the researcher presents findings and meaningful patterns in a theme before supporting them through existing studies in subsequent discussions. Alternatively, the research may merge the presentation and discussion into one section. The goal of either approach is to convince the reader that the thematic analysis is valid.

## *Limitation of Data Presentation*

(1) **Graphs :**

(a) They do not provide detailed information.

(b) Graphs can be easily misinterpreted.

(c) Graphs can take much time and labour.

(d) Only experts can interpret the graphs with accuracy.

(e) Exact measurements are not possible in graphs.

(2) **Tables :**

(a) It cannot be used with the more visually-oriented audience.

(b) We can only readily see increasing and decreasing values as the $x$ – value increments uniformly (aka, the $x$ – values are evenly spaced apart).

(c) It's tough to find, say, the y-intercept if the $x$-value 0 and its pair is not given, or the slope. You have to run a series of tests to determine what kind of functions the relationship between $x$, and $y$ is (whether it's linear, exponential, quadratic; etc.)

(3) **Thematic Presentation of Data :**

(a) Difficulty to compare the values with the naked eye.
(b) Impossible to design value overlapping.
(c) Poor accuracy of value estimation
(d) Difficulty to design symbol overlapping.
(e) Symbols do not express exact values.
(f) Need quite a lots of space
(g) Boundaries may suggest that densities change abruptly at the lines.

It depends upon the analyst which method to use to present the data so that it is valid, reliable and practical.

❑❑

# MEASURES OF CENTRAL TENDENCY

**1. *Find the geometric mean of 4, 9, 12 and 48.* (2023-24)**

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

$G.M = (4*9*12*48)^{1/4}$

$= (20736)^{1/4}$

We can write $\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$

So,

$= \left((20736)^{1/2}\right)^{\frac{1}{2}} = (144)^{1/2} = 12$ **(Ans.)**

**2. *By using grouping method, locate Mode from the following data.***

| Mid Value | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|---|---|---|---|---|---|---|---|
| Frequency | 7 | 12 | 17 | 29 | 31 | 5 | 3 |

**(2023-24)**

***Grouping Method***

In discrete and continuous series if the items concentrate at more than one value, attempts are made to find out the point of maximum concentration with the help of grouping method. In this method values are first arranged in ascending order and the frequencies against each value are written down. These frequencies are then added in *two's* and the totals are written in lines between the values added. Frequencies can be added in *two's* in two ways :

**(1)** By adding frequencies of item numbers 1 and 2; 3 and 4; 5 and 6 and so on.

**(2)** By adding frequencies of item numbers 2 and 3; 4 and 5; 6 and 7 and so on. After this the frequencies are added in *three's*. This can be done in three ways :

**(a)** By adding frequencies of item numbers 1, 2 and 3; 4, 5 and 6; 7, 8 and 9 and so on.

**(b)** By adding frequencies of item numbers 2, 3 and 4; 5, 6 and 7; 8, 9 and 10 and so on.

(c) By adding the frequencies of item numbers 3, 4 and 5; 6, 7 and 8, 9, 10 and 11 and so on.

| Mid vales | Class Interval | Frequency |
|---|---|---|
| 30 | 25-35 | 7 |
| 40 | 35-45 | 12 |
| 50 | 45-55 | 17 |
| 60 | 55-65 | 29 |
| 70 | 65-75 | 31 |
| 80 | 75-85 | 5 |
| 90 | 85-95 | 3 |

| Mid values | F (I) | II (1+2) | III (2+3) | IV (1+2+3) | V (2+3+4) | VI (3+4+5) |
|---|---|---|---|---|---|---|
| 30 | 7 | 19 | | 36 | | |
| 40 | 12 | | 29 | | 58 | |
| 50 | 17 | 46 | | | | 77 |
| 60 | 29 | | 60 | 65 | | |
| 70 | 31 | 36 | | | 39 | |
| 80 | 5 | | 8 | | | |
| 90 | 3 | | | | | |

| Column | Mid Values 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|---|---|---|---|---|---|---|---|
| I | | | | | √ | | |
| II | | | √ | √ | | | |
| III | | | | √ | √ | | |
| IV | | | | √ | √ | √ | |
| V | | √ | √ | √ | | | |
| VI | | | √ | √ | √ | | |
| TOTAL | | 1 | 3 | 5 | 4 | 1 | |

From the analysis table, it is clear that 60 repeats the maximum number of times. Thus, mode is 60.

**3.** ***The average marks of 100 students were found to be 30. Later, it was discovered that a score of 32 was misread as 23. Find the corrected mean of the 100 students.*** **(2022-23)**

To find the corrected mean of the 100 students after the misread score is corrected, we need to adjust the average by considering the error in the original calculation. Here's how we can calculate the corrected mean:

(1) **Calculate the Sum of the Original Scores :** Sum = Average * Number of students = 30 * 100 = 3000

(2) Adjust the sum by subtracting the misread score and adding the correct score :

Corrected Sum = Sum - Misread score + Correct score
= 3000 - 23 + 32
= 3009

**(3) Calculate the Corrected Mean :** Corrected Mean = Corrected Sum / Number of students
= 3009 / 100
= 30.09

Therefore, the corrected mean of the 100 students is 30.09.

**4. *Find the harmonic mean of 8, 16, 24.*** **(2022-23)**

To find the harmonic mean of a set of numbers, we use the formula :

Harmonic Mean = (Number of values) / (Sum of reciprocals of the values)

In this case, we have three numbers: 8, 16, and 24. So, the harmonic mean can be calculated as follows :

$$\text{Harmonic Mean} = \frac{3}{\left(\frac{1}{8}+\frac{1}{16}+\frac{1}{24}\right)}$$

$$= \frac{3}{\frac{11}{48}}$$

$$= 3 \times \frac{48}{11}$$

$$= \frac{144}{11} = 13.9$$

**5. *By Grouping method locate mode from the following data :*** **(2022-23)**

| Class | Frequency |
|---|---|
| 0-5 | 5 |
| 5-10 | 7 |
| 10-15 | 9 |
| 15-20 | 18 |
| 20-25 | 16 |
| 25-30 | 15 |
| 30-35 | 6 |
| 35-40 | 3 |

According to A.M.Tuttle 'Mode is the value which has the greatest frequency density in its immediate neighbor hood'.

## Grouping Method

In discrete and continuous series if the items concentrate at more than one value, attempts are made to find out the point of maximum concentration with the help of grouping method. In this method values are first arranged in ascending order and the frequencies against each value are written down. These frequencies are then added in two'. and the totals are written in lines between the values added.

Frequencies can be added in two's in two ways :

(1) By adding frequencies of item numbers 1 and 2; 3 and 4; 5 and 6 and so on.

(2) By adding frequencies of item numbers 2 and 3; 4 and 5; 6 and 7 and so on. After this the frequencies are added in three's. This can be done in three ways :

(a) By adding frequencies of item numbers 1, 2 and 3; 4, 5 and 6; 7, 8 and 9 and so on.

(b) By adding frequencies of item numbers 2, 3 and 4; 5, 6 and 7; 8, 9 and 10 and so on.

(c) By adding the frequencies of item numbers 3, 4 and 5; 6, 7 and 8, 9, 10 and 11 and so on.

If necessary frequencies can be added in four's and five's also after this the size of items containing the maximum frequencies are noted down and the item which has the maximum frequency the largest number of times is called the mode.

| Class | Frequency | I | II | III | IV | V | VI |
|---|---|---|---|---|---|---|---|
| 0-5 | 5 | 5 | 12 | | | | |
| 5-10 | 7 | 7 | | 16 | | | |
| 10-15 | 9 | 9 | | | 21 | | |
| 15-20 | 18 | **18** | 27 | | | 34 | |
| 20-25 | 16 | 16 | | 34 | | | 43 |
| 25-30 | 15 | 15 | 31 | | | | |
| 30-35 | 6 | 6 | | | 49 | 37 | |
| 35-40 | 3 | 3 | 9 | 21 | | | |

| | 0-5 | 5-10 | 10-15 | 15-20 | 20-25 | 25-30 | 30-35 | 35-40 |
|---|---|---|---|---|---|---|---|---|
| I | | | | ✔ | | | | |
| II | | | | | | | | |
| III | | | | | ✔ | ✔ | | |
| IV | | | | ✔ | ✔ | | | |
| V | | | | ✔ | ✔ | ✔ | | |
| VI | | | | | ✔ | ✔ | ✔ | |
| | | | ✔ | ✔ | ✔ | | | |

Therefore :

Mode =16

**6. *The mean of 200 items was 50. Later on it was discovered that two items were misread as 92 and 8 instead of 192 and 88. Find out the correct mean.*** **(2021)**

**Given :** Incorrected mean of 200 items = 50

i.e. $\sum x / 200 = 50$

Incorrected $\sum x = 200 * 50 = 10000$

Corrected $\sum x$ = Incorrected $\sum x$ – (Sum of Incorrect values) + (Sum of correct values)

Hence, Corrected $\sum x = 10000 - (92 + 8) + (192 + 88)$

$= 10000 - 100 + 280 = 10180$

Corrected Mean = Corrected $\sum x / n$

Corrected Mean = 10180 / 200 = 50.9 **(Ans.)**

**7. *Find the geometric mean of 4, 8, 16.*** **(2021)**

$$GM = \sqrt[n]{X_1 * X_2 * X_3}$$

$$= (4 * 8 * 16)^{1/3} = \left(2^2 * 2^3 * 2^4\right)^{1/3} = 2^3 = 8$$ **(Ans.)**

**8. *Find median and mode from the following distribution :*** **(2021)**

| Income (₹) | No. of Persons |
|---|---|
| 100-200 | 15 |
| 100-300 | 33 |
| 100-400 | 63 |
| 100-500 | 83 |
| 100-600 | 100 |

| Income | Income Class | No of Person(c.f) | Person(f) |
|---|---|---|---|
| 100-200 | 100-200 | 15 | 15 |
| 100-300 | 200-300 | 33 | 33 – 15 = 18 |
| 100-400 | 300-400 | 63 | 63 – 33 = 30 |
| 100-500 | 400-500 | 83 | 83-63 = 20 |
| 100-600 | 500-600 | 100 | 100 – 83 = 17 |
| | | | N = 100 |

Median is the value of N/2 = 100/2 = 50th item which lies in (300-400)

**Median :**

$$M = l + \frac{(n/2 - cf)}{30} * h = 300 + \frac{(50 - 30)}{30} * 100 = 356.67$$

$f = 30$ **(Ans.)**

**9. *The mean marks of 50 students in Section A is 40 and mean marks of 40 students in Section B is 45. Find the combined mean of the 90 students in both the Sections.*** **(2019)**

Let us put the given data in a tabular form for convenience in calculations :

| | Size (n) | Arithmetic Mean ($\bar{X}$) |
|---|---|---|
| Section A | 50 | 40 |
| Section B | 40 | 45 |

**Combined Arithmetic Mean :**

$$\overline{x_{12}} = \frac{n_1\overline{x_1} + n_2\overline{x_2}}{n_1 + n_2} = \frac{50 * 40 + 40 * 45}{50 + 40}$$

$$= \frac{2000 + 1800}{90}$$

$$= \frac{3800}{90} = 42.22\ Marks$$

**(Ans.)**

**10. *The arithmetic mean of two observations is 25 and the harmonic mean is 9. What is geometric mean of the series?*** **(2019)**

***Relationship between A.M, H.M and G.M***

$$GM = \sqrt{A.M * H.M}$$

$$GM = \sqrt{25 * 9}$$

$$GM = \sqrt{225} = 15$$

**(Ans.)**

**11. *Calculate median and mode for the following frequency distribution :*** **(2019)**

| Marks | Frequency |
|---|---|
| 25 – 30 | 22 |
| 30 – 35 | 34 |
| 35 – 40 | 50 |
| 40 – 45 | 42 |
| 45 – 50 | 38 |
| 50 – 55 | 14 |

$$\text{Mode} = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} * h$$

Where : Where $L$ = Lower Limit of Modal Class (Class with highest frequency)

$f_1$ = Frequency of Modal Class

$f_0$ = Frequency of Class preceding modal class

$f_2$ = Frequency of Class succeeding modal class

$h$ = Width of modal class

Since, highest frequency is 80, hence 50 – 60 is the modal class

| Marks | No. of Students |
|---|---|
| 25 – 30 | 22 |
| 30 – 35 | 34 $f_0$ |
| 35 – 40 | 50 $f_1$ |
| 40 – 45 | 42 $f_2$ |
| 45 – 50 | 38 |
| 50 - 55 | 14 |

Hence substituting the values :

$$\text{Mode} = 35 + \frac{50-34}{100-34-42} * 5$$

Mode = 35 + 3.33 = 38.33 **(Ans.)**

**12.** ***The average marks of 80 students were found to be 40. Later, it was discovered that a score of 54 was misread as 84. Find the corrected mean of the 80 students.*** **(Dec. 2013, 2018)**

Incorrected mean of 80 students = 40

i.e. $\sum x$ / 80 = 40

Incorrected $\sum x$ = 80 x 40 = 3200 Marks

Incorrect total of marks of 80 students = 80 x 40 = 3200 Marks

Corrected $\sum x$ = Incorrected $\sum x$ – (Sum of Incorrect values) + (Sum of correct values)

Therefore, the corrected total of the marks of 80 students = 3200 – 84 + 54 = 3170

Corrected Mean = Corrected $\sum x$ / $n$

The corrected average of 80 students = 3170/ 80 = 39.625 marks **(Ans.)**

**13. *Find the harmonic mean of 4, 8, 16, 24.*** **(2018)**

Harmonic Mean of 4, 8, 16, 24

$$HM = \frac{n}{\sum \frac{1}{x}}$$

Where, $x$ : Observation Values

$n$ : Total number of observations

| x | $\frac{1}{x}$ |
|---|---|
| 4 | $\frac{1}{4} = 0.25$ |
| 8 | $\frac{1}{8} = 0.125$ |
| 16 | $\frac{1}{16} = 0.0625$ |
| 24 | $\frac{1}{24} = 0.041667$ |
| | $\sum \frac{1}{x} = 0.479167$ |

Here, $n = 4$

$$\sum \frac{1}{x} = 0.479167$$

$$HM = \frac{4}{0.479167}$$

$$HM = 8.3478$$

**(Ans.)**

**14. *Locate mode from the following data :*** **(2018)**

| Mid Value | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|---|---|---|---|---|---|---|---|
| Frequency | 7 | 12 | 17 | 29 | 31 | 5 | 3 |

In this case, see the nature of the given data. It is not a discrete frequency distribution. Instead, the mid-values of the classes with corresponding frequencies are given. Thus, class intervals are first determined by using the given mid-values. In this question, maximum concentration of frequencies is at two places, as such, the modal class is determined by grouping method.

## Determination of Modal Class by Grouping Method

| Mid-Values | Classes | Frequency | | | | | |
|---|---|---|---|---|---|---|---|
| | | Col. 1 | Col. 2 | Col. 3 | Col. 4 | Col. 5 | Col. 6 |
| 30 | 25-35 | 7 | 19 | | 36 | | |
| 40 | 35-45 | 12 | | 29 | | 58 | |
| 50 | 45-55 | 17 | 46 | | | | 77 |
| 60 | 55-65 | 29 | | 60 | 65 | | |
| 70 | 65-75 | 31 | 36 | | | 39 | |
| 80 | 75-85 | 5 | | 8 | | | |
| 90 | 85-95 | 3 | | | | | |

**Analysis Table :**

| Column No. | Classes Contributing to Highest Frequency | | | | | | |
|---|---|---|---|---|---|---|---|
| | 25–35 | 35–45 | 45–55 | 55–65 | 65–75 | 75–85 | 85–95 |
| 1 | | | | | × | | |
| 2 | | | × | × | | | |
| 3 | | | | × | × | | |
| 4 | | | | × | × | × | |
| 5 | | × | × | × | | | |
| 6 | | | × | × | × | | |
| No. of items | - | 1 | 3 | 5 | 4 | 1 | - |

∴ The analysis table shows that (55-65) is the modal class. So, by using the interpolation formula.

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

Here, $f_o = 17, f_1 = 29, f_2 = 31, l_1 = 55, h = 10$

$$\therefore \text{Mode} = 55 + \frac{29-17}{58-17-31} \times 10 = 55 + 12 = 67.0$$

But 67.0 does not lie in the modal class (55-65) and therefore our answer is not correct. To overcome this difficulty we use interpolation formula (16), i.e.

$$\text{Mode} = l_1 + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times h$$

$$= 55 + \frac{|29-17|}{|29-17| + |29-31|} \times 10 = 55 + \frac{12 \times 10}{12+2}$$

$$= 55 + 8.57$$

$$= 63.57$$

**(Ans.)**

**15. *Prove that if values of the variable are multiplied (or divided) by a constant value, the arithmetic mean so obtained is same as the initial arithmetic mean is multiplied (or divided) by the constant value.*** **(2017)**

The mean of $n$ observations $x_1, x_2, \ldots\ldots, x_n$ is $x$. If each observation is multiplied by a nonzero number $p$, the mean of the new observations is $px$.

Now we will proof the Property :

$$x = (x_1 + x_2 + \ldots\ldots + x_n)/n \qquad \ldots(A)$$

$$\Rightarrow x_1 + x_2 + \ldots\ldots + x_n = nx$$

Mean of $px_1, px_2, \ldots\ldots, px_n$,

$$= (px_1 + px_2 + \ldots\ldots + px_n)/n$$

$$= \{p(x_1 + x_2 + \ldots\ldots + x_n)\}/n$$

$$\{p(n_x)\}/n, \text{[using (A)]}.$$

$$= p_x$$

Hence, the mean of the new observations is $p_x$

**Property :**

The mean of $n$ observations $x_1, x_2, \ldots, x_n$ is $x$. If each observation is divided by a nonzero number $p$, the mean of the new observations is $(x/p)$.

Now we will proof the Property :

$$x = (x_1 + x_2 + \ldots\ldots + x_n)/n$$

$$\Rightarrow x_1 + x_2 + \ldots\ldots + x_n = nx \qquad \ldots(A)$$

$$= \text{Mean of } (x_1/p), (x_2/p), \ldots\ldots\ldots, (x_n/p)$$

$$= \text{Mean of } (x_1/p), (x_2/p), \ldots\ldots (x_n/p)$$

$$= (1/n).(x_1/p + x_2/p + \ldots\ldots x_n/p)$$

$$= (x_1 + x_2 + \ldots\ldots + x_n)/np$$

$$= (nx)/(np), \text{[using (A)]}.$$

$$= (x/p).$$

**16. *If the A.M of two numbers is 4.5 and their H.M. is 4, then find the numbers.*** **(2017)**

Let the two numbers be $x$ and $y$

$$\text{Then : A.M} = \frac{x+y}{2} \qquad \ldots(1)$$

And H.M. $= \dfrac{2xy}{x+y}$ ...(2)

Given A.M = 4.5 and H.M = 4 in eq (1) and (2) respectively we get

From (1) : $4.5 = \dfrac{x+y}{2}$

$x + y = 9$ ...(3)

From (2) :

$4 = \dfrac{2xy}{x+y}$ ...(4)

Substituting $x + y = 9$ in Eq (4) we get :

$4 = \dfrac{2xy}{9} = 2xy = 36$

$xy = 18$ ...(5)

From Eq (3) : $x + y = 9$ or $y = 9 - x$

Substituting $y = 9 - x$ in Eq (5) we get :

$xy = 18 \Rightarrow x(9 - x) = 18$

$\Rightarrow 9x - x^2 = 18$

Or $x^2 - 9x + 18 = 0$

On solving the above equation we get :

$(x-3)(x-6) = 0$

i.e. $x = 3$ or $x = 6$

We know from Eq (3) : $x + y = 9$

if $x = 3$, $y = 6$

if $x = 6$, $y = 3$

Therefore, the two numbers are 3 and 6 whose :

And H.M $= \dfrac{2xy}{x+y} = \dfrac{2(3\times 6)}{3+6} = \dfrac{36}{9} = 4$

**Therefore, the two numbers are 3 and 6** **(Ans.)**

**17.** ***Find the mode from the given data :*** **(2017)**

| Class | Frequency |
|---|---|
| 0-5 | 5 |
| 5-10 | 7 |
| 10-15 | 9 |
| 15-20 | 18 |
| 20-25 | 16 |

| | |
|---|---|
| 25-30 | 15 |
| 30-35 | 6 |
| 35-40 | 3 |

**Calculating Mode of**

| Class | Frequency |
|---|---|
| 0-5 | 5 |
| 5-10 | 7 |
| 10-15 | 9 |
| 15-20 | 18 |
| 20-25 | 16 |
| 25-30 | 15 |
| 30-35 | 6 |
| 35-40 | 3 |

Here, the maximum frequency is 18, corresponding to the class interval (15 - 20) which is the modal class.

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

Hence substituting the values :

Where $L = 15,\ f = 18,\ f_0 = 9, f_2 = 16,\ h = 5$

$$\text{Mode} = 15 + \frac{18 - 9}{36 - 9 - 16} \times 5$$

$$\text{Mode} = 15 + (45/11)$$

$$= 15 + 4.09$$

**Mode = 19.09**

**(Ans.)**

**18.** ***If the arithmetic average of data given below be 165 rupees, find the missing term :*** **(2016)**

| Monthly Wages in (₹) | Number of Labourers |
|---|---|
| 100 | 30 |
| 150 | 20 |
| 200 | 15 |
| – | 10 |
| 300 | 4 |
| 500 | 1 |

$$\text{Arithmetic Mean A. M} = \frac{\sum fx}{N}$$

Let the missing term be '$a$'

| Monthly Wages (in ₹) (x) | Number of Labourers (f) | fx |
|---|---|---|
| 100 | 30 | 3000 |
| 150 | 20 | 3000 |
| 200 | 15 | 3000 |
| a | 10 | 10a |
| 300 | 4 | 1200 |
| 500 | 1 | 500 |
| | N = 80 | Σ fx = 10700 + 10a |

$$\frac{\sum fx}{N} = \frac{10700+10a}{80}$$

$$= 165$$

$$10700+10a = 13200$$

$$10a = 13200-10700 = 2500$$

$$a = 250$$

Therefore, missing term is 250

**19. *Calculate the median and mode from the following series :*** **(2016)**

| Marks | No. of Students |
|---|---|
| 0 – 10 | 15 |
| 10 – 20 | 25 |
| 20 – 30 | 52 |
| 30 – 40 | 56 |
| 40 – 50 | 78 |
| 50 – 60 | 80 |
| 60 – 70 | 70 |

***Mode***

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

Where, $L$ = Lower Limit of Modal Class (Class with highest frequency)

$f_1$ = Frequency of Modal Class

$f_0$ = Frequency of Class preceding modal class

$f_2$ = Frequency of Class succeeding modal class

$h$ = Width of modal class

Since highest frequency is 80, hence 50 – 60 is the modal class

| Marks | No. of Student |
|---|---|
| 0 – 10 | 15 |
| 10 – 20 | 25 |
| 20 – 30 | 52 |
| 30 – 40 | 56 |
| 40 – 50 | 78 $f_0$ |
| 50 – 60 | 80 $f_1$ |
| 60 – 70 | 70 $f_2$ |

Hence substituting the values :

$$\text{Mode} = 50 + \frac{80-78}{160-78-70} \times 10$$

$$\text{Mode} = 50 + 1.67 = 51.67$$ **(Ans.)**

**Median**

$$Md = Md = L + \left[\frac{\left(\frac{N}{2} - c.f\right) \times h}{f}\right]$$

Where : $L$=Lower Limit of Median Class
$c.f.$=$c.f.$ of the class preceding median class
$f$=Frequency of median class
$h$=Width of Median Class
$N=\Sigma f$

**Median Class :** Steps in finding Median Class

**(1)** Find $N/2$

**(2)** Look for $c.f.$ just greater than $N/2$

**(3)** Corresponding Class is Median Class

Hence,

| Marks | No. of Students | c.f. |
|---|---|---|
| 0 – 10 | 15 | 15 |
| 10 – 20 | 25 | 40 |
| 20 – 30 | 52 | 92 |
| 30 – 40 | 56 | 148 |
| 40 – 50 | 78 | 226 Md Class |
| 50 – 60 | 80 | 306 |
| 60 – 70 | 70 | 376 |
| | N = 376 | |

**Median Class :**

$$\frac{N}{2} = \frac{376}{2} = 188$$

C.f. just greater than 188 = 226, hence (40 – 50) is median class.

$$\text{Therefore, Median} = 40 + \left[\frac{(188-148)\times 10}{78}\right]$$

$$Md = 40 + 5.13 = 45.13$$ **(Ans.)**

**20.** ***By using Grouping method locate mode from the following data :*** **(2016)**

| Mid Value | Frequency |
|---|---|
| 30 | 7 |
| 40 | 12 |
| 50 | 17 |
| 60 | 29 |
| 70 | 31 |
| 80 | 5 |
| 90 | 3 |

***Process***

In order to find mode, a grouping table and an analysis table are to be prepared in the following manner.

***Grouping Table***

A grouping table consists of 65 columns.

**(1)** Arrange the values in ascending order and write down their corresponding frequencies in the column – I .

**(2)** In column – 2 the frequencies are grouped into two's and added.

**(3)** In column – 3 the frequencies are grouped into two's, leaving the first frequency and added.

**(4)** In column – 4 the frequencies are grouped into three's, and added.

**(5)** In column – 5 the frequencies are grouped into three's, leaving the first frequency and added.

**(6)** In column – 6 the frequencies are grouped into three's, leaving the first and second frequencies and added.

**(7)** Now in each these columns mark the highest total with a circle.

***Analysis Table***

After preparing a grouping table, prepare an analysis table. While preparing this table take the column numbers as rows and the values of the variable as columns.

Now for each column number see the highest total in the grouping table (Which is marked with a circle) and mark the corresponding values of the variable to which the frequencies are related by using bars in the relevant boxes.

Now the value of the variable (class) which gets highest number of bars is the modal value (model class). Calculation of mode – continuous series : In a continuous series, to find out the mode we need one step more than those used for discrete series. As explained in the discrete series, modal class it determined by inspection or by preparing grouping and analysis tables. Then we apply the following formula :

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

Where, $L$= Lower Limit of Modal Class (Class with highest frequency)

$f_1$ = Frequency of Class preceding modal class

$f_0$ = Frequency of Class succeeding modal class

$h$ = Width of modal class

OR

If the frequency of the Modal class is less than the frequency of the Succeeding or Preceding class, in such a case if we calculate Mode by the general formula than the value of the Mode will lie outside the range of the class and will be inappropriate. Therefore, in such case the Mode should be calculated by the following formula :

$$\text{Mode} = L + \frac{f_2}{f_1 + f_2} \times h$$

Where the notations have their usual meanings.

**Finding the Modal Class by Grouping Method :**

| Mid Value | Class interval | Frequency Column 1 | Col 2 | Col 3 | Col 4 | Col 5 | Col 6 |
|---|---|---|---|---|---|---|---|
| 30 | 25 – 35 | 7 | | | | | |
| | | | 19 | | | | |
| 40 | 35 – 45 | 12 | | | | | |
| | | | | | 36 | | |
| 50 | 45 – 55 | 17 | | 29 | | | |
| | | | | | | **58** | |
| 60 | 55 – 65 | 29 | **46** | | | | |
| | | | | | | | **77** |
| 70 | 65 – 75 | **31** | | **60** | | | |
| | | | | | **65** | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | 35 | | | | |
| 80 | 75 – 85 | 5 | | | | 39 | |
| | | | | 8 | | | |
| 90 | 85 – 95 | 3 | | | | | |

**Analysis Table :**

| Columns | Class Interval | | | | | | |
|---|---|---|---|---|---|---|---|
| | 25–35 | 35–45 | 45–55 | 55–65 | 65–75 | 75–85 | 85–95 |
| 1 | | | | | × | | |
| 2 | | | × | × | | | |
| 3 | | | | × | × | | |
| 4 | | | | × | × | × | |
| 5 | | × | × | × | | | |
| 6 | | | × | × | × | | |

From the analysis table the Modal Class is 55 – 65 as it has maximum number of '$x$'

| Class Interval | Frequency |
|---|---|
| 25 – 35 | 7 |
| 35 – 45 | 12 |
| 45 – 55 | 17 $f_0$ |
| 55 – 65 | **29 (Modal Class) $f_1$** |
| 65 – 75 | 31 $f_2$ |
| 75 – 85 | 5 |
| 85 – 95 | 3 |

The frequency of the Modal class is less than the frequency of the Succeeding class, therefore, in such case the Mode should be calculated by the following formula :

$$\text{Mode} = L + \frac{f_2}{f_1 + f_2} \times h$$

Where,

$L$ = Lower Limit of Modal Class (Class with highest frequency)

$f_1$ = Frequency of Modal Class

$f_0$ = Frequency of Class preceding modal class

$f_2$ = Frequency of Class succeeding modal class

$h$ = Width of modal class

$$\text{Mode} = 55 + \frac{31}{29+31} \times 10 = 55 + \frac{31}{60} \times 10$$

$$= 55 + 5.17 = 60.17$$

**(Ans.)**

**21.** ***Calculate the mode for the following frequency distribution :*** **(2015)**

| Class | 25-30 | 30-35 | 35-40 | 40-45 | 45-50 | 50-55 |
|---|---|---|---|---|---|---|
| Frequency | 25 | 34 | 50 | 42 | 38 | 14 |

***Mode***

$$M_o = L + \left[\left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times (h)\right]$$

Where;
$L$ : Lower Limit of Modal Class
$f_1$ = frequency of Modal Class
$f_0$ = frequency of class proceeding Modal Class
$f_2$ = frequency of class succeeding Modal Class
$h$ = width of Modal Class

**Modal Class :** Class corresponding to highest frequency

| Class | Frequency |
|---|---|
| 25 – 30 | 25 |
| 30 – 35 | 34 |
| 35 – 40 | 50 (Modal Class) |
| 40 – 45 | 42 |
| 45 – 50 | 38 |
| 50 – 55 | 14 |
| | N = 203 |

Because highest frequency is 50 hence, 35 - 40 is Modal Class

$L : 35$
$f_1 = 50$
$f_0 = 34$
$f_2 = 42$
$h = 5$

$$M_o = L + \left[\left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times (h)\right]$$

$$M_o = 35 + \left[\left(\frac{50 - 34}{100 - 34 - 42}\right) \times (5)\right]$$

$$M_o = 35 + \left[\left(\frac{2}{3}\right) \times (5)\right]$$

$$M_o = 35 + 3.33$$

Mode = 38.33

**(Ans.)**

**22.** ***The monthly incomes of 10 families are given by the following table :***

| A | B | C | D | E | F | G | H | I | J | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 80 | 70 | 15 | 75 | 500 | 20 | 45 | 250 | 40 | 36 | 1136 |

***Compute the arithmetic mean, geometric mean and harmonic mean of the above incomes. Which of the above three means represents the data most suitable? Give reasons.*** **(2015)**

***Arithmetic Mean***

$$A.M.\ \bar{x} = \frac{1}{n}\Sigma x$$

| Families | Income (x) |
|---|---|
| A | 80 |
| B | 70 |
| C | 15 |
| D | 75 |
| E | 500 |
| F | 20 |
| G | 45 |
| H | 250 |
| I | 40 |
| J | 36 |
| | $\Sigma x = 1131$ |

$$\bar{x} = \frac{1}{n}\sum x = \frac{1131}{10} = 113.1$$

**Harmonic Mean :**

$$HM = n / \Sigma \frac{1}{x}$$

Where; $x$ : Observation values

$n$ : Total number of observations

| Families | Income | 1/x |
|---|---|---|
| A | 80 | 0.0125 |
| B | 70 | 0.014286 |
| C | 15 | 0.066667 |
| D | 75 | 0.013333 |
| E | 500 | 0.002 |
| F | 20 | 0.05 |
| G | 45 | 0.022222 |
| H | 250 | 0.004 |
| I | 40 | 0.025 |
| J | 36 | 0.027778 |
| | | $\Sigma$ 1/x = 0.237786 |

$$HM = n / \sum \frac{1}{x}$$

$n = 10$

$HM = 10/0.237786$

$HM = 42.05467$

**Geometric Mean :**

$$\text{G.M.} = \text{Antilog} \frac{\sum(\log x)}{n}$$

| Families | Income (x) | log x |
|---|---|---|
| A | 80 | 1.90309 |
| B | 70 | 1.845098 |
| C | 15 | 1.176091 |
| D | 75 | 1.875061 |
| E | 500 | 2.69897 |
| F | 20 | 1.30103 |
| G | 45 | 1.653213 |
| H | 250 | 2.39794 |
| I | 40 | 1.60206 |
| J | 36 | 1.556303 |
| | n = 10 | $\sum$(logx) = 18.00886 |

$$\text{G.M.} = \text{Antilog} \frac{18.00886}{10}$$

$$\text{G.M.} = \text{Antilog}\,(1.800886)$$

Therefore,

$$\text{G.M.} = 63.09573$$

Arithmetic Mean = 113.1 **(Ans.)**

Harmonic Mean = 42.05467 **(Ans.)**

Geometric Mean = 63.09573 **(Ans.)**

Harmonic Mean represents the data more accurately as it is least affected by the two large values 500 and 250.

**23.** ***Find the harmonic mean from the following observations :***

**2574, 475, 75, 5, 0.8, 0.08, 0.005, 0.0009** **(2014)**

***Harmonic Mean***

$$HM = n / \sum \frac{1}{x}$$

Where; $x$ : observation values

$n$ : Total number of observations

| x | 1/x |
|---|---|
| 2574 | 0.000389 |
| 475 | 0.002105 |
| 75 | 0.013333 |
| 5 | 0.2 |
| 0.8 | 1.25 |
| 0.08 | 12.5 |
| 0.005 | 200 |
| 0.0009 | 1111.111 |
| | $\sum \frac{1}{x} = 1325.077$ |

$$HM = n / \sum \frac{1}{x}$$

$$n = 8$$

$$HM = 8/1325.077$$

$$HM = 165.6346$$

**(Ans.)**

**24.** ***Calculate arithmetic mean, median and mode from the following table :*** **(2014)**

| Income between (₹) | No. of Persons |
|---|---|
| 100 and 200 | 15 |
| 100 and 300 | 30 |
| 100 and 400 | 63 |
| 100 and 500 | 83 |
| 100 and 600 | 100 |

The above frequency distribution is less than cumulative frequency distribution. Therefore we need to convert it to the Frequency Distribution before calculating Arithmetic Mean, Median and Mode.

| Income between (₹) | No. of Persons (f) |
|---|---|
| 100 - 200 | 15 |
| 200 - 300 | 15 |
| 300 - 400 | 33 |
| 400 - 500 | 20 |
| 500 - 600 | 17 |

**Arithmetic Mean :**

$$\bar{x} = \frac{1}{N} \sum fx$$

| Income between (₹) | No. of Persons (f) | Mid Point x | fx |
|---|---|---|---|
| 100 - 200 | 15 | 150 | 2250 |
| 200 - 300 | 15 | 250 | 3750 |
| 300 - 400 | 33 | 350 | 11550 |
| 400 - 500 | 20 | 450 | 9000 |
| 500 - 600 | 17 | 550 | 9350 |
| | N = 100 | | Σfx = 35900 |

$$\bar{x} = \frac{1}{N}\sum fx = \frac{1}{100}(35900)$$

Arithmetic Mean = ₹(359.00)

**Median :**

$$Md = L + \left[\frac{\left(\frac{N}{2} - c.f.\right)}{f} \times h\right]$$

Where : $L$ = Lower limit of Median Class

$c.f.$ = cumulative frequency of the class preceding Median Class.

$f$ = frequency of Median Class.

$h$ = width of Median Class.

$N$ = Sum of the frequency or Total number of items in the series.

**Median Class :** Following is the method of finding the Median Class :

**(1)** Calculate cumulative frequency (c.f.)

**(2)** Look for c.f. "just greater" than N/2

**(3)** The corresponding Class is Median Class.

| Income between (₹) | No. of Persons (f) | c.f. |
|---|---|---|
| 100 - 200 | 15 | 15 |
| 200 - 300 | 15 | 30 (c.f.) |
| 300 - 400 | 33 (f) | 63 |
| M 400 - 500 | 20 | 83 |
| 500 - 600 | 17 | 100 |
| | N = 100 | |

**Median Class :**

$$\frac{N}{2} = \frac{100}{2} = 50$$

$c.f.$ just greater than 50 is 63

Hence, 300 – 400 is Median Class.
Therefore,

$L = 300$

$c.f. = 30$

$f = 33$

$h = 400 - 300 = 100$

Substituting the values in the formula :

$$Md = 300 + \left[\frac{(50-30)}{33} \times 100\right]$$

$$Md = 300 + \left[\frac{(20)}{33} \times 100\right]$$

$$Md = 300 + [66.67]$$

$$\text{Median} = ₹\,366.67$$

**Mode :**

$$Mo = L + \left[\left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times (h)\right]$$

Where;

$L$ : Lower Limit of Modal Class
$f_1$ : frequency of Modal Class
$f_0$ : frequency of class preceding Modal Class
$f_1$ : frequency of class succeeding Modal Class
$h$ : width of Modal Class

**Modal Class :** Class corresponding to highest frequency

| Income between (₹) | No. of Persons (f) |
|---|---|
| 100 - 200 | 15 |
| 200 - 300 | 15 ($f_0$) |
| 300 - 400 | 33 ($f_1$) |
| 400 - 500 | 20 ($f_2$) |
| 500 - 600 | 17 |
| | **N = 100** |

Because highest frequency is 33 hence, 300 – 400 is Modal Class

$L = 300$

$f_0 = 15$

$f_1 = 33$

$f_2 = 20$

$h = 100$

$$Mo = L + \left[\left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times (h)\right]$$

$$Mo = 300 + \left[\left(\frac{33-15}{66-15-20}\right) \times (100)\right]$$

$$Mo = 300 + \left[\left(\frac{15}{31}\right) \times (100)\right]$$

$$Mo = 300 + 48.39$$

Mode = ₹ 348.39

Therefore,

Arithmetic Mean = ₹ 359.00 **(Ans.)**

Median = ₹ 366.67 **(Ans.)**

Mode = ₹ 348.39 **(Ans.)**

❑❑

# MEASURES OF DISPERSION

**1. *Compare between mean deviation and standard deviation.*** **(2017, 2023-24)**

***Comparison of Mean Deviation and Standard Deviation***

| Mean Deviation | Standard Deviation |
|---|---|
| Calculating Mean Deviation algebraic signs are ignored | In calculating Standard Deviation algebraic signs are taken into account |
| Mean or Median is used to calculate Mean Deviation | Arithmetic Mean is used to calculate Standard Deviation |
| It is not possible to compute combined mean deviation | We can find the, combined standard deviation of two or more series. |
| If some wrong item is taken into account, we have to do the computations from the beginning to get the correct value of mean deviation. | If some wrong item is taken into account, we can correct the value of standard deviation without doing computations from the beginning. |
| Mean Deviation (MD) = $\frac{\Sigma\lvert x - AM\rvert}{n}$ or $\frac{\Sigma\lvert x - Md\rvert}{n}$ | Standard Deviation : $\sigma = \sqrt{\frac{\Sigma(x_1 - \bar{x})^2}{n}}$ |

**2. *Calculate Standard deviation and coefficient of variation from the following data :***

| Class Interval | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 |
|---|---|---|---|---|---|---|
| Frequency | 6 | 12 | 15 | 28 | 20 | 14 |

**(2023-24)**

Standard deviation, usually denoted by the letter σ (small sigma) of the Greek alphabet was first suggested by Karl Pearson as a measure of dispersion in 1893. It is defined as the positive square root of the arithmetic mean of the squares of the deviations of the given observations from their arithmetic mean.

[C.2]

| Class | Frequency | Mid Values (X) | d = x-A/h = x-35/10 A = 35, h = 10 | f.d | $f.d^2$ |
|---|---|---|---|---|---|
| 0 - 10 | 6 | 5 | -3 | -18 | 54 |
| 10 - 20 | 12 | 15 | -2 | -24 | 48 |
| 20 - 30 | 15 | 25 | -1 | -15 | 15 |
| 30 - 40 | 28 | 35=A | 0 | 0 | 0 |
| 40 - 50 | 20 | 45 | 1 | 20 | 20 |
| 50 - 60 | 14 | 55 | 2 | 28 | 56 |
| | N=95 | | | $\sum$f.d = -9 | $\sum f.d^2$ = 193 |

$$Mean\ \bar{x} = A + \frac{\sum fd}{n}.h$$

$= 35 + (-9/95)*10$

$= 35 + (-0.0947)10$

$= 35 + (-0.947)$

$= 35 - 0.947 = 34.052$

Population Standard deviation $\sigma = \sqrt{\dfrac{\sum f d^2 - \dfrac{(\sum f.d)^2}{n}}{n}}.h$

$$= \sqrt{\frac{193 - \frac{(-9)^2}{95}}{95}}.10$$

$$= \sqrt{\frac{193 - 0.8526}{95}}.10$$

$$= \sqrt{\frac{192.1474}{95}}.10$$

$$= \sqrt{2.0226}.10$$

$$= 1.4222.10$$

$$= 14.2218$$

Coefficient of Variation (Population) $= \left(\dfrac{\sigma}{\bar{x}}\right) * 100$

$$= \left(\frac{14.2218}{34.0526}\right) * 100$$

$$= 41.76\%$$

**3. *The smallest value in a set of observations is 20 with a range of 45. Find the largest observation and the co-efficient of range.*** **(2022-23)**

Range is the simplest possible measure of dispersion. It is the difference between the values of the extreme items of the series.

To find the largest observation and the coefficient of range, we can use the given information about the smallest value and the range.

**Smallest Value :** 20

The smallest value in the set of observations is given as 20.

**Range :** 45

The range is the difference between the largest and smallest values in a dataset. In this case, the range is given as 45.

To find the largest observation, we can add the range to the smallest value :

$$\text{Largest observation} = \text{Smallest value} + \text{Range}$$
$$= 20 + 45$$
$$= 65$$

Therefore, the largest observation in the dataset is 65.

$$\text{Coefficient of range} = \frac{L-S}{L+S}$$
$$= \frac{65-20}{65+20}$$
$$= 0.529$$

**4. *Calculate semi-interquartile range and coefficient of quartile deviation from the following data :***

| Marks | No. of students |
|---|---|
| 0-10 | 6 |
| 10-20 | 5 |
| 20-30 | 8 |
| 30-40 | 15 |
| 40-50 | 7 |

**(2022-23)**

Semi-inter-quartile range as the name suggests is the midpoint of the inter-quartile-range. In other words, it is one half of the difference between the third quartile and the first quartile. Symbolically, Semi-inter-quartile range

or

Quartile deviation $=Q_3-Q_1/2$

Where $Q_3$ and $Q_1$ stand for the upper and lower quartiles respectively.

$Q_1=(N/4)^{th}$ item

$Q_3=3(N/4)^{th}$ item

| Marks | No.of students | c.f |
|---|---|---|
| 0-10 | 6 | 6 |
| 10-20 | 5 | 11 |
| 20-30 | 8 | 19 |
| 30-40 | 15 | 34 |
| 40-50 | 7 | 41 |
| 50-60 | 6 | 47 |
| 60-70 | 3 | 50 |

$\therefore$ N = 50

$Q_1$= 50/4 = 12.5th item which lies in class interval 20-30

$Q_1=L_1+((N/4-c.f)/f)h$

Where $L_1$ = lower limit of class interval

c.f = cumulative frequency of preceding class

$f$ = corresponding frequency of class interval

$h$ = class interval difference

$Q_1=20+((50/4-11)/8)10$

= 20 + ((12.5-11)/8)10

= 20 + (1.875)

= 21.875

$Q_3=3(N/4)^{th}$ item

= 3*12.5

= 37.5 th item which lies in 40-50 class interval

$Q_3=L_1+((3N/4-c.f)/f)h$

= 40+ ((37.5-34)/7)10

= 40+5 = 45

Now

Semi-inter-quartile range $=(Q_3-Q_1)/2$

= (45-21.875)/2

= 11.5625

Coefficient of quartile deviation = $Q_3$-$Q_1$/$Q_3$+$Q_1$

= 45-21.875/45+21.875

= 23.125/66.875

= **0.3457**

5. ***Calculate standard deviation and coefficient of variation from the following data giving the age distribution of 542 members :*** **(2022-23)**

| Age group (in years) | No. of members |
|---|---|
| 20-30 | 3 |
| 30-40 | 61 |
| 40-50 | 132 |
| 50-60 | 153 |
| 60-70 | 140 |
| 70-80 | 51 |
| 80-90 | 2 |

| Age-group | No .of Members(f) | x(mid-values) | d(x-A) | fd | $d^2$ | $f d^2$ |
|---|---|---|---|---|---|---|
| 20-30 | 3 | 25 | -30 | -90 | 900 | 2700 |
| 30-40 | 61 | 35 | -20 | -1220 | 400 | 24400 |
| 40-50 | 132 | 45 | -10 | -1320 | 100 | 13200 |
| 50-60 | 153 | 55 | 0 | 0 | 0 | 0 |
| 60-70 | 140 | 65 | 10 | 1400 | 100 | 14000 |
| 70-80 | 51 | 75 | 20 | 1020 | 400 | 20400 |
| 80-90 | 2 | 85 | 30 | 60 | 900 | 1800 |
| | | N=542 | | $\sum$=150 | | $\sum$=76500 |

Let $A = 55$

**Standard deviation** $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$

$\sigma = \sqrt{76500/542} - (150/542)^2$

$= \sqrt{141.1439} - 0.076592094$

$= \sqrt{141.0673}$

$= 11.8771764$

Coefficient of variation (c.v) = $\sigma / mean * 100$

Now

Mean = $\sum fx / N$

| No. of members(f) | x(mid-values) | fx |
|---|---|---|
| 3 | 25 | 75 |
| 61 | 35 | 2135 |
| 132 | 45 | 5940 |
| 153 | 55 | 8415 |
| 140 | 65 | 9100 |
| 51 | 75 | 3825 |
| 2 | 85 | 170 |
| $\sum$=542 | | $\sum$=29660 |

Mean(X) = 29660/542 =54.72325

C.V = 11.88/54.72 *100

= 21.70

Therefore, the standard deviation of the age distribution is approximately 11.88, and the coefficient of variation is approximately 21.70%.

**6. *The smallest value in a set of observations is 10 with a range of 25. Find the largest observation and the co – efficient of range.*** **(2021)**

Given $R = 10$, $S = 25$ to find the largest observation $L$,

$$R = L - S$$

$$25 = L - 10$$

$$L = 35$$

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{35-10}{35+10} = \frac{25}{45} = 0.55$$ **(Ans.)**

**7. *Find standard deviation and coefficient of variation from the following data :*** **(2021)**

| Marks | No. of Students |
|---|---|
| 0-10 | 10 |
| 10-20 | 15 |
| 20-30 | 25 |
| 30-40 | 30 |
| 40-50 | 15 |
| 50-60 | 10 |
| 60-70 | 5 |

**Solve :**

| Marks | f | x | d=x-35 | Fd | $d^2$ | $fd^2$ |
|---|---|---|---|---|---|---|
| 0-10 | 10 | 5 | -30 | -300 | 900 | 9000 |
| 10-20 | 15 | 15 | -20 | -300 | 400 | 6000 |
| 20-30 | 25 | 25 | -10 | -250 | 100 | 2500 |
| 30-40 | 30 | 35 | 0 | 0 | 0 | 0 |
| 40-50 | 15 | 45 | 10 | 150 | 100 | 1500 |
| 50-60 | 10 | 55 | 20 | 200 | 400 | 4000 |
| 60-70 | 5 | 65 | 30 | 150 | 900 | 4500 |
| | N=110 | | | $\Sigma fd = -350$ | | $\Sigma fd^2 = 27500$ |

**Standard Deviation** $\sigma_x = \sqrt{\left[fd^2{}_x / N - (fd_x / N)^2\right]}$

$$\sqrt{27500/110-(-350/110)^2} = \sqrt{250-10.11} = 15.49$$

**Coefficient of Variation c.v.=**

$$(\sigma_x / x)*100 = (15.49/110)*100 = 14.08$$

**8.** ***Calculate mean deviation from median and its coefficient from the following data :*** **(2021)**

| Size | Frequency |
|---|---|
| 0-10 | 4 |
| 10-20 | 8 |
| 20-30 | 11 |
| 30-40 | 15 |
| 40-50 | 11 |
| 50-60 | 7 |
| 60-70 | 4 |

| Marks | f | cf | x | $\lvert x-M \rvert$ | $\lvert x-M \rvert$ | $f\lvert x-M \rvert$ |
|---|---|---|---|---|---|---|
| 0-10 | 4 | 4 | 5 | -29.67 | 29.67 | 118.68 |
| 10-20 | 8 | 12 | 15 | -19.67 | 19.67 | 157.36 |
| 20-30 | 11 | 23 | 25 | -9.67 | 9.67 | 106.37 |
| 30-40 | 15 | 38 | 35 | 0.33 | 0.33 | 4.95 |
| 40-50 | 11 | 49 | 45 | 10.33 | 10.33 | 113.63 |
| 50-60 | 7 | 56 | 55 | 20.33 | 20.33 | 142.31 |
| 60-70 | 4 | 60 | 65 | 30.33 | 30.33 | 121.32 |
| | $\Sigma f = N = 60$ | | | | | $\Sigma f\lvert x-\bar{x} \rvert = 764.62$ |

$$\Sigma f = 60$$

$$\Sigma f\lvert x-M \rvert = 764.62$$

$$N = 60, \text{ then,}$$

$\therefore$ Median class $=30-40$

Then we have, $l = 30, cf = 23, f = 15, h = 10$

$$\text{Median } (M) = l + \left(\frac{n/2-cf}{f}\right)$$

$$f = 30 + (30-23)/15$$

$$= 30 + 4.67 = 34.67$$

Mean deviation from median

$$= \frac{\Sigma f\lvert xi - M \rvert}{\Sigma f} = \frac{764.62}{60} = 12.74$$

Coefficient of Mean Deviation

$$= \frac{M.D}{M} = \frac{12.74}{34.67} = 0.37$$ **(Ans.)**

**9.** ***Find the range and coefficient of range for the following data :*** **(2019)**

| Variable | Frequency |
|---|---|
| 83 | 6 |
| 88 | 8 |
| 93 | 9 |
| 98 | 13 |
| 103 | 5 |
| 108 | 3 |

Range = Largest Observation − Smallest Observation

$= L - S$

Range $108 - 83 = 25$

Coefficient of range

$$= \frac{Largest\ Observation - Smallest\ Observation}{Largest\ Observation + Smallest\ Observation} = \frac{L-S}{L+S}$$

$$\text{Coefficient of Range} = \frac{108-83}{108+83} = \frac{25}{191} = 0.13$$

Range = 25

Coefficient of Range =0.13 **(Ans.)**

**10.** ***Compute quartile deviation and its coefficient from the information given below :*** **(2019)**

| Mid Value | Frequency |
|---|---|
| 5 | 5 |
| 15 | 8 |
| 25 | 15 |
| 35 | 20 |
| 45 | 16 |
| 55 | 10 |
| 65 | 6 |

Quartile deviation is based on the lower quartile $Q_1$ and the upper quartile $Q_3$. The difference $Q_3 - Q_1$ is called the inter quartile range. The difference $Q_3 - Q_1$ divided by 2 is called semi-inter-quartile range or the quartile deviation. Thus

$$Q.D = \frac{Q_3 - Q_1}{2}$$

## Coefficient of Quartile Deviation

A relative measure of dispersion based on the quartile deviation is called the coefficient of quartile deviation. It is defined as :

$$\text{Coefficient of Quartile Deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$\text{Quartiles : } Q_r = L + \left[\frac{\left(\frac{rN}{4} - c.f\right) * h}{f}\right],$$

Where $r = 1, 2, 3$, for First, Second and third quartile resp.

Where : $L$ = Lower Limit of Quartile Class

$c.f$ = c.f. of the class preceding Quartile Class

$f$ = frequency of Quartile Class

$h$ = width of Quartile Class

$N = \Sigma f$

## Steps in Finding rth Quartile Class

**(1)** Find $rN/4$

**(2)** Look for c.f. just greater than $rN/4$

**(3)** Corresponding Class is $r$th Quartile Class

**Calculating Semi-Interquartile Range and Coefficient of Quartile Deviation for the given Data :**

| Marks | No of Students | Less than cf |
|---|---|---|
| 0 - 10 | 5 | 5 |
| 10 – 20 | 8 | 13 |
| 20 – 30 | 15 | 28 ($Q_1$ Class) |
| 30 – 40 | 20 | 48 ($Q_2$ Class) |
| 40 – 50 | 16 | 64 ($Q_3$ Class) |
| 50 – 60 | 10 | 74 |
| 60 - 70 | 6 | 80 |
| | N= 80 | |

**First Quartile :** $Q_1 = L + \left(\frac{\frac{N}{4} - c.f}{f}\right) * h$

$Q_1$ class = class corresponding to c.f just greater than $N/4$

$$\frac{N}{4} = \frac{80}{4} = 20$$

c.f. just greater than 20 is 28, hence (20 – 30) is $Q_1$ class

Hence :

$$L = 20, c.f = 13, f = 15, h = 5 \text{ and } \frac{N}{4} = 20$$

Substituting the values in the formula and calculating we get :

$$Q_1 = 20 + \left(\frac{20-13}{15}\right) * 5$$

$$Q_1 = 20 + \left(\frac{7}{3}\right)$$

$$Q_1 = 20 + (2.33) = 22.33$$

$$Q_1 = 22.33$$

**Second Quartile :** $Q_2 = L + \left(\frac{\frac{2N}{4} - c.f}{f}\right) * h$

$Q_2$ =class corresponding to c.f. just greater than $2N/4$

$$\frac{2N}{4} = \frac{2(80)}{4} = 40$$

c.f. just greater than 40 is 48

Hence, 30-40 is $Q_2$ Class.

Therefore,

$L = 30$

$c.f = 28$

$f = 20$

$h = 25 - 30 = 5$

Substituting the values in the formula :

$$Q_2 = 30 + \left[\frac{(40-28)}{20} * 5\right]$$

$$Q_2 = 30 + \left[\frac{12}{20} * 5\right]$$

$$Q_2 = 30 + [3]$$

$$Q_2 = 33 \ Marks$$

(Note : Median and Second Quartile are same as II Quartile divides the series into two parts and Median also does the same)

Third Quartile : $Q_3 = L + \left(\frac{\frac{3N}{4} - c.f}{f}\right) * h$

$Q_3$ class = Class corresponding to c.f. just greater than $3N/4$

$$\frac{3N}{4} = \frac{3(80)}{4} = 60$$

c.f. just greater than 60 is 64, hence (40 – 50) is $Q_3$ class

Hence :

$$L = 40,\ c.f = 48, f = 16, h = 5 \text{ and } \frac{3N}{4} = 60$$

Substituting the values in the formula and calculating we get :

$$Q_3 = 40 + \left(\frac{60-48}{16}\right) * 5$$

$$Q_3 = 40 + \left(\frac{12}{16}\right) * 5$$

$$Q_3 = 40 + 3.75 = 43.75$$

$$Q_3 = 43.75$$

Hence :

**Quartiles :**

$Q_1 = 22.33\ Marks$

$Q_2 = 33\ Marks$

$Q_3 = 43.75\ Marks$

$$Q.D = \frac{43.75 - 22.33}{2} = 10.71$$

**Coefficient of Quartile Deviation =**

$$\frac{43.75 - 22.33}{43.75 + 22.33} = \frac{10.71}{66.08} = 0.162$$ **(Ans.)**

11. *Find standard deviation and coefficient of variation (C.V.) of the following data :* (2019)

| Class Interval | Frequency |
|---|---|
| 0 – 10 | 10 |
| 10 – 20 | 15 |
| 20 – 30 | 25 |
| 30 – 40 | 25 |
| 40 – 50 | 10 |
| 50 – 60 | 10 |
| 60 – 70 | 5 |

**Standard Deviation**

Standard Deviation $(\sigma) = \sqrt{\Sigma f(x)^2 / N - (\Sigma fx / N)^2}$

$x$ is the Mid point of class Interval = (Upper Limit + Lower Limit)/2

| Class Interval | Frequency | x (Mid Pt) | $x^2$ | fx | $fx^2$ |
|---|---|---|---|---|---|
| 0 – 10 | 10 | 5 | 25 | 50 | 250 |
| 10 – 20 | 15 | 15 | 225 | 225 | 3375 |
| 20 – 30 | 25 | 25 | 625 | 625 | 15625 |
| 30 – 40 | 25 | 35 | 1225 | 875 | 30625 |
| 40 – 50 | 10 | 45 | 2025 | 450 | 20250 |
| 50 - 60 | 10 | 55 | 3025 | 550 | 30250 |
| 60 - 70 | 5 | 65 | 4225 | 325 | 21125 |
| | N=100 | | | $\Sigma$fx =3100 | $\Sigma f(x)^2 = 121500$ |

**Standard Deviation** $(\sigma) = \sqrt{121500/100 - (3100/100)^2}$

$= \sqrt{121500/100 - (3100/100)^2}$

$= \sqrt{254}$

Standard Deviation $(\sigma) = 15.94$

Arithmetic mean

$(\bar{x}) = \dfrac{\Sigma fx}{N} = \dfrac{3100}{100} = 31$ **(Ans.)**

Coefficient of Variance $= \dfrac{\sigma}{x} * 100 = \dfrac{15.94}{31} * 100 = 51.42\%$

**(Ans.)**

**12. *Find the range and its coefficient for the following data :*** **(2018)**
**11, 9, 7, 10, 8, 6, 5, 7, 8**

Data given : 11, 9, 7, 10, 8, 6, 5, 7, 8

Range $(R) = L - S$ Where, $L$ = Largest observation
$= 11 - 5$ $S$ = Smallest observation

Range $(R) = 6$ **(Ans.)**

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{11-5}{11+5}$$

$$= \frac{6}{16} = \frac{3}{8} = 0.375$$ **(Ans.)**

**13. *Find standard deviation and its coefficient of variation (C.V.) from the following distribution :*** **(2016, 2018)**

| Age group (years) | No. of Workers |
|---|---|
| 25-30 | 10 |
| 30-35 | 12 |
| 35-40 | 25 |
| 40-45 | 40 |
| 45-50 | 10 |
| 50-55 | 3 |

***Standard Deviation***

$$\text{Standard Deviation } (\sigma) = \sqrt{\sum f(x)^2 / N - \left(\sum fx / N\right)^2}$$

$X$ is the Midpoint of Class Interval = (Upper Limit + Lower Limit) / 2

| Age Groups (yrs) | No. of Workers | x | $x^2$ | fx | $fx^2$ |
|---|---|---|---|---|---|
| 20 – 30 | 10 | 27.5 | 756.25 | 275 | 7562.5 |
| 30 – 35 | 12 | 32.5 | 1056.25 | 390 | 12675 |
| 35 – 40 | 25 | 37.5 | 1406.25 | 937.5 | 35156.3 |
| 40 – 45 | 40 | 42.5 | 1806.25 | 1700 | 72250 |
| 45 – 50 | 10 | 47.5 | 2256.25 | 475 | 22562.5 |
| 50 – 55 | 3 | 52.5 | 2756.25 | 157.5 | 8268.5 |
| | N = 100 | | Σfx = 3935 | | $\Sigma f(x)^2$ = 158475 |

$$\text{Standard Deviation } (\sigma) = \sqrt{158475/100 - (3935/100)^2}$$

$$= \sqrt{36.325}$$

$$\text{Standard Deviation } (\sigma) = 6.03$$ **(Ans.)**

$$\text{Coefficient of Variance} = \frac{\sigma}{\bar{x}} \times 100 = \frac{6.03}{39.35} \times 100$$
$$= 15.32\%$$ **(Ans.)**

**14. *On the basis of the following distribution, compute mean deviation from median :*** **(2018)**

| Marks | Frequency |
|---|---|
| 140-150 | 4 |
| 150-160 | 6 |
| 160-170 | 10 |
| 170-180 | 18 |
| 180-190 | 9 |
| 190-200 | 3 |

***Computation of M.D. about Median***

| Marks | M.V(x) | f | c.f. | \|x - Md\| | f \|x - Md\| |
|---|---|---|---|---|---|
| 140-150 | 145 | 4 | 4 | 27.78 | 111.12 |
| 150-160 | 155 | 6 | 10 | 17.78 | 106.68 |
| 160-170 | 165 | 10 | 20 | 7.78 | 77.80 |
| 170-180 | 175 | 18 | 38 | 2.22 | 39.96 |
| 180-190 | 185 | 9 | 47 | 12.22 | 109.98 |
| 190-200 | 195 | 3 | 50 | 22.22 | 66.66 |
| | | N = 50 | | | $\sum f\lvert x - Md\rvert = 512.2$ |

Median = Value of $\left(\frac{N}{2}\right)$ *th* item = Value of $\left(\frac{50}{2}\right)$ *th* item = Value of 25-th item.

This belongs to the class 170-180. Thus,
$L_1 = 170, i = 10, f = 18, c = 20$
Using the formula

$$\text{Median} = L_1 + \frac{i}{f}\left(\frac{N}{2} - c\right)$$
$$= 170 + \frac{10}{18}(25 - 20) = 170 + \frac{50}{18}$$
$$= 170 + 2.78 = 172.78$$

$$\text{M.D. (about median)} = \frac{1}{N}\sum_i f_i \lvert x_i - \text{Median}\rvert$$

i.e., $$= \frac{512.20}{50} = 10.24$$ **(Ans.)**

**15.** ***Calculate semi-interquartile range and coefficient of quartile deviation from the following data :***
**(2017)**

| Class | Frequency |
|---|---|
| 0-5 | 29 |
| 5-10 | 95 |
| 10-15 | 225 |
| 15-20 | 93 |
| 20-25 | 29 |
| 25-30 | 7 |
| 30-35 | 9 |
| 35-40 | 6 |
| 40-45 | 4 |
| 45-50 | 3 |

### *Calculating Semi-Inter quartile Range and Coefficient of Quartile Deviation for the Given Data*

| Class | Frequency(f) | Cumulative Frequency (c.f) |
|---|---|---|
| 0-5 | 29 | 29 |
| 5-10 | 95 | 124 |
| 10-15 | 225 | 349($Q_1$ class) |
| 15-20 | 93 | 442($Q_3$ class) |
| 20-25 | 29 | 471 |
| 25-30 | 7 | 478 |
| 30-35 | 9 | 487 |
| 35-40 | 6 | 493 |
| 40-45 | 4 | 497 |
| 45-50 | 3 | 500 |
| | N=500 | |

First Quartile : $Q_1 = L + \left[\dfrac{\left(\dfrac{N}{4} - c.f\right) \times h}{f}\right]$

$Q_1$ class = Class corresponding to $c.f$ just greater $\dfrac{N}{4}$

$$\frac{N}{4} = \frac{500}{4} = 125$$

$c.f$ just greater than 125 is 349, hence (10 - 15) is $Q_1$ class

Hence :

$$L = 10,\ c.f = 124,\ f = 225,\ h = 5 \text{ and } \frac{N}{4} = 125$$

Substituting the values in the formula and calculating we get :

$$Q_1 = 10 + \left(\frac{125-124}{225}\right) \times 5$$

$$Q_1 = 10 + \left(\frac{1}{225}\right) \times 5$$

$$Q_1 = 10 + (0.02) = 10.02$$

$$Q_1 = 10.02$$

Third Quartile : $Q_3 = L + \left(\frac{\frac{3N}{4} - c.f}{f}\right) \times h$

$Q_3$ Class = Class corresponding to $c.f$ just greater than $\frac{3N}{4}$

$$\frac{3N}{4} = \frac{3(500)}{4} = 375.$$

$c.f$ just greater than 375 is 442, hence (15 - 20) is $Q_3$ class

Hence :

$$L = 15,\ c.f = 349,\ f = 93,\ h = 5 \text{ and } \frac{3N}{4} = 375$$

Substituting the values in the formula and calculating we get :

$$Q_3 = 15 + \left(\frac{375-349}{93}\right) \times 5$$

$$Q_3 = 15 + \left(\frac{26}{93}\right) \times 5$$

$$Q_3 = 15 + 1.397 = 16.397$$

$$Q_3 = 16.397$$

**Hence :**

$Q_3 = 16.397$

$Q_1 = 10.02$

$Q_3 = 16.397$

Now, semi-interquartile range

$$\frac{(Q_3 - Q_1)}{2} = \frac{(16.397 - 10.02)}{2} = \frac{6.377}{2} = 3.19$$

Coefficient of Quartile Deviation :

$$\frac{(Q_3 - Q_1)}{Q_3 + Q_1} = \frac{(16.397 - 10.02)}{(16.397 + 10.02)} = \frac{6.377}{26.417} = 0.24$$

Therefore,

**Semi-interquartile range = 3.19**

**Coefficient of Quartile Deviation = 0.24**

16. ***In the production of certain rods, "a process is said to be in control if the outside diameters have a mean 2.5 and a S.D. of 0.002". Find the control limits for the mean of random sample of size 4.***
**(2017)**

$\mu_p = 2.5$ and $\sigma_p = 0.002, n = 4$

Control limits for $\overline{X}$ chart :

The central line (C.L.) = $\mu_p = 2.5$

Upper Control (UCL) =

$$= \mu_p + \frac{3\sigma_p}{\sqrt{n}} = 2.5 + \frac{3(0.002)}{\sqrt{4}} = 2.5 + 0.003 = 2.503$$

Upper Control limit (UCL) = 2.503

Lower Control Limit (LCL) =

$$= \mu_p - \frac{3\sigma_p}{\sqrt{n}} = 2.5 - \frac{3(0.002)}{\sqrt{4}} = 2.5 - 0.003 = 2.497$$

Lower Control Limit (LCL) = 2.497

Therefore, the control limits for the mean of the random samples of size 4 are :

The central line (C.L) = 2.5

Upper Control Limit (UCL) = 2.503

Lower Control Limit (LCL) = 2.497

17. ***Find the range of the following distribution of families according to its size (No. of children).***
**(2016)**

| Size of Family | No. of Families |
|---|---|
| 0 | 8 |
| 1 | 12 |
| 2 | 15 |
| 3 | 25 |
| 4 | 4 |
| 5 | 6 |

Range = Largest Observation – Smallest Observation

Range = 5 – 0 = 5

Therefore, the range of the following distribution of families according to its size (No. of children) = 5

**18.** ***The smallest value in a set of observation is 20 with a range of 25. Find the largest observation and the coefficient of range.*** **(2015)**

Range = Largest Observation – Smallest Observation

$$\text{Coefficient of Range} = \frac{\text{Largest Observation} - \text{Smallest Observation}}{\text{Largest Observation} + \text{Smallest Observation}}$$

Hence,

25 = Largest Observation – 20

Largest Observation = 25 + 20 = 45

$$\text{Coefficient of Range} = \frac{45-20}{45+20}$$

$$\text{Coefficient of Range} = \frac{25}{65} = 0.384$$

Therefore, Largest Observation = 45

Coefficient of Range = 0.384 **(Ans.)**

□□

est Observation

ıg distribution of
:en) = 5

*ɔation is 20 with*
*ervation and the*
**(2015)**

st Observation
ervation
Observation
servation
Observation

: 45

: 45

**(Ans.)**

□□

# PERMUTATIONS AND COMBINATIONS

**1.** ***Find the number of Permutations of the letters of the word 'MATHEMATICS'. How many of these begin with H and end with S.*** **(2023-24)**

The word 'MATHEMATICS' contains 11 letters, with the following repetitions :

M : 2 times

A : 2 times

T : 2 times

Total permutations = 11! /2! *2! *2! = 39916800/2*2*2
= 4989600

To find the number of permutations that begin with '*H*' and end with '*S*', we need to fix '*H*' at the beginning and '*S*' at the end. Then, we can permute the remaining 9 letters in between.

The remaining 9 letters are: *M, A, T, E, M, A, T, I, C*

Out of these, there are 2 M's, 2 A's, and 2 T's.

So, the number of permutations of these 9 letters is:

Permutations = 9! /2! ×2! ×2!

Let's calculate this.

$$\text{Permutations} = \frac{362880}{2\times2\times2} = \frac{362880}{8} = 45360$$

**2.** ***In how many ways can a student choose 5 courses out of 8 courses if 2 courses are compulsory for every student.*** **(2022-23)**

If 2 courses are compulsory for every student and they need to choose 5 courses out of 8 total courses, it means they have to select 3 additional courses from the remaining 6 courses.

To calculate the number of ways to choose 5 courses out of 8 with 2 compulsory courses, we can use the combination formula:

$$C(n, r) = n!/(r!(n-r)!)$$

where $n$ is the total number of courses and $r$ is the number of courses to be chosen.

In this case, $n = 6$ (remaining courses after selecting the 2 compulsory courses) and $r = 3$ (additional courses to be chosen).

So the number of ways to choose 5 courses out of 8 with 2 compulsory courses is :

$$C(6, 3) = 6!/(3!(6-3)!) = (6*5*4)/(3*2*1) = 20$$

Therefore, there are 20 ways for a student to choose 5 courses out of 8 courses if 2 courses are compulsory for every student.

**3. *There are 5 men and 4 ladies to dine at a round table. In how many ways can they seat themselves so that no two ladies are together?* (2022-23)**

Given :

Number of men = 5

Number of ladies = 4

**Concept Used :**

The total number of ways to arrange $n$ person $= (n - 1)!$

The number of all combinations of $n$ things, taken $r$ at a time $= {}^nC_r$

**Calculation :** According to the question, we have

The total number of ways to arrange 5 men = (5 - 1)!

$\Rightarrow 4!$

$\Rightarrow 4 \times 3 \times 2 \times 1$

$\Rightarrow 24$

The arrangement of 4 ladies $= {}^5c_4 \times 4!$

$= 5 \times 4!$

$= 5 \times 4 \times 3 \times 2 \times 1$

$= 120$

Now,

The total number of ways to arrangements $= 24 \times 120$

$= 2880$

Therefore, there are 2880 ways to seat themselves so that no two ladies are together.

**4. ◆ *If $1 \le r \le n$, prove that :***

$$C(n,r) + C(n,r-1) = C(n+1, r)$$

**(2022-23)**

◆ ***To prove that :*** **(2018)**
$\mathbf{C(n, r)+C(n, r-1)=C((n+1), r)}$

**LHS :** $C(n, r) = {}^{n}C_{r} = \dfrac{n!}{r!(n-r)!}$

$C(n, r-1) = {}^{n}C_{r-1} = \dfrac{n!}{(r-1)!(n-r+1)!}$

Adding them together requires a common denominator, which can be $r!(n-r+1)!$

$$= \frac{n!}{r!(n-r)!}\cdot\frac{(n-r+1)}{(n-r+1)} + \frac{n!}{(r-1)!(n-r+1)!}\cdot\frac{r}{r}$$

$$= \frac{n!(n-r+1)+n!r}{r!(n-r+1)!}$$

$$= \frac{n!(n-r+1+r)}{r!(n-r+1)!}$$

$$= \frac{n!(n+1)}{r!(n-r+1)!}$$

$$= \frac{(n+1)!}{r!(n+1-r)!}$$

$$= {}^{n+1}C_{r}$$

*i.e.* $= C((n+1), r)$ RHS (Hence Proved)

**5. *Find n, if :*** **(2016, 2021)**
$\mathbf{(n+1)! = 12 \times (n-1)!}$

$$(n+1)! = (n+1)n(n-1)! = 12(n-1)!$$

$$(n+1)n = 12$$

$$(n^2+n) = 12$$

$$(n^2+n-12) = 0$$

$$(n^2+4n-3n-12) = 0$$

$$n(n+4)-3(n+4) = 0$$

$$(n+4)(n-3) = 0$$

$$n = 3, \quad n = -4$$

'$n$' cannot be negative therefore $n = 3$

**6. *Find the number of permutations of the letters of the word 'EXAMINATION'?*** **(2021)**

Here there are 11 objects, out of which there are 2I's, 2N's, 2A's

The number of such arrangements is

$$\frac{11\,!}{2!\,2!\,2!}=\frac{11*10*9*8*7*6*5*4*3*2}{8}=4{,}989{,}600$$

**(Ans.)**

**7. *In how many ways can 10 books be arranged on a shelf so that a particular pair of books shall be :***

**(1) *Always together***

**(2) *Never together*** **(2021)**

Given that 10 books can be arranged on a shelf is 10 ! ways

**(1)** Pair of books always together i.e. 2 books will take 1 unit and remaining 8 books on 8 units then total number of units = 1 + 8 = 9.

These 9 can be arranged in 9! ways and two Pair of books can be arranged in 2! ways.

$\therefore$ Hence the total number of arrangements = $9!\times 2!$ ways.

**(2)** First 8 books can be arranged in 8! ways and there are 9 gaps and 2 books can be arranged in ${}^9P_2$ ways.

$\therefore$ Hence the total number of arrangements $= 8!\times {}^9P_2$ ways.

**8. *If $1\le r\le n$, prove that $c(n,r)+c(n,r+1)=c(n+1,\,r+1)$*** **(2021)**

$${}^nC_r+{}^nC_{r+1}={}^{n+1}C_{r+1}$$

$$\text{LHS}\Rightarrow {}^nC_r+{}^nC_{r+1}\qquad\left[\text{we know } {}^nC_r=\frac{n!}{(n-r)!\,r!}\right]$$

$$\Rightarrow \frac{n!}{(n-r)!\,r!}+\frac{n!}{(n-r-1)!\,(r+1)!}$$

$$\Rightarrow \frac{n!}{(n-r)(n-r-1)!\,r!}+\frac{n!}{(n-r-1)!\,(r+1)!\,r!}$$

Now we can written as :

$$\Rightarrow \frac{n!}{(n-r-1)!\,r!}\left(\frac{1}{(n-r)}+\frac{1}{(r+1)}\right)$$

$$\Rightarrow \frac{n!}{(n-r-1)!\,r!}\left(\frac{r+1+n-r}{(n-r)(r+1)}\right)$$

$$=\frac{(n+1)n!}{(n-r)(n-r-1)!(r+1)r!}$$

So, we can written

$$=\frac{(n+1)!}{(n-r)!(r+1)!}={}^{n+1}C_{r+1}$$

Hence, ${}^{n}C_{r}+{}^{n}C_{r+1}={}^{n+1}C_{r+1}$

LHS = RHS **(Proved)**

**9.** ***How many words can be formed from the letters of the word 'SUNDAY'? How many of these begin with D?*** **(2019)**

There are a total of 6 letters, so if all the letters were different, the number of arrangements would be 6! and there are no duplicates, so the number of words that can be formed with the word SUNDAY are :

$6! = 720$

Number of these words that begin with $D$ :

Fixing the first alphabet as $D$ the remaining alphabets are 5 so, the number of these words that begin with $D = 5! = 120$

720 different words can be formed with the letters of the word, SUNDAY' and 120 of these begin with D.

**10.** ***If ${}^{n}P_{r}=1680$ and ${}^{n}C_{r}=70$, find r.*** **(2019)**

Given ${}^{n}P_{r}=1680$ ...(1)

$$\frac{n!}{(n-r)!}=1680$$

And ${}^{n}C_{r}=70$ ...(2)

$$\frac{n!}{r!(n-r)!}=70$$

Substituting the $\dfrac{n!}{(n-r)!} = 1680$ in Equation 2 :

$$\frac{1680}{r!} = 70$$

$$r! = \frac{1680}{70} = 24$$

**(Ans.)**

$$r! = 24 \Rightarrow r = 4$$

**11.** ***Prove that :*** **$P(n,r) = P(n-1,r) + r.P(n-1, r-1)$ (Dec. 2013, 2019)**

We have : $P(n,r) = P(n-1,r) + r.P(n-1,r-1)$

$$\frac{(n-1)!}{[(n-1)-r]!} + \frac{r(n-1)!}{[(n-1)-(r-1)]!}$$

(because $P(n, r) = \dfrac{n!}{n-r!}$ )

$$= \frac{(n-1)!}{[n-r-1]!} + \frac{r(n-1)!}{[n-r]!}$$

$$= \frac{(n-1)!}{[n-r-1]!} + \frac{r(n-1)!}{[(n-r)\times(n-r-1)!]}$$

$$= \frac{(n-1)!}{[n-r-1]!}\left[1 + \frac{r}{n-r}\right] = \frac{(n-1)!}{[n-r-1]!}\left[\frac{n-r+r}{n-r}\right]\left[\frac{n-r+x}{n-r}\right]$$

$$= \frac{(n-1)!}{[n-r-1]!}\left[\frac{n-r+r}{n-r}\right] = \frac{(n-1)!}{[n-r-1]!}\left[\frac{n}{n-r}\right]$$

$$= \frac{n(n-1)!}{n-r(n-r-1)!} = \frac{n!}{n-r!} = P(n, r)$$ **(Hence Proved)**

**12.** ***Find the number of ways of which 5 boys and 3 girls can be seated in a row so that no two girls are together.*** **(2019)**

Five Boys can be placed in 5! ways

Set the positions for boys first

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | B | | B | | B | | B | | B | |

Now see the Blank Spaces. At these spaces Girls can be placed fulfilling the requirement of the question.

So there are 3 Girls and 6 places for Girls to occupy.
It can be done in 6*5*4 ways = 120.
So total no. of ways = 5! * 120 = 120*120 = 14,400 ways

**13. A box contains 4 different red and 6 different white balls. In how many ways can 6 balls be selected so that there are at least two balls of each colour? (2019)**

For selecting 6 balls so that atleast two balls of each colour

$= {}^5C_2 \times {}^6C_4 + {}^5C_3 \times {}^6C_3 + {}^5C_4 \times {}^6C_2$

$= 10 \times 15 + 10 \times 20 + 5 \times 15$

$= 150 + 200 + 75$

$= 425$

In 424 ways 6 balls can be selected so that there are at least two balls of each colour.

**14. What is the chance of throwing a total of 5 or 10 with two dice? (2019)**

When two dices are thrown total outcomes $n(S) = 36$

Let $A$ be the event of getting a total of 10 with two dice :

Favorable out comes for getting 10 are (6,4)(4,6)(5,5)

$P(A) = 3/36$

Let $B$ be the event of getting a total of 5 with two dice

Favorable outcomes for getting 5 are (4,1)(1,4)(2,3)(3,2)

$P(B) = 4/36$

Both are mutually exclusive events, therefore using the Addition Theorem of probability :

**Addition Theorem on Probability**

If $A$ and $B$ are any two events then the probability of happening of at least one of the events is defined as

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

If $A$ and $B$ are mutually exclusive then,

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

Hence the required probability of getting a total of 5 or 10 with two dice is :

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

$P(A \cup B) = 3/36 + 4/36 = 7/36$ **(Ans.)**

**15. *How many permutations can be made out the letters of the word 'TRIANGLE'. How many of these will begin with T and end with E?*** **(2018)**

(1) The word 'TRIANGLE' consists of 8 different letters. They can be arranged among themselves in $P$ (8, 8) ways = 8! = 40320 ways. Hence, the total number of permutations = 40320 ways. **(Ans.)**

(2) If $T$ is fixed at the first place and $E$ at the last place, then we are to fill up the remaining 6 places by the remaining 6 letters which can be done in $P$ (6, 6) = 6! = 720 ways. Hence, the total number of such permutations is 720. **(Ans.)**

**16. Find $n$, if :** **(2018)**

$^nC_7 = {}^nC_9$

$$^nC_7 = {}^nC_9$$

$$^nC_r = {}^nC_{n-r}$$

$$^nC_7 = {}^nC_{n-9}$$

$$7 = n - 9$$

$$n = 16$$

**(Ans.)**

**17. *If two balls are drawn from a bag containing 2 white, 4 red and 5 black balls. What is the chance that both the balls are red?*** **(2018)**

No. of white balls = 2
No. of Red balls = 4
No. of Black balls = 5

There are 11 balls in the bag out of which 2 balls can be drawn in $^{11}C_2$ ways.

So, total number of elementary events = $^{11}C_2$

$$= \frac{11!}{2!(11-2)!}$$

$$= \frac{11!}{2!9!}$$

$$= \frac{11.10.9!}{2.1.9!}$$

$$= 55$$

There are 4 red balls out of which 2 balls can be drawn in ${}^4C_2$ ways.

$\therefore$ Favorable number of elementary events $= {}^4C_2$

$$= \frac{4!}{2!2!}$$

$$= \frac{4.3.2!}{2.1.2!}$$

$$= 6$$

So, required probability is $\frac{6}{55}$. **(Ans.)**

18. ***In how many arrangements of word 'GOLDEN' will the vowels never occur together?*** **(2018)**

Given word : GOLDEN

Number of vowels = 2

Number of consonants = 4

4 consonants can be arranged first in 4! Ways

Now there are 5 places to arrange 2 vowels

It can be done in ${}^5P_2$ ways.

Number of Arrangements $= 4! \times {}^5P_2$

$$= 4.3.2.1 \times \frac{5!}{(5-2)!}$$

$$= 24 \times 20 = 480 \text{ ways.}$$ **(Ans.)**

19. ***Find the number of permutations of the letters of the word 'English'. How many of these begin with E and end with I?*** **(2016, 2017)**

The permutations of a word is given by

$$\frac{(\text{Total number of alphabets})!}{\Pi(\text{Repetitions of each letter})!}$$

Since ENGLISH has 7 letters and none is getting repeated hence the number of permutations of the letters of the word ENGLISH = 7!

= 5040 ways.

To know how many of them begins with $E$ and ends with $I$:

First and last place of each word is reserved for letters $E$ and $I$ respectively. The remaining places can be filled up by remaining 5 letters. This can be done in 5! Ways.

5! = 120 permutations.

**20.** ***Find the number of arrangements of a multi-set objects some of which are alike say $n_1$ are alike of first kind, $n_2$ are alike of second kind,...... $n_r$ are alike of rth kind.*** **(2017)**

Lets solve the question taking an example.

For example: We can arrange $A_1A_2B$ in 3! ways. These are

$A_1A_2B$, $A_1BA_2$, $BA_1A_2$

$A_2A_1B$, $A_2BA_1$, $BA_2A_1$

However, as soon as we remove the subscripts on the $A$'s the second row is the same as the first row. i.e., we have only 3 distinct arrangements since each arrangement appears twice as the $A_1$ and $A_2$ are interchanged. In general, there would be $n!$ Arrangements if all $n$ objects were distinct. However, each arrangement would appear $n_1!$ times as the 1st type was interchanged with itself, $n_2!$ times as the 2nd type was interchanged with itself, etc. Hence only

$(n!)/(n_1!n_2! \cdots n_r!)$ of the $n!$ arrangements are distinct.

**21.** ***If $P_n^r = 720$ and $^nC_r = 120$ find r.*** **(2016)**

$$P(n,r) = \frac{n!}{(n-r)!} = 720 \quad \ldots (1)$$

$$C(n,r) = \frac{n!}{r!(n-r)!} = 120 \quad \ldots (2)$$

Dividing equation (1) by equation (2) we get :

$$\frac{\dfrac{n!}{(n-r)!}}{\dfrac{n!}{r!(n-r)!}} = \frac{720}{120}$$

$$r! = 6$$

$$\Rightarrow \quad r = 3$$

**(Ans.)**

**22.** ***To prove that 0! = 1.*** **(2015)**

We can also use the following pattern. We know that

$$n! = n(n-1)(n-2)(n-3)...3(2)(1)$$

Which means that

$$n! = n(n-1)!$$

Dividing both sides of the equation by $n$, we have.

$$(n-1)! = \frac{n!}{n}$$

Using this fact, we can check the following pattern.

$$4! = \frac{5!}{5} = \frac{(5)(4)(3)(2)(1)}{5} = 24$$

$$3! = \frac{4!}{4} = \frac{(4)(3)(2)(1)}{4} = 6$$

$$2! = \frac{3!}{3} = \frac{(3)(2)(1)}{(3)} = 2$$

$$1! = \frac{2!}{2} = \frac{(2)(1)}{(2)}$$

Now, we go to 0!

$$0! = \frac{1!}{1} = 1$$

**(Hence Proved)**

**23.** ***If $^{15}C_{11} + {}^{11}C_X = {}^{16}C_{12}$ find x. This question is incorrect. The correct question may be $^{15}C_{11} + {}^{11}C_X = {}^{16}C_{12}$*** **(2015)**

We know that ...(1)

$$^nC_r + {}^nC_{r+1} = {}^{n+1}C_{r+1} \quad ...(2)$$

Given $^{15}C_{11} + {}^{15}C_x = {}^{16}C_{12}$

Comparing Equation (1) and (2)

We deduce $x = r + 1$
Where $r = 11$
Thus, $x = 11 + 1 = 12$
Hence $x = 12$ **(Ans.)**

**24. *If $1 < r < n$ prove that***
***$^nC_r + ^nC_{r+1} = ^{n+1}C_{r+1}$*** **(2015)**

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

$$^nC_{r+1} = \frac{n!}{(n-(r+1))!r+1!}$$

$$^{n+1}C_{r+1} = \frac{n+1!}{((n+1)-(r+1))!r+1!} = \frac{n+1!}{(n-r)!r+1!}$$

**L.H.S.**

$$^nC_r + ^nC_{r+1} = \frac{n!}{(n-r)!r!} + \frac{n!}{(n-(r+1))!(r+1)!}$$

$$= \frac{n!}{(n-r)!r!} + \frac{n!}{(n-r-1)!(r+1)!}$$

$$= \frac{n!(r+1)}{(n-r)!r!(r+1)} + \frac{n!(n-r)}{(n-r-1)!(r+1)!(n-r)}$$

$$= \frac{n!(r+1)}{(n-r)!(r+1)!} + \frac{n!(n-r)}{(r+1)!(n-r)!}$$

$$= \frac{n!(r+1)+n!(n-r)}{(n-r)!(r+1)!}$$

$$= \frac{n![(r+1)+(n-r)]}{(n-r)!(r+1)!} = \frac{n![r+1+n-r]}{(n-r)!(r+1)!}$$

$$= \frac{n!(n+1)}{(n-r)!(r+1)!}$$

$$= \frac{(n+1)!}{(n-r)!(r+1)!} = ^{n+1}C_{r+1} = \text{R.H.S}$$

Therefore, $^nC_r + ^nC_r + 1 = n + ^1C_r + 1$ **(Hence Proved)**

**25. *If $^nP_4 = 2 \times ^5P_3$ find n.*** **(2015)**

We know that $= ^nP_r = \dfrac{(n)!}{(n-r)!}$

Therefore,

$$^{n}P_4 = \frac{(n)!}{(n-4)!} = 2 \times \frac{(5)!}{(5-3)!} = \frac{2(5)!}{(2)!} = \frac{2(5)!}{(2\times 1)} = 5!$$

$$\frac{(n)!}{(n-4)!} = \frac{n(n-1)(n-2)(n-3)(n-4)!}{(n-4)!}$$

$$= n\,(n-1)(n-2)(n-3) = 5!$$

$$n\,(n-1)(n-2)(n-3) = 5!$$

$$n\,(n-1)(n-2)(n-3) = 5 \times (5-1)(5-2)(5-3)$$

Therefore, if $^{n}P_4 = 2 \times {}^{5}P_3;\ n = 5$ **(Ans.)**

**26. *Evaluate :* $^{12}P_4$, $^{75}P_2$, $^{8}P_8$.** **(2014)**

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

**(1)** $$^{12}P_4 = \frac{12!}{(12-8)!} = \frac{12!}{(8)!} = \frac{12\times 11\times 10\times 9\times 8!}{(8)!}$$

$$= 12\times 11\times 10\times 9 = 11{,}880$$ **(Ans.)**

**(2)** $$^{75}P_2 = \frac{75!}{(75-2)!} = \frac{75!}{(73)!} = \frac{75\times 74\times 73!}{73!}$$

$$= 75\times 74 = 5{,}550$$ **(Ans.)**

**(3)** $$^{8}P_8 = \frac{8!}{(8-8)!} = \frac{8!}{(0)!} = \frac{8!}{(1)!} = 8\times 7\times 6\times 5\times 4\times 3\times 2\times 1$$

$$= 40{,}320$$ **(Ans.)**

**27. *How many words with or without meaning can be formed by using all the letters of the word DELHI, using each letter exactly once?*** **(2014)**

The word 'DELHI' has 5 letters and all these letters are different. Total words (with or without meaning) formed by using all these 5 letters using each letter exactly once = Number of arrangements of 5 letters taken all at a time

$$^{5}P_5 = \frac{5!}{(5-5)!} = \frac{5!}{(0)!} = \frac{5!}{1} = 5! = 5\times 4\times 3\times 2\times 1 = 120$$ **(Ans.)**

**28. *A committee of 5 is to be formed out of 6 gentlemen and 4 ladies. In how many ways this can be done when :*** **(2014)**

**(1) *At least 2 ladies are included?***

**(2) *At most 2 ladies are included?***

The number of all combinations of n things, taken r at a time is :

$$^{n}C_{r}=\frac{n!}{r!\times(n-r)!}$$

**(1) At Least Two Ladies are Included :** We may have (2 ladies and 3 gentlemen) or (3 ladies and 2 gentlemen) or (4 ladies and 1 gentleman).

Therefore, required number of ways are :

$$=(4C2\times 6C3)+(4C3\times 6C2)+(4C4\times 6C1)$$

$$=\left(\frac{4!}{2!\times(2)!}\times\frac{6!}{3!\times(3)!}\right)+\left(\frac{4!}{3!\times(1)!}\times\frac{6!}{2!\times(4)!}\right)+\left(\frac{4!}{4!\times(0)!}\times\frac{6!}{1!\times(5)!}\right)$$

$$=(6\times 20)+(4\times 15)+(1\times 6)$$

$$=120+60+6=186 \text{ ways}$$ **(Ans.)**

**(2) At Most 2 Ladies are Included :** We may have (2 ladies and 3 gentlemen) or (1 lady and 4 gentlemen) or (0 lady and 5 gentlemen).

Therefore, required number of ways are :

$$\left(^{4}C_{2}\times^{6}C_{3}\right)+\left(^{4}C_{1}\times^{6}C_{4}\right)+\left(^{4}C_{0}\times^{6}C_{5}\right)$$

$$=\left(\frac{4!}{2!\times(2)!}\times\frac{6!}{3!\times(3)!}\right)+\left(\frac{4!}{3!\times(1)!}\times\frac{6!}{2!\times(4)!}\right)+\left(\frac{4!}{4!\times(0)!}\times\frac{6!}{1!\times(5)!}\right)$$

$$=(6\times 20)+(4\times 15)+(1\times 6)$$

$$=120+60+6=186 \text{ ways}$$ **(Ans.)**

**Note :** The above problem is the case of $^{n}C_{r}={}^{n}C_{n-r}$

□□

# Sample Space, Events and Probability

**1. *If P(A) = 0.3 and* $P(A \cup B) = 0.7$. *Find P(B), If A and B are mutually exclusive.* (2023-24)**

Since $A$ and $B$ are mutually exclusive, $P(A \cap B) = 0$.

Therefore, $P(B) = P(A \cup B) - P(A) = 0.7 - 0.3 = 0.4$

So the answer is 0.4

**2. *A bag contains 8 balls of which 5 are white and 3 are black. Two balls are drawn at random. What is the probability that both are white?* (2023-24)**

The total number of ways in which, two balls can be drawn out of 8 is $= {}^8C_2 = \frac{8!}{2!6!} = 28$.

The number of ways in which two white balls can be drawn out of five $= {}^5C_2 = \frac{5}{2!\,3!} = 10$.

Therefore, the desired probability $= {}^5C_2 / {}^8C_2$

${}^5C_2 = 5 * 4 / 2 * 1 = 20 / 2 = 10$

${}^8C_2 = 8 * 7 / 2 * 1 = 56 / 2 = 28$

$P = 10 / 28 = 0.35$

**3. *Find the number of ways in which 5 boys and 3 girls can be seated in a row so that no two girls are together.* (2023-24)**

We have been given 5 boys and 3 girls. We have to arrange them in such a way that no 2 girls sit together. Thus, we can say it as,

_B_B_B_B_B_

Let us take boy as $B$ and girl as $G$.

Now let us find all the cases of filling 3 girls.

| | | | | |
|---|---|---|---|---|
| GBGBGBBB | GBBGBBGB | BGBGBGBB | BGBBBGBG | GBGBBGBB |
| GBBGBBBG | BGBGBBGB | BBGBGBGB | GBGBBBGB | GBBBGBGB |
| BGBGBBBG | GBGBBBBG | BBGBGBBG | GBBBGBBG | BGBBGBGB |
| BBGBBGBG | GBBGBGBB | GBBBBGBG | BGBBGBBG | BBBGBGBG |
| BBBGBGBG | | | | |

Thus counting all the cases there are 20.

Now we have been given 5 boys. Now these 5 boys can be arranged in 5! Ways.

$\therefore$ Ways to arrange boys= $5! = 5 \times 4 \times 3 \times 2 = 120$ ways

Ways to arrange girls $= 3! = 3 \times 2 \times 1 = 6$ ways

$\therefore$ Total no of ways = 20 $\times$ways to arrange boys $\times$ ways to arrange girls$\Rightarrow 20 \times 120 \times 6 = 14400$ total ways

Thus we have arranged 5 boys and 3 girls in 14400 ways in such a way that no 2 girls are together.

**4. *Three group of children contain respectively 3 girls and 1 boy, 2 girls and 2 boys and 1 girl and 3 boys. One child selected at random from each group. Show that the chance that the three selected consist of 1 girl an d2 boys is $\frac{13}{32}$.* (2023-24)**

| Group 1 | | Group 2 | | Group 3 | |
|---|---|---|---|---|---|
| Girls | Boys | Girls | Boys | Girls | Boys |
| 3 | 1 | 2 | 2 | 1 | 3 |

Let $G_1, G_2, G_3$ denote events for selecting a girl and $B_1, B_2, B_3$ denotes event for selecting a boy from 1st, 2nd and 3rd groups respectively.

Then

$$P(G_1) = 3/4$$
$$P(G_2) = 2/4$$
$$P(G_3) = 1/4$$

And

$$P(B_1) = 1/4$$
$$P(B_2) = 2/4$$
$$P(B_3) = 3/4$$

Where $G_1, G_2, G_3$ and $B_1, B_2, B_3$ are mutually exclusive events.

Let $E$ be the event that 1 girl and 2 boys are selected.

Therefore

$$E = (G_1 \cap B_2 \cap B_3) \cup (B_1 \cap G_2 \cap B_3) \cup (B_1 \cap B_2 \cap G_3)$$

Therefore

$$P(E) = P(G_1 \cap B_2 \cap B_3) + P(B_1 \cap G_2 \cap B_3) + P(B_1 \cap B_2 \cap G_3)$$

$$= P(G_1) \times P(B_2) \times P(B_3) + P(B_1) \times P(G_2) \times P(B_3) + P(B_1) \times P(B_2) \times P(G_3)$$

$$= \left(\frac{3}{4} \times \frac{2}{4} \times \frac{3}{4}\right) + \left(\frac{1}{4} \times \frac{1}{2} \times \frac{3}{4}\right) + \left(\frac{1}{4} \times \frac{2}{4} \times \frac{1}{4}\right) = \frac{26}{64} = \frac{13}{32}$$

**5. *How many different words can be formed with the letters of the word 'MISSISSIPPI'?* (2022-23)**

There are a total of 11 letters, so if all the letters were different, the number of arrangements would be11!

However, there are a lot of duplicates. There are 4 S's, 2 P's and 4 I's. So you have to divide the total number by (4! × 2! × 4!).

i.e. 11!/(4!4!2!)=34650. **(Ans.)**

**6. *If two balls are drawn from a bag containing 2 white, 4 red and 5 black balls, what is the chance that both the balls are black?* (2022-23)**

To calculate the probability of drawing two black balls from the given bag, we need to consider the total number of balls and the number of black balls in the bag.

Total number of balls in the bag = 2 (white) + 4 (red) + 5 (black) = 11

Number of black balls in the bag = 5

To calculate the probability, we need to find the ratio of the number of favorable outcomes (drawing two black balls) to the total number of possible outcomes (drawing any two balls).

The probability of drawing the first black ball is 5/11 because there are 5 black balls out of 11 total balls.

After drawing the first black ball, there are now 10 balls left in the bag. The probability of drawing a second black ball, given that the first ball was black, is 4/10 since there are 4 black balls remaining out of the 10 remaining balls.

To find the probability of both events occurring (drawing two black balls), we multiply the probabilities:

Probability of drawing two black balls = (5/11) * (4/10) = 20/110 = 2/11

Therefore, the chance of drawing both balls as black is 2/11.

**7. *A pair of dice is tossed twice. Find the probability of scoring 7 points :***
**(1) *once***
**(2) *at least once*** **(2022-23)**

To find the probability of scoring 7 points once when a pair of dice is tossed twice, we need to consider the different ways in which we can achieve this outcome.

**(1) Scoring 7 Points Once :** Let's consider the possibilities for getting a sum of 7 points when rolling two dice:

(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)

Now, let's consider the possible outcomes for the first toss and the second toss. We can either score 7 in the first toss and not score 7 in the second toss or vice versa. Therefore, there are two favorable outcomes.

The total number of possible outcomes when tossing a pair of dice twice is 6 * 6 = 36 (since each die has 6 possible outcomes).

Therefore, the probability of scoring 7 points once is 1/6*1/5 + 1/6*1/5 = 10/36 = 5/18

**(2) To Find the Probability of Scoring 7 Points at Least Once :** There are two possibilities to score 7 points at least once: scoring 7 on the first toss or scoring 7 on the second toss.

Probability of scoring 7 on the first toss = Favorable outcomes / Total outcomes = 6 / 36 = 1 / 6

Probability of scoring 7 on the second toss = Favorable outcomes / Total outcomes = 1 / 6 (since there are only 6 possible outcomes, and one of them is favorable)

To find the probability of scoring 7 at least once, we can subtract the probability of not scoring 7 in both tosses from 1 (since the sum of probabilities of all possible outcomes is 1).

Probability of not scoring 7 in both tosses = (5/6) * (5/6) = 25/36 (probability of not rolling a 7 on the first toss multiplied by the probability of not rolling a 7 on the second toss)

Probability of scoring 7 at least once = 1 - Probability of not scoring 7 in both tosses = 1 - 25/36 = 11/36

Therefore, the probability of scoring 7 points at least once when tossing a pair of dice twice is 11/36.

**8. *A problem in statistics is given to the three students A, B and C. Whose chance of solving it are $\frac{1}{2}, \frac{2}{3}, \frac{1}{4}$ respectively. What is the probability that the problem is solved?*** **(2022-23)**

To find the probability that the problem is solved, we need to consider the individual probabilities of each student solving it and calculate the overall probability.

***Let's Denote the Events***

**A :** Student A solves the problem.

**B :** Student B solves the problem.

**C :** Student C solves the problem.

The probability of each student solving the problem is given as follows :

$P(A) = 1/2$

$P(B) = 2/3$

$P(C) = 1/4$

To calculate the probability that at least one student solves the problem, we can use the complement rule, which states that the probability of an event occurring is equal to 1 minus the probability of its complement (the event not occurring).

The probability that none of the students solves the problem is given by :

$P(A') = 1 - P(A) = 1 - 1/2 = 1/2$

$P(B') = 1 - P(B) = 1 - 2/3 = 1/3$

$P(C') = 1 - P(C) = 1 - 1/4 = 3/4$

Now, let's calculate the probability that none of the students solve the problem :

$P(A' \cap B' \cap C') = P(A') * P(B') * P(C')$

$= (1/2) * (1/3) * (3/4) = 1/8$

Using the complement rule, the probability that at least one student solves the problem is:

$P(A \cup B \cup C) = 1 - P(A' \cap B' \cap C') = 1 - 1/8 = 7/8$

Therefore, the probability that the problem is solved by at least one student is 7/8.

**9. *A bag contains tickets numbered from 1 to 25. Two tickets are drawn. Find probability that both the numbers are prime.* (2021)**

Total number of tickets = 20

If two tickets are drawn, the total number of cases $= {}^{20}C_2$

Favorable Cases : There are 8 prime numbers from 1 to 20. {2, 3, 5, 7, 11, 13, 17, 19}

Total number of favorable cases $= {}^{8}C_2$

Hence, the required probability $= {}^{8}C_2 / {}^{20}C_2$

$= 28/190$

$= 14/95$. **(Ans.)**

**10. *A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is 1/7 and that of wife's selection is 1/5. What is the probability that***

***(1) Both of them will be selected***

***(2) Only one of them will be selected*** **(2021)**

**A :** Event of husband being selected

**B :** Event of wife being selected

Then $P(A) = 1/7$, $P(B) = 1/5$

$P(\overline{A}) = 1 - 1/7 = 6/7 \quad P(\overline{B}) = 1 - 1/5 = 4/5$

(1) $P$ (Both are selected ) $= P(\overline{A}) = 1/7 * 1/5 = 1/35$

(2) $P$ (only one is selected ) = P (A selected)* P (B not selected) + P (A not selected) * P (B selected)

$= P(A) * P(\overline{B}) + P(\overline{A}) * P(\overline{B})$

$= 1/7 * 4/5 + 6/7 * 1/5$

$= 10/35$

$= 2/7$

**(Ans.)**

**11.** ***A bag contains 4 white and 2 black balls and a second bag contains 3 of each colour. A bag is selected at random, and a ball is drawn at random from the bag chosen. What is the probability that the ball drawn is white?*** **(2021)**

Let $E$ denote the event of drawing white ball

$$\Rightarrow P(E) = P(B1)P(E/B1) + P(B2)P(E/B2)$$

$$= \frac{1}{2} * \frac{4}{6} + \frac{1}{2} * \frac{3}{7}$$

$$= \frac{1}{2}\left[\frac{2}{3} + \frac{3}{7}\right]$$

$$= \frac{1}{2}\left[\frac{14+9}{21}\right]$$

$$= \frac{23}{42}$$

**(Ans.)**

**12.** ***Explain the following :*** **(2019)**

**(1)** ***Mutually exclusive event***

**(2)** ***Conditional probability***

**(1)** **Mutually Exclusive Events :** In probability theory, events $E_1, E_2, \ldots E_n$ are said to be mutually exclusive if the occurrence of any one of them automatically implies the non-occurrence of the remaining $n - 1$ events. Therefore, two mutually exclusive events cannot both occur. Formally said, the intersection of each two of them is empty (the null event) : $A$ and $B = \Phi$. In consequence, mutually exclusive events have the property: $P$ ($A$ and $B$) = 0.

For example, one cannot draw a card that is both red and a club because clubs are always black. If one draws just one card from the deck, either a red card or a black card (club or spade) can be drawn. When A and B are mutually exclusive,

$P(A \text{ or } B) = P(A) + P(B)$.

An example is tossing a coin once, which can result in either heads or tails, but not both.

In the coin-tossing example, both outcomes are collectively exhaustive, which means that at least one of the outcomes must happen, so these two possibilities together exhaust all the possibilities. However, not all mutually exclusive events are collectively exhaustive.

For example, the outcomes 1 and 4 of a single roll of a six-sided die are mutually exclusive (cannot both happen) but not collectively exhaustive (there are other possible outcomes; 2, 3, 5, 6).

(2) **Conditional Probability :** Conditional probability is the probability of some event $A$, given the occurrence of some other event $B$. Conditional probability is written $P(A \mid B)$, and is read "the probability of $A$, given $B$". It is defined by :

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

If $P(B) = 0$ then $P(A \mid B)$ is formally undefined by this expression. However, it is possible to define a conditional probability for some zero-probability events using a σ-algebra of such events (such as those arising from a continuous random variable).

**13. *A problem in statistics is given to the three students A, B and C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{3}{4}$ respectively. What is the probability that the problem will be solved?*** **(2019)**

Probability that $A$ is able to solve = $P(A)$ = (1 / 2).
Probability that $B$ is able to solve = $P(B)$ = (1 / 4).
Probability that $C$ is able to solve = $P(C)$ = (3 / 4).
Probability that $A$ is not able to solve = $P(a)$ = 1 - $P(A)$ = (1 / 2).
Probability that $B$ is not able to solve = $P(b)$ = (3 / 4).
Probability that $C$ is not able to solve = $P(c)$ = (1 / 4).

When $A$, $B$ and $C$ are independent events; the probability that none of them will be able to solve if all of them try independently = $P(a)*P(b)*P(c)$ = (1 / 2)*(3 / 4)*(1 / 4) = (3 / 32).

Therefore, the probability that the problem will be solved if all of them try independently = $1 - P(a)*P(b)*P(c)$ = (29 / 32).

**(Ans.)**

**14.** ***A bag contains 3 red and 4 black balls and a second bag contains 5 black balls and 4 red balls. A ball is drawn from each bag. Find the probability that one is red and other is black.*** **(2019)**

Probability of getting a red ball from bag 1 = 3/7
Probability of getting a black ball from bag 1 = 4/7
Probability of getting a red ball form bag 2 = 4/9
Probability of getting a black ball from bag 2 = 5/9
Probability of getting a red ball form bag 1 and a black ball from bag 2

= 3/7*5/9 = 15/63 = 5/21

Probability of getting a black ball from bag 1 and red ball from bag 2 = 4/7* 4/9 = 16/63

Hence the required probability that one is red and other is black ball is 5/21 + 16/63 = 31/63 **(Ans.)**

**15.** ***There are 5 men and 5 ladies to dine at a round table. In how many ways can they seat themselves so that no two ladies are together?*** **(2016, 2018)**

In a Circular arrangement, Total Permutation is $(n-1)!$

Here we don't want women to sit together, hence we follow these steps.

**Step 1 :** Fix all the boys first around the table. This can be done in $(5-1)!$

**Step 2 :** Now we have 5 places in between these men where we can fit available 5 women. This can be done in 5! Ways.

Total number of ways = 4!*5! = 2880 ways.

**16.** ***Out of 6 teachers and 8 students, a committee of 11 is to be formed. In how many ways can this be done, if the committee contains at least 4 teachers?*** **(2018)**

No. of teachers = 6
No. of Students = 8

A committee of 11 is to be formed. Since, committee contains at least 4 teachers, so committee can consist of

**(1)** 4 teachers and 7 students

(2) 5 teachers and 6 students

(3) 6 teachers and 5 students

We have to calculate all these combinations separately and then add it.

**Case 1 : 4 Teachers and 7 Students :**

| | Total Number | Number to be Chosen | Number of Ways to Choose |
|---|---|---|---|
| Teachers | 6 | 4 | $^6C_4$ |
| Students | 8 | 7 | $^8C_7$ |

$$\text{Total Ways} = {}^6C_4 \times {}^8C_7$$

$$= \frac{6!}{4!2!} \times \frac{8!}{7!11!}$$

$$= \frac{6.5.4!}{4!2!} \times \frac{8.7!}{7!1!}$$

$$= 15 \times 8 = 120.$$

**Case 2 : 5 Teachers and 6 Students :**

| | Total Number | Number to be Chosen | Number of Ways to Choose |
|---|---|---|---|
| Teachers | 6 | 5 | $^6C_5$ |
| Students | 8 | 6 | $^8C_6$ |

$$\text{Total Ways} = {}^6C_5 \times {}^8C_6$$

$$= \frac{6!}{5!1!} \times \frac{8!}{6!2!}$$

$$= 6 \times \frac{8.7.6!}{6!2}$$

$$= 6 \times 28 = 168$$

**Case 3 : 6 Teachers and 5 Students :**

| | Total Number | Number to be Chosen | Number of Ways to Choose |
|---|---|---|---|
| Teachers | 6 | 6 | $^6C_6$ |
| Students | 8 | 5 | $^8C_5$ |

$$\text{Total Ways} = {}^6C_6 \times {}^8C_5$$

$$= 1 \times \frac{8!}{5!3!}$$

$$= \frac{8.7.6.5!}{5!3.2.1}$$

$$= 56$$

Hence, Total ways = 120 + 168 + 56

= 344 **(Ans.)**

> **17.** ***Two dice are thrown. Find the probability that at least a sum of 9 occurs.*** **(2018)**

Total number of outcomes possible when a dice is rolled = 6 ($\because$ any one face out of the 6 faces)

Hence, total number of outcomes possible when two dice are rolled,

$$n(S) = 6 \times 6 = 36$$

Let $E$ be the event = Getting a sum of at least 9 when the two dice fall = $\{(3,6),(4,5),(5,4),(6,3),(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}$.

Hence, $n(E) = 10$

$$P(E) = \frac{n(E)}{n(S)} = \frac{10}{36} = \frac{5}{18}$$ **(Ans.)**

> **18.** ***Probability that a boy will pass an examination is $\frac{3}{5}$ and that for a girl is $\frac{2}{5}$. What is the probability that at least one of them passes?*** **(2018)**

Probability of passing examination by boy = $\frac{3}{5}$,

Probability of not passing examination by boy = $1 - \frac{3}{5} = \frac{2}{5}$

Probability of passing examination by girl = $\frac{2}{5}$

Probability of not passing examination by girl = $1 - \frac{2}{5} = \frac{3}{5}$

Probability of examination not passing by any one

$$= \frac{2}{5} \times \frac{3}{5} = \frac{6}{25}$$

Probability of at least one passes examination

$$1-\frac{6}{25}=\frac{25-6}{25}$$

$$=\frac{19}{25}$$ **(Ans.)**

| | |
|---|---|
| 19. ◆ | *A and B are disjoint events : $P(A) = .5$, $P(A \cup B) = .6$, then find $P(B) = ?$* **(2017)** |
| ◆ | *If $P[A]=0.5$ and $P[A \cup B]=0.6$. Find $P[B]$ if A and B are mutually exclusive.* **(2016)** |

$P[A \cup B] = P[A] + P[B] - P[A \cap B]$ ... (Addition Theorem)

In case, $A$ and $B$ are mutually exclusive $P[A \cap B] = 0$

$\therefore$ $P[A \cup B] = P[A] + P[B]$

$P[B] = P[A \cup B] - P[A] = 0.6 - 0.5 = 0.1$

$P[B] = 0.1$ **(Ans.)**

**20.** ***In how many ways a football eleven can be chosen out of 17 players when (i) five particular players are to be always included, (ii) two particular players are to be always excluded.*** **(2017)**

**(1)** 11 players can be selected out of 17 in ${}^{17}C_{11}$ ways $= {}^{17}C_{11} = (17\times16\times15\times14\times13\times12) / (1\times2\times3\times4\times5\times6)$ = 12376 ways.

**(2)** Since five particular players must always be included, we have to select 6 more out of remaining 12 players. This can be done in ${}^{12}C_6 = (12\times11\times10\times9\times8\times7)/(1\times2\times3\times4\times5\times6) = 924$ ways.

**(3)** Since two particular players must be always excluded, we have to choose 11 players out of remaining 15. This can be done in ${}^{15}C_{11}$ ways $= {}^{15}C_4 = (15\times14\times13\times12)/(1\times2\times3\times4) = 1365$ ways.

**21.** ***A and B take turns in throwing two dice, the first to throw 10 being awarded the prize. Show that if A has the first throw, their chances of winning are in the ratio 12:11.*** **(2017)**

Let $E$ be the event that person $A$ gets a sum of 10

Let $F$ be the event that person $B$ gets a sum of 10

Sum of 10 is obtained in the forms :

(4, 6), (5, 5), (6, 4)

Therefore, $P(E) = 3/36 = 1/12$

And $P(\overline{E})$

$$= 1 - P(E) = 1 - \frac{1}{12} = \frac{11}{12}$$

Similarly, $P(F) = \frac{1}{12}$

$$P(\overline{F}) = 1 - P(\overline{F}) = 1 - \frac{1}{12} = \frac{11}{12}$$

Now A will win in the 1st or 3rd or 5th or......throws

So, probability of winning of A

$$P(A) = P\left[E \cup \left(\overline{E} \cap \overline{F} \cap E\right) \cup \left(\overline{E} \cap \overline{F} \cap \overline{E} \cap \overline{F} \cap E\right) \cup ........\right]$$

$$P(A) = P(E) + P\left(\overline{E} \cap \overline{F} \cap E\right) + P\left(\overline{E} \cap \overline{F} \cap \overline{E} \cap \overline{F} \cap E\right) + ....$$

$$= \frac{1}{12} + \left[\left(\frac{11}{12}\right)^2 \times \frac{1}{12}\right] + \left[\left(\frac{11}{12}\right)^4 \times \frac{1}{2}\right] + ...$$

$$= \frac{1}{12} \times \left[\left(\frac{11}{12}\right)^2 + \left(\frac{11}{12}\right)^4 + ....\right]$$

$$= \frac{1}{12} \times \left[\frac{1}{1 - \left(\frac{11}{12}\right)^2}\right]$$

$$= \frac{1}{12} \times \left[\frac{144}{144 - 121}\right] = \frac{1}{12} \times \frac{144}{23} = \frac{12}{23}$$

$$P(A) = \frac{12}{23}$$

Probability of winning of $B = 1 -$ (prob of winning of $A$)

$$P(B) = 1 - P(A) = 1 - \frac{12}{23} = \frac{11}{23}$$

$$P(B) = \frac{11}{23}$$

$$P(A) : P(B) = \frac{12}{23} : \frac{11}{23} = 12 : 11$$

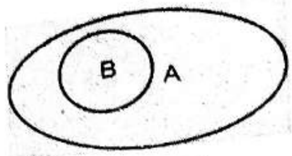
Therefore, if $A$ is the first to throw, the chances of winning of players $A$ and $B$ are in ratio of 12:11

**(Hence Proved)**

 (2016)

22. *If $B \subset A$ then*
(1) $P[A \cap \overline{B}] = P[A] - P[B]$
(2) $P[B] \le P[A]$

To prove the below given relations we make use of Venn Diagram for :
$B \subset A$ :




**(Figure)**

From the Venn Diagram we can say :
$A = (A \cap B) \cup (A \cap B') = B \cup (A \cap B')$

**(1)** $P[A \cap \overline{B}] = P[A] - P[B]$ :

Since,
$A = (A \cap B) \cup (A \cap B') = B \cup (A \cap B')$
From Venn Diagram

$P(A) = P(B \cup (A \cap B'))$
$P(A) = P(B) + P(A \cap B') - P(B \cap A \cap B'))$
$P(A) = P(B) + P(A \cap B') - P(\phi)$
$\therefore \quad P(A \cap B') = P(A) - P(B)$ **(Hence Proved)**

**(2)** $P[B] \le P[A]$ :

Since,
$A = (A \cap B) \cup (A \cap B') = B \cup (A \cap B')$
From Venn Diagram

$P(A) = P(B \cup (A \cap B'))$
$P(A) = P(B) + P(A \cap B') - P(B \cap A \cap B'))$
$P(A) = P(B) + P(A \cap B') - P(\phi)$

Since, $P(A \cap B') \ge 0$, we have $P(A) \ge P(B)$
$P(A) \ge P(B)$
$\therefore \quad P(B) \le P(A)$ **(Hence Proved)**

23. ***A husband and wife appear in an interview for two vacancies in the same post. The probability of the husband's selection is 1/7 and that of wife's selection is 1/5. What is the probability that only one of them will be selected?*** (2016)

Let $A$ = Event that the husband is selected
And $B$ = Event that the wife is selected

Then, $P(A) = \frac{1}{7}$ and $P(B) = \frac{1}{5}$.

$$P(\overline{A}) = \left(1-\frac{1}{7}\right) = \frac{6}{7} \text{ and } P(\overline{B}) = \left(1-\frac{1}{5}\right) = \frac{4}{5}$$

Required probability $= P$ [($A$ and not $B$) or ($B$ and not ($A$)] $= P[(A \text{ and } \overline{B}) \text{ or } (B \text{ and } \overline{A})]$

$$= P(A \text{ and } \overline{B}) + P(B \text{ and } \overline{A})$$

$$= P(A)\cdot P(\overline{B}) + P(B)\cdot P(\overline{A}) = \left(\frac{1}{7}\times\frac{4}{5}\right) + \left(\frac{1}{5}\times\frac{6}{7}\right)$$

$$= \frac{10}{35} = \frac{2}{7}$$ **(Ans.)**

**24.** ***The probability that a person can hit a target is 3/5 and the probability that another person can hit the same target is 2/5. But the first person can fire 8 shots in the time. The second person fires 10 shots. They fire together. What is the probability that the second person shoots the target?*** **(2016)**

We are given :

$$P(E/E_1) = 3/5$$
$$P(E/E_2) = 2/5$$

The ratio of the shots of the first person to those of the second in the same time :

$$= 8/10 = 4/5$$

Then $\frac{P(E_1)}{P(E_2)} = 4/5$

$$P(E_1) = \frac{4}{5}P(E_2)$$

By Bay's Theorem :

$$P(E_2/E) = \frac{P(E_2)P(E/E_2)}{P(E_1)P(E/E_1) + P(E_2)P(E/E_2)}$$

$$= \frac{P(E_2)2/5}{\frac{4}{5}P(E_2)\frac{3}{5} + P(E_2)\frac{2}{5}} = \frac{2/5}{\frac{12}{25}+\frac{2}{5}} = 5/11$$

Therefore, the probability that the second person shoots the target = 5/11.

**25. *A drilling machine bores holds with a mean diameter 0.5230 cm and a standard deviation of 0.0032 cm. Calculate the 2-sigma and 3-sigma upper and lower control limits for means of sample of 4.*** **(2016)**

Mean Diameter = $\bar{\bar{x}} = 0.5230$

Standard Deviation = $\sigma = .0032$

Sample Size $n = 4$

**(1) 2σ Limits :**

Control Limit : $CL = \bar{\bar{x}} = 0.5230 \ cm$

Upper Control Limit :

$$UCL = \bar{\bar{x}} + 2\frac{\sigma}{\sqrt{n}} = 0.5230 + \frac{2 \times 0.0032}{\sqrt{4}} = 0.5262$$

*cm*

Lower Control Limit :

$$LCL = \bar{\bar{x}} - 2\frac{\sigma}{\sqrt{n}} = 0.5230 - \frac{2 \times 0.0032}{\sqrt{4}} = 0.5198$$

*cm*

**(2) 3σ Limits :**

Control Limit : $CL = \bar{\bar{x}} = 0.5230 \ cm$

Upper Control Limit :

$$UCL = \bar{\bar{x}} + 3\frac{\sigma}{\sqrt{n}} = 0.5230 + \frac{3 \times 0.0032}{\sqrt{4}} = 0.5278 \ cm$$

Lower Control Limit :

$$LCL = \bar{\bar{x}} - 3\frac{\sigma}{\sqrt{n}} = 0.5230 - \frac{3 \times 0.0032}{\sqrt{4}} = 0.5182 \ cm$$

**26. *From a pack of 52 cards, two cards are drawn at random. Find the chance that one is a king and the other is a queen.*** **(2015)**

Let $A$ be the event of drawing a king.

$c(A) = 4$

(Total number of king is 4 out of 52 cards)

Let $B$ be the event of drawing a Queen.

$c(B) = 4$

(Total number of queen is 4 out of 52 cards)

Number of favorable outcomes i.e. 'a king or a queen' is $4 + 4 = 8$ out of 52 cards.

$c(A \text{ or } B) = 8$

Therefore, probability of getting 'a king or a queen'

$$= \frac{\text{Number of Favourable Outcomes}}{\text{Number of Total Possible Outcomes}} = \frac{8}{52} = \frac{2}{13} \quad \textbf{(Ans.)}$$

**27. *A bag contains 8 balls of which 5 are red and 3 are black. Two balls are drawn at random. What is the probability that both are black?* (2015)**

Let $A$ be the event that both the balls drawn are Red.

Two balls can be drawn from 8 balls in ${}^8C_2$ equally likely ways.

$$\text{Therefore, } n(S) = {}^8C_2 = \frac{8!}{2!6!} = 28$$

Two balls can be drawn from 5 red balls in ${}^5C_2$ ways

$$n(A) = {}^5C_2 = \frac{5!}{2!3!} = 10 \text{ ways}$$

Thus, probability of drawing 2 red balls from 8 balls is :

$$p(A) = \frac{n(A)}{n(S)} = \frac{10}{28} = 0.36 \quad \textbf{(Ans.)}$$

**28. *A book contains 100 pages numbering from 1 to 100. A page is opened at random and is selected. Find the probability that opened page is a multiple of 6 or 10.* (2015)**

Since any of the 100 pages can be opened and selected, therefore exhaustive number of cases are

$c(S) = 100$

(1) Let $A$ be the event of getting page which is multiple of 6 :

$A = \{6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 90, 96\}$

$c(A) = 16$

$$P(A) = \frac{c(A)}{c(S)} = \frac{16}{100}$$

(2) Let $B$ be the even of getting a page which is multiple of 10 :

$$B = \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$$

$$c(B) = 10$$

$$P(B) = \frac{c(B)}{c(S)} = \frac{10}{100}$$

Since both the events are equally likely but not mutually exclusive because the pages 30 and 60 are multiple of 6 and 10 both therefore :

$$(A \cap B) = \{30, 60\}$$

$$c(A \cap B) = 2$$

$$P(A \cap B) = \frac{2}{100}$$

To calculate the probability of getting a page which is multiple of 6 or 10, we make use of Addition Theorem of probability.

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

$$P(A \cup B) = \frac{16}{100} + \frac{10}{100} - \frac{2}{100}$$

$$P(A \cup B) = \frac{24}{100} = 0.24$$

**29. *A committee of 5 is to be formed out of 6 gents and ladies. In how many ways this can be done when***
**(1) *At least 2 ladies are included***
**(2) *At most 2 ladies are included*** **(2015)**

(1) **When At least 2 Ladies are Included :** There are three possibilities :

(a) **3 Ladies, 2 Gents :** It can be done in $C(6, 2) \cdot C(4, 3)$ ways

(b) **4 Ladies, 1 Man :** It can be done in $C(6, 1)$ $C(4, 4)$ ways

(c) **2 Ladies, 3 Gents :** It can be done in $C(6, 3) \cdot C(4, 2)$ ways.

The required number of ways :

$$= 4C_2 \times 6C_3 + 4C_3 \times 6C_2 + 4C_4 \times 6C_1$$

$$= \frac{4!}{2!2!} \times \frac{6!}{3!3!} + \frac{4!}{3!1!} \times \frac{6!}{2!4!} + \frac{4!}{4!0!} \times \frac{6!}{5!1!}$$

$$= 6 \times 20 + 4 \times 15 + 1 \times 6$$

$$= 120 + 60 + 6 = 186 \text{ ways}$$

**(Ans.)**

(2) **When At Most 2 Ladies are Included :** There are three possibilities :

(a) **5 Gents and No Lady :** This can be done in $C(6, 5)\ C(4, 0)$ ways

(b) **4 Gents and 1 Lady :** This can be done in $C(6, 4)\ C(4, 1)$ ways

(c) **3 Gents and 2 Ladies :** This can be done in $C(6, 3)\ C(4, 2)$ ways

The required number of ways :

$C(6, 5).C(4, 0) + C(6, 4).C(4, 1) + C(6, 3).C(4, 2)$

$$= \frac{4!}{4!0!} \times \frac{6!}{5!1!} + \frac{4!}{3!1!} \times \frac{6!}{2!4!} + \frac{4!}{2!2!} \times \frac{6!}{3!3!}$$

$= 6 + 60 + 120 = 186$ Ways

Therefore,

When at least two ladies are included : 186 ways **(Ans.)**

When at most two ladies are included: 186 ways **(Ans.)**

**30.** ***There are 100 cards. These cards are numbered from 1 to 100. One card is drawn at random. What is the probability that the number on the card is a square.*** **(2015)**

In the experiment of drawing/picking a card

Total Number of Possible Choices = Number of ways in which one card can be drawn from the total $r$ :

$$\Rightarrow n = {}^{r}c_1$$

Let "A" be the event of the card drawn having a number marked on it which is a square

**For Event "A" :** Numbers which are squares from 1 to 100 = 10 {1, 4, 9, 16, 25, 36, 49, 64, 81, 100}

| | Favorable | Unfavorable | Total |
|---|---|---|---|
| Available | 10 | 90 | 100 |
| To Choose | 1 | 0 | 1 |
| Choices | ${}^{10}C_1$ | ${}^{90}C_0$ | ${}^{100}C_1$ |

Number of Favorable Choices = the number of ways in which one card with a number which is a square marked on it, from the total 10 cards

$$\Rightarrow m_A = {}^{10}c_1 = \frac{10}{1} = 10$$ **(Ans.)**

**31.** ***Three horses A, B and C are in race. A is twice as likely to win as B is twice as likely to win as C. What are their respective probabilities of winning?*** **(2014)**

Let $A$ be the event that Horse $A$ will win
Let $B$ be the event that Horse $B$ will win
Let $C$ be the event that Horse $C$ will win

We know that $\Pr(A) = 2\Pr(B)$, and $\Pr(B) = 2\Pr(C)$; hence, $\Pr(A) = 4\Pr(C)$.

But $\Pr(A) + \Pr(B) + \Pr(C) = 1$.

Consequently, $4\Pr(C) + 2\Pr(C) + \Pr(C) = 1$ i.e. $7\Pr(C) = 1$

$\Pr(C) = 1/7$

$\Pr(A) = 4\Pr(C)$, hence $P(A) = 4/7$ and

$\Pr(B) = 2\Pr(C)$, hence $P(B) = 2/7$

Thus their respective probabilities of winning is :

Horse $A = 4/7$ **(Ans.)**

Horse $B = 2/7$ **(Ans.)**

Horse $C = 1/7$ **(Ans.)**

❑❑

# STATISTICAL QUALITY CONTROL

**1.** ***In the production of certain rods a process is said to be in control if the outside diameters have a mean 2.532 and a standard deviation of 0.002. Find the central limits for the mean of random samples of size 4.*** **(2023-24)**

**Given :**

**(1)** Population mean $(\mu) = 2.532$

**(2)** Population standard deviation $(\sigma) = 0.002$

**(3)** Sample size $(n) = 4$

We can calculate the standard deviation of the sample means $(\sigma \overline{X})$ using the formula:

$$\sigma \overline{X} = \frac{\sigma}{\sqrt{n}}$$

Substituting the given values :

$$\sigma \overline{X} = \frac{.002}{\sqrt{4}}$$

$$\sigma \overline{X} = \frac{.002}{2}$$

Therefore, the standard deviation of the sample means $(\sigma \overline{X})$ is 0.001.

Now, to find the central limits for the mean of random samples of size 4, we use the formula:

$$\text{Central Limits} = \mu \pm z \times \sigma \overline{X}$$

Where :

**(1)** $\mu$ is the population mean

**(2)** $\sigma \overline{X}$ is the standard deviation of the sample means

**(3)** $z$ is the z-score corresponding to the desired level of confidence (typically 1.96 for a 95% confidence level)

Let's calculate the central limits.

**Substituting the Values :**

Central Limits $= 2.532 \pm 1.96 \times 0.001$

Now, let's calculate the central limits.

The central limits for the mean of random samples of size 4, assuming a 95% confidence level, are:

**Upper Limit :**

Upper Limit= $2.532 + 1.96 \times 0.001$

Upper Limit = $2.532 + 0.00196$

Upper Limit $\approx 2.53396$

**Lower Limit :**

Lower Limit = $2.532 - 1.96 \times 0.001$

Lower Limit = $2.532 - 0.00196$

Lower Limit $\approx 2.53004$

Therefore, the central limits for the mean of random samples of size 4, assuming a 95% confidence level, are approximately 2.53004 and 2.53396.

**2. *Calculate interquartile range. Quartile Deviation and coefficient of quartile deviation from the following data :***

| Marks | 4-8 | 9-13 | 14-18 | 19-23 | 24-28 |
|---|---|---|---|---|---|
| No. of Students | 3 | 4 | 3 | 2 | 4 |

**(2023-24)**

The interquartile range denotes the difference between the third quartile and the first quartile. Symbolically, interquartile range $= Q_3 - Q_1$

$$\text{Quartile deviation} = \frac{Q_3 - Q_1}{2}$$

$$\text{Coefficient of } QD = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

| Class Interval | Class Interval | Frequency | Cumulative Frequency |
|---|---|---|---|
| 4-8 | 3.5-8.5 | 3 | 3 |
| 9-13 | 8.5-13.5 | 4 | 7 |
| 14-18 | 13.5-18.5 | 3 | 10 |
| 19-23 | 18.5-23.5 | 2 | 12 |
| 24-28 | 23.5-28.5 | 4 | 16 |

In order to calculate quartiles, $Q_1$ and $Q_3$, it is necessary that we should use the cumulative frequencies.

**Finding $Q_1$ :**

$$\textbf{Formula :}\ Q_1 = l + \left( \frac{\left( \frac{N}{4} - c.f \right)}{f} \right) h$$

$\frac{N}{4} = \frac{16}{4} = 4$, so class will be 8.5-13.5

$l$ (lower limit of class) = 8.5

N/4 = 16

c.f (cumulative frequency of preceding class) = 3

$f$ (frequency of class) = 4

$h$ (class width) = 5

$$Q_1 = l + \left( \frac{\left( \frac{N}{4} - c.f \right)}{f} \right) h$$

$$= 8.5 + \left( \frac{\left( \frac{16}{4} - 3 \right)}{4} \right) 5$$

$$= 8.5 + \left( \frac{1}{4} \right) 5 = 8.5 + 1.25 = 9.25$$

**Finding $Q_3$ :**

**Formula :** $Q_3 = l + \left( \frac{\left( \frac{3N}{4} - c.f \right)}{f} \right) h$

$\frac{3N}{4} = 3 * \frac{16}{4} = 12$

So class will be 18.5-23.5

$l$ (lower limit of class) = 18.5

$\frac{3N}{4} = 12$

c.f (cumulative frequency of preceding class) = 10

$f$ (frequency of class) = 2

$h$ (class width) = 5

$$Q_3 = l + \left( \frac{\left( \frac{3N}{4} - c.f \right)}{f} \right) h$$

$$= 18.5 + \left(\frac{(12-10)}{2}\right)5 = 23.5$$

$$IQR = Q_3 - Q_1 = 23.5 - 9.25 = 13.75$$

Quartile Deviation

$$= \frac{Q_3 - Q_1}{2} = \frac{13.75}{2} = 6.875$$

Coefficient of QD

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{13.75}{33.25} = 0.4135$$

**3. *Two dice are thrown. Find the probability that at least a sum of 10 occurs.* (2023-24)**

To find the probability that at least a sum of 10 occurs when two dice are thrown, we need to determine the number of favorable outcomes and divide it by the total number of possible outcomes.

**Total Number of Possible Outcomes :** When two dice are thrown, each die has 6 sides, so there are $6 \times 6 = 36$ possible outcomes.

**Favorable Outcomes for a Sum of at Least 10 :**

To find the favorable outcomes, we list all combinations of two dice rolls that result in a sum of at least 10 :

(4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6, 6)

There are 6 favorable outcomes.

**Calculation of Probability :**

Probability (P) = Number of favorable outcomes/Total number of possible outcomes

Total number of possible outcomes = 36

Number of favorable outcomes = 6

$$P = \frac{6}{36} = \frac{1}{6} = 0.1667$$

**(Ans.)**

The probability that at least a sum of 10 occurs when two dice are thrown is 1/6 or approximately 0.1667.

**4. *A box contains 5 different red and 6 different white balls. In how many ways can 6 balls be selected so that there are at least two balls of each colour.***

**(2023-24)**

The no. of red color balls in the box = 5
The no. of white color balls in the box = 6

We have to select 6 balls from the box so that there should be 2 balls at least of each color, which means that there can be more than two balls from each color.

Here selecting the no. of combinations in total 6 balls choosing balls from red and white colors, which have at least two from each color.

Considering the cases of selection as given below :

Selecting 4 red balls from 5 red balls and 2 white balls from 6 white balls which can be done in the following ways, as given below :

$$^5C_4 \times {}^6C_2 = 75$$

Now selecting 3 red balls from 5 red balls and 3 white balls from 6 white balls which can be done in the following ways, as given below :

$$^5C_3 \times {}^6C_3 = 200$$

Now selecting 2 red balls from 5 red balls and 4 white balls from 6 white balls which can be done in the following ways, as given below :

$$^5C_2 \times {}^6C_4 = 150$$

So the total no. of ways that 6 balls be selected so that there are at least two balls of each color, is given by :

75 + 200 + 150 = 425

**Final Answer :** The no. of ways that 6 balls be selected so that there are at least two balls of each color is 425.

**5. *Mobile charger supplier drawn randomly constant sample size of 500 chargers every day for quality control test. Defects in each charger are recorded during testing. 12, 14, 16, 18, 16, 14, 12, 12, 32, 16, 18, 16, 12, 14, 16, 18, 12, 19, 18, and 21.***

***Draw control chart for the number of defects (C-chart) and give your comments.*** **(2023-24)**

c-chart is also known as the control chart for defects (counting of the number of defects). It is generally used to monitor the number of defects in constant size units.

c chart, the number of defects is plotting on the y-axis and the number of units on the x-axis. The centerline of the c chart $\bar{c}$ is the total number of defects divided by the number of samples.

$$\bar{c} = \frac{\textit{Total number of defects}}{\textit{Number of samples}}$$

$$\bar{c} = \frac{\sum c}{k}$$

$$UCL_c = \bar{c} + 3\sqrt{\bar{c}}$$

$$LCL_c = \bar{c} - 3\sqrt{\bar{c}}$$

Where $c$ = number of defects

$k$ = number of samples

| Lot | Sample size | No.of Defects | $\bar{c}$ | UCLx | LCLx |
|---|---|---|---|---|---|
| 1. | 500 | 12 | 16.3 | 28.4 | 4.19 |
| 2. | 500 | 14 | 16.3 | 28.4 | 4.19 |
| 3. | 500 | 16 | 16.3 | 28.4 | 4.19 |
| 4. | 500 | 18 | 16.3 | 28.4 | 4.19 |
| 5. | 500 | 16 | 16.3 | 28.4 | 4.19 |
| 6. | 500 | 14 | 16.3 | 28.4 | 4.19 |
| 7. | 500 | 12 | 16.3 | 28.4 | 4.19 |
| 8. | 500 | 12 | 16.3 | 28.4 | 4.19 |
| 9. | 500 | 32 | 16.3 | 28.4 | 4.19 |
| 10. | 500 | 16 | 16.3 | 28.4 | 4.19 |
| 11. | 500 | 18 | 16.3 | 28.4 | 4.19 |
| 12. | 500 | 16 | 16.3 | 28.4 | 4.19 |
| 13. | 500 | 12 | 16.3 | 28.4 | 4.19 |
| 14. | 500 | 14 | 16.3 | 28.4 | 4.19 |
| 15. | 500 | 16 | 16.3 | 28.4 | 4.19 |
| 16. | 500 | 18 | 16.3 | 28.4 | 4.19 |
| 17. | 500 | 12 | 16.3 | 28.4 | 4.19 |
| 18. | 500 | 19 | 16.3 | 28.4 | 4.19 |
| 19. | 500 | 18 | 16.3 | 28.4 | 4.19 |
| 20. | 500 | 21 | 16.3 | 28.4 | 4.19 |

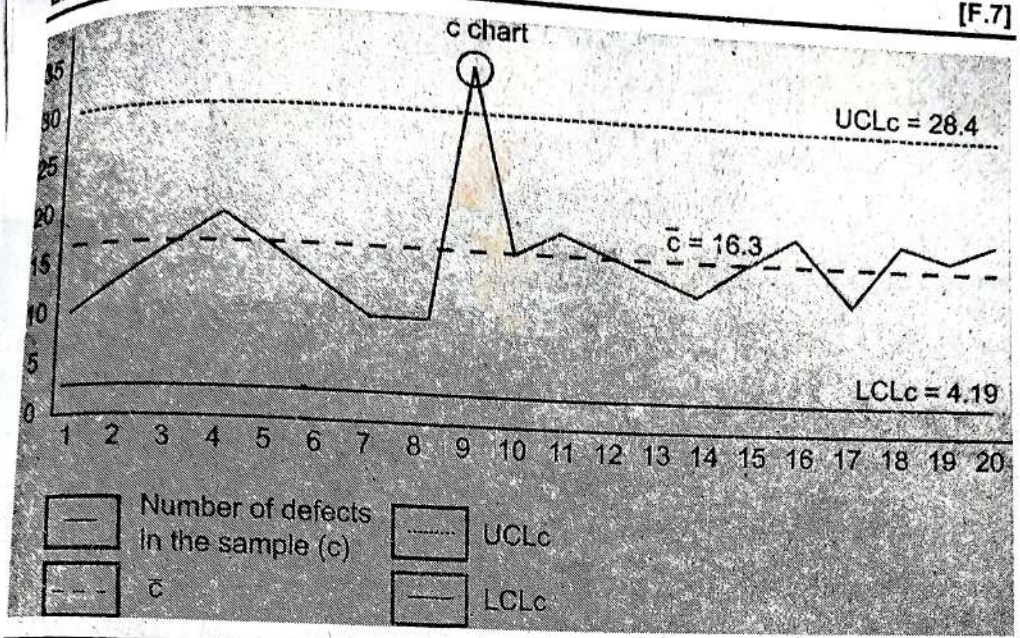
$\bar{c}$ = total number of defects / total number of lots

$$= \frac{\sum c}{k} = \frac{326}{20} = 16.3$$

Then calculate upper control limit (UCL) and low control limit (LCL)

$$UCL_c = \bar{c} + 3\sqrt{\bar{c}} = 16.3 + 3\sqrt{16.3} = 28.4$$

$$LCL_c = \bar{c} - 3\sqrt{\bar{c}} = 16.3 - 3\sqrt{16.3} = 4.19$$




**6.** ***The data shows the sample mean and range for 10 samples for size 5 each. Draw mean chart, range chart and comment on the state of control of the process.***

| Sample No. | Mean $(\overline{X})$ | Range (R) |
|---|---|---|
| 1 | 21 | 5 |
| 2 | 26 | 6 |
| 3 | 23 | 9 |
| 4 | 18 | 7 |
| 5 | 19 | 4 |
| 6 | 15 | 7 |
| 7 | 14 | 4 |
| 8 | 20 | 9 |
| 9 | 16 | 8 |
| 10 | 16 | 6 |

**Assume for $n = 5$, $A_2 = 0.58$, $D_3 = 0$ and $D_4 = 2.115$.**

**(2023- 24)**

| Sample No. | Mean $(\overline{X})$ | Range (R) |
|---|---|---|
| 1 | 21 | 5 |
| 2 | 26 | 6 |
| 3 | 23 | 9 |
| 4 | 18 | 7 |
| 5 | 19 | 4 |
| 6 | 15 | 7 |
| 7 | 14 | 4 |
| 8 | 20 | 9 |
| 9 | 16 | 8 |

| | | |
|---|---|---|
| 10 | 16 | 6 |
| Total | 188 | 65 |

$$\overline{\overline{X}} = \frac{\sum \overline{X}}{number\ of\ samples} = \frac{188}{10} = 18.8$$

$$\overline{R} = \frac{\sum R}{n} = \frac{65}{10} = 6.5$$

The control limit for Mean Chart is

$$CL = \overline{\overline{X}} = 18.8$$

$$UCL = \overline{\overline{X}} + A_2\overline{R} = 18.8 + 0.58 * 6.5 = 22.57$$

$$LCL = \overline{\overline{X}} - A_2\overline{R} = 18.8 - 0.58 * 6.5 = 15.03$$

The control limit for Range Chart is

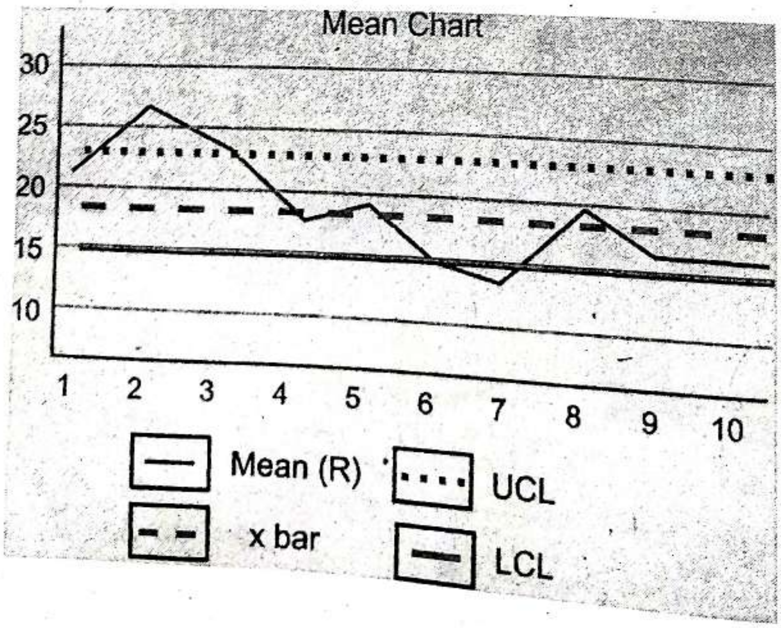
$$UCL = D_4\overline{R} = 2.115 * 6.5 = 13.7475$$

$$CL = \overline{R} = 6.5$$

$$LCL = D_3\overline{R} = 0 * 6.5 = 0$$

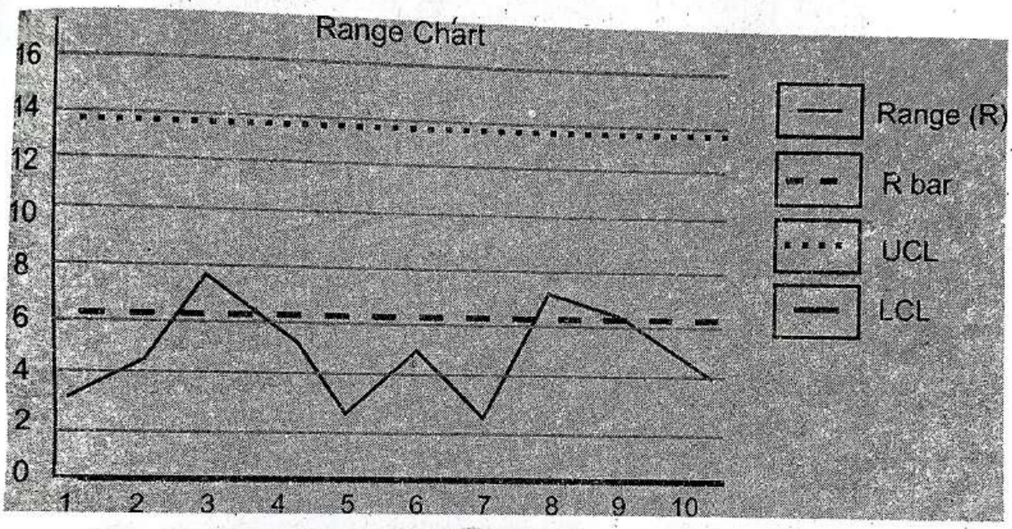
| Sample No. | Mean(R) | x bar | UCL | LCL |
|---|---|---|---|---|
| 1 | 21 | 18.8 | 22.57 | 15.03 |
| 2 | 26 | 18.8 | 22.57 | 15.03 |
| 3 | 23 | 18.8 | 22.57 | 15.03 |
| 4 | 18 | 18.8 | 22.57 | 15.03 |
| 5 | 19 | 18.8 | 22.57 | 15.03 |
| 6 | 15 | 18.8 | 22.57 | 15.03 |
| 7 | 14 | 18.8 | 22.57 | 15.03 |
| 8 | 20 | 18.8 | 22.57 | 15.03 |
| 9 | 16 | 18.8 | 22.57 | 15.03 |
| 10 | 16 | 18.8 | 22.57 | 15.03 |

**Mean Chart :**



| Sample No. | Range (R) | R bar | UCL | LCL |
|---|---|---|---|---|
| 1 | 5 | 6.5 | 13.74 | 0 |
| 2 | 6 | 6.5 | 13.74 | 0 |
| 3 | 9 | 6.5 | 13.74 | 0 |
| 4 | 7 | 6.5 | 13.74 | 0 |
| 5 | 4 | 6.5 | 13.74 | 0 |
| 6 | 7 | 6.5 | 13.74 | 0 |
| 7 | 4 | 6.5 | 13.74 | 0 |
| 8 | 9 | 6.5 | 13.74 | 0 |
| 9 | 8 | 6.5 | 13.74 | 0 |
| 10 | 6 | 6.5 | 13.74 | 0 |

**Range Chart**




The Process is under Control.

| | |
|---|---|
| **7.** | ***Write a short note on statistical quality control.*** **(2016, 2021, 2022-23)** |

***Statistical Quality Control***

Statistical quality control refers to the use of statistical methods in the monitoring and maintaining of the quality of products and services. One method, referred to as acceptance sampling, can be used when a decision must be made to accept or reject a group of parts or items based on the quality found in a sample. A second method, referred to as statistical process control, uses graphical displays known as control charts to determine whether a process should be continued or should be adjusted to achieve the desired quality.

**Acceptance Sampling :** Assume that a consumer receives a shipment of parts called a lot from a producer. A sample of parts will be taken and the number of defective items counted. If the number of defective items is low, the entire lot will be accepted. If the number of defective items is high, the entire lot will be rejected. Correct decisions correspond to accepting a good quality lot and rejecting a poor quality lot. Because sampling is being used, the probabilities of erroneous decisions need to be considered. The error of rejecting a good-quality lot creates a problem for the producer the probability of this error is called the producer's risk. On the other hand, the error of accepting a poor quality lot creates a problem for the purchaser or consumer the probability of this error is called the consumer's risk.

The design of an acceptance sampling plan consists of determining a sample size $n$ and an acceptance criterion $c$, where $c$ is the maximum number of defective items that can be found in the sample and the lot still be accepted. The key to understanding both the producer's risk and the consumer's risk is to assume that a lot has some known percentage of defective items and compute the probability of accepting the lot for a given sampling plan. By varying the assumed percentage of defective items in a lot, several different sampling plans can be evaluated and a sampling plan selected such that both the producer's and consumer's risk are reasonably low.

**Statistical Process Control :** Statistical process control uses sampling and statistical methods to monitor the quality of an ongoing process such as a production operation. A graphical display referred to as a control chart provides a basis for deciding whether the variation in the output of a process is due to common causes (randomly occurring variations) or to out-of-the-ordinary assignable causes. Whenever assignable causes are identified, a decision can be made to adjust the process in order to bring the output back to acceptable quality levels.

Control charts can be classified by the type of data they contain :

(1) For instance, an $x$-chart is employed in situations where as sample mean is used to measure the quality of the output.

(2) Quantitative data such as length, weight, and temperature can be monitored with an $x$-chart.

(3) Process variability can be monitored using a range or $R$-chart.

(4) In cases in which the quality of output is measured in terms of the number of defectives or the proportion of defectives in the sample, an $np$-chart or a $p$-chart can be used.

All control charts are constructed in a similar fashion. For example, the centre line of an $x$-chart corresponds to the mean of the process when the process is in control and producing output of acceptable quality. The vertical axis of the control chart identifies the scale of measurement for the variable of interest. The upper horizontal line of the control chart, referred to as the upper control limit, and the lower horizontal line, referred to as the lower control limit, are chosen so that when the process is in control there will be a high probability that the value of a sample mean will fall between the two control limits. Standard practice is to set the control limits at three standard deviations above and below the process mean. The process can be sampled periodically. As each sample is selected, the value of the sample mean is plotted on the control chart. If the value of a sample mean is within the control limits, the process can be continued under the assumption that the quality standards are being maintained. If the value of the sample mean is outside the control limits, an out-of-control conclusion points to the need for corrective action in order to return the process to acceptable quality levels.

**8. *The number of defects in 18 rolls of cloth each of 1-50 meters length is given by 3, 5, 8, 9, 4, 2, 5, 9, 6, 4, 8, 12, 7, 5, 10, 10, 7 and 5.***
***Draw C – chart and give your comment.*** **(2022-23)**

A C control chart is a statistical tool used to monitor the number of defects in a process over time. The C chart plots the count of defects in each sample on the y-axis, while the x-axis represents the sample number. The chart also includes a center line, which is the average count of defects, as well as upper and lower control limits (UCL and LCL) that indicate the expected range of variation in the process.

To draw a C-chart for the given data on the number of defects in 18 rolls of cloth, we need to calculate the average number of defects per roll and determine the control limits.

First, let's calculate the average number of defects per roll :

Sum of defects = 3 + 5 + 8 + 9 + 4 + 2 + 5 + 9 + 6 + 4 + 8 + 12 + 7 + 5 + 10 + 10 + 7 + 5 = 119

Average defects per roll = (Sum of defects) / (Number of rolls) = 119 / 18 =6.61

Next, we can calculate the control limits for the C-chart :

Upper control limit (UCL) = Average defects per roll + 3 * sqrt(Average defects per roll) = 6.61 + 3 * sqrt(6.61)

Lower control limit (LCL) = Average defects per roll - 3 * sqrt(Average defects per roll) =6.61 - 3 * sqrt(6.61)

Now, let's calculate the square root of the average defects per roll :

sqrt(6.61) ≈ 2.57

UCL = 6.61 + 3 * 2.57 ≈ 14.32

LCL = 6.61 - 3 * 2.57 ≈ -1.1

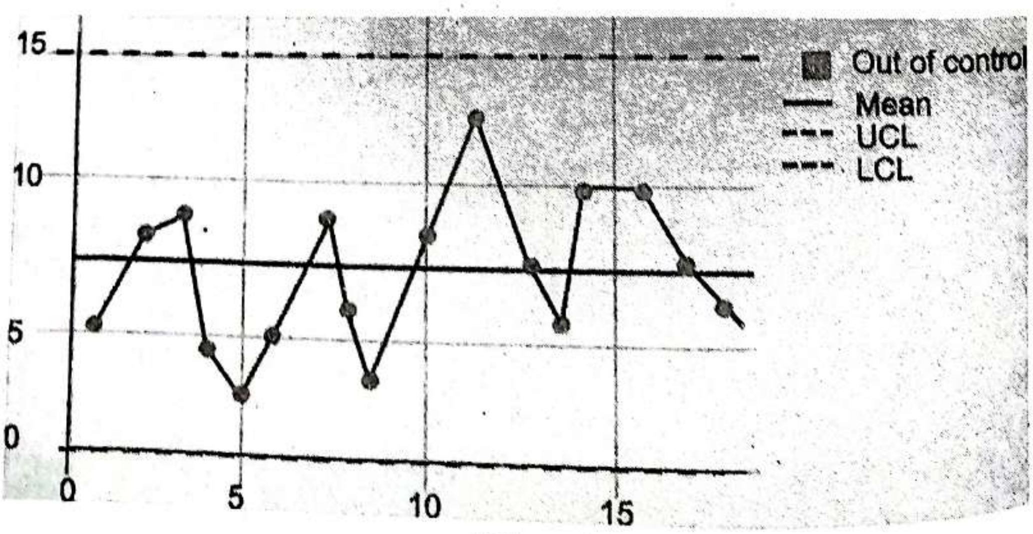
However, the lower control limit cannot be negative in this case, so we set the LCL to be 0.

Now we have the control limits for the C-chart :

UCL ≈ 14.32

LCL = 0

To draw the C-chart, we plot the number of defects in each roll on the y-axis and the roll numbers on the x-axis. We then draw the UCL and LCL as horizontal lines on the chart.




(Figure)

Based on the C-chart, we can make the following observations :

(1) Most of the data points fall within the control limits, indicating that the process is stable.

(2) There are a few data points close to the upper control limit, suggesting some variability in the number of defects.

(3) No data points fall below the lower control limit, which is expected since we set the LCL to be 0.

Overall, the C-chart can be used as a tool for monitoring the number of defects in rolls of cloth. If any data points fall outside the control limits or if there are consistent patterns or trends, it may indicate the need for further investigation or process improvement.

**9. *The following table gives the result of inspection of 15 samples of 100 items each taken on working days. Draw a np-chart :***

| Sample No. | No. of Defectives |
|---|---|
| 1 | 9 |
| 2 | 17 |
| 3 | 8 |
| 4 | 7 |
| 5 | 12 |
| 6 | 5 |
| 7 | 11 |
| 8 | 16 |
| 9 | 14 |
| 10 | 15 |
| 11 | 10 |
| 12 | 6 |
| 13 | 7 |
| 14 | 18 |
| 15 | 10 |

**(2022-23)**

np chart is also known as the control chart for defectives (d-chart). It generally monitors the number of non-conforming or defective items in the measurement process. It uses binomial distribution to measure the number of defective or non-conforming units in a sample. An np chart is very similar to the $p$ chart. np chart plots the number of items, while the $p$ chart plots the proportion of defective items.

These charts are modified form of P-chart and are used when the analyst is more interested in number of defectives than fraction of defectives.

**Steps to Construct Np-Chart is**

**1st Step :** Calculate

$$\overline{P} = \frac{\text{Total number of defectives}}{\text{Total number of items inspected}}$$

**2nd Step :** Determine the Central line :

Central line = $n\overline{P}$

Here $n$ = size of sample

**3rd Step :** Calculate Control Limits :

$$\text{Control Limits} = n\overline{P} \pm 3\sqrt{n\overline{P}\left(1-\overline{P}\right)}$$

$$UCL = n\overline{P} + 3\sqrt{n\overline{P}\left(1-\overline{P}\right)}$$

$$LCL = n\overline{P} - 3\sqrt{n\overline{P}\left(1-\overline{P}\right)}$$

Np-charts are more preferred than P-charts as in np-charts; the number of defective can be directly plotted from the inspection report.

$\overline{P}$ = Total number of defectives items/Total number of items inspected

= 165/15*100

=0.11

Central line $n * \overline{P} = 100 * 0.11 = 11$

UCL = $n * \overline{P} + 3\sqrt{n} * \overline{P}\left(1-\overline{P}\right)$

= 11 + 3√11 (1-.011)

= 11 + 3√11*0.89

= 11 + 3√9.79

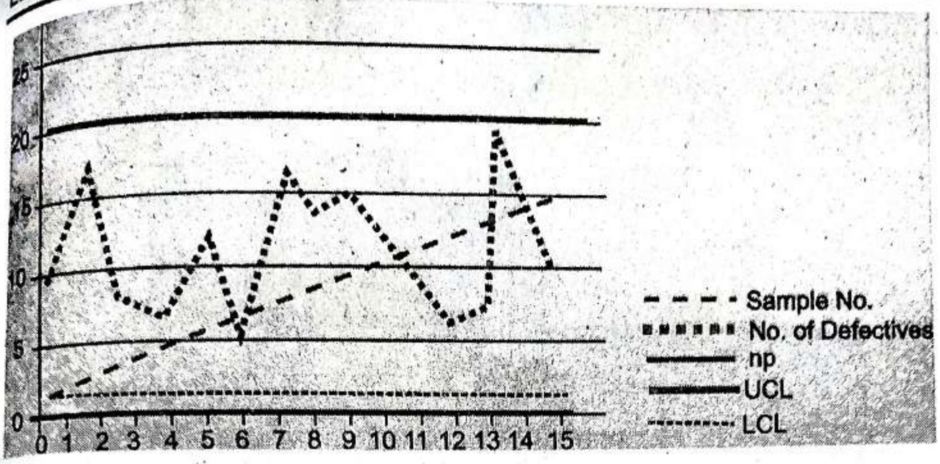
= 11 + 9.38

= 20.39

LCL = $n * \overline{P} - 3\sqrt{n} * \overline{P}\left(1-\overline{P}\right)$

= 11 -9.38 =1.62

Since the number of defectives in each sample are within the two limits i.e. UCL and LCL so process is said to be under control.




(Figure)

**10.** ***The following table gives the result of inspection of 15 samples of 100 items each taken on working days. Draw a n. p. chart.*** **(2021)**

| Sample No. | No. of Defectives |
|---|---|
| 1 | 9 |
| 2 | 17 |
| 3 | 8 |
| 4 | 7 |
| 5 | 12 |
| 6 | 5 |
| 7 | 11 |
| 8 | 16 |
| 9 | 14 |
| 10 | 15 |
| 11 | 10 |
| 12 | 6 |
| 13 | 7 |
| 14 | 8 |
| 15 | 5 |

We have $n = 100$

$\sum np =$ total number of defectives $= 150$

$\sum n =$ total number inspected $= 100*15$

$$\overline{p} = \sum np / \sum n = 150/100*15 = 0.1$$

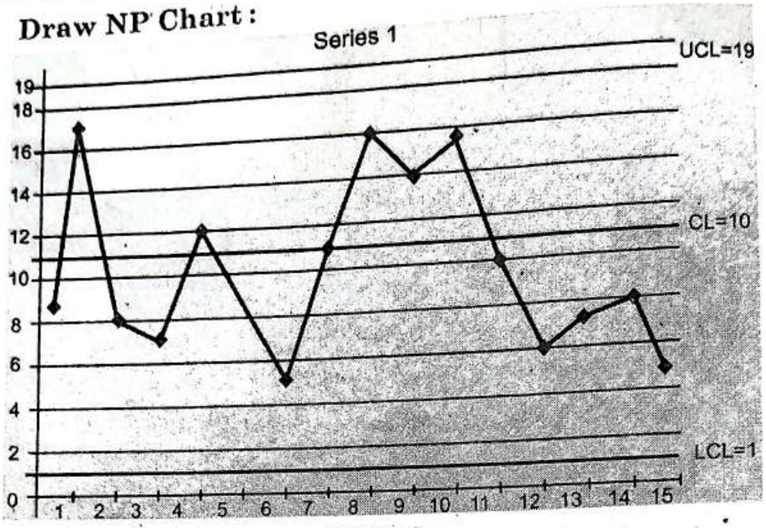
$$n\overline{p} = 100 * 0.1 = 10$$

Control limits $\quad C.L. = n\overline{p} = 10$

$$\text{U.C.L.} = n\overline{p} + 3\sqrt{np}\left(1-\overline{p}\right) = 10 + 3\sqrt{10}\left(1-0.1\right) = 19$$

$$\text{L.C.L.} = n\overline{p} - 3\sqrt{np}\left(1-\overline{p}\right) = 10 - 3\sqrt{10}\left(1-0.1\right) = 1$$

**Draw NP Chart :**




(Figure)

**11.** ***The following data shows the value of samples mean X and range R for 10 samples of size 5 each. Calculate the values for control line and control limits for mean chart and range chart and determine whether process is under control***

| Sample No. | Mean (x) | Range (R) |
|---|---|---|
| 1 | 15 | 7 |
| 2 | 17 | 7 |
| 3 | 15 | 4 |
| 4 | 18 | 9 |
| 5 | 17 | 8 |
| 6 | 14 | 7 |
| 7 | 18 | 12 |
| 8 | 15 | 4 |
| 9 | 17 | 11 |
| 10 | 16 | 5 |

**Give for**

$n = 5, A_2 = 0.577, D_3 = 0, D_4 = 2.115$ **(2021)**

**(1) Control Limit for $\overline{X}$ Chart :**

We have 20 samples and size of each is 4

$K = 10, n = 2, \Sigma\overline{X} = 162, \Sigma R = 74$

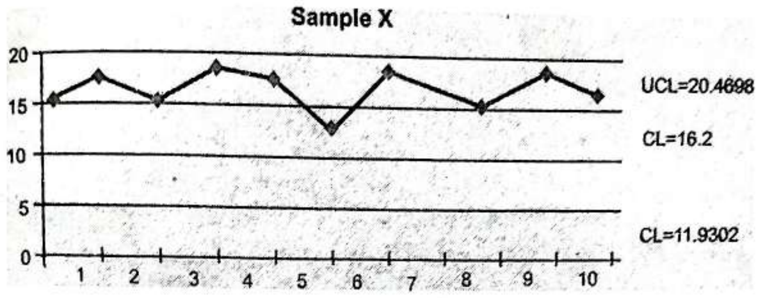
$\overline{\overline{x}} = \Sigma\overline{X}/k = 162/10 = 16.2$

$\overline{R} = \Sigma R/k = 74/10 = 7$

Control line $\overline{\overline{X}} = 16.2$

$$U.C.L = \overline{\overline{X}} + A_2 R = 16.2 + 0.577 * 7.4 = 20.4698$$

$$L.C.L. = \overline{\overline{X}} - A_2 R = 16.2 - 0.577 * 7.4 = 11.9302$$



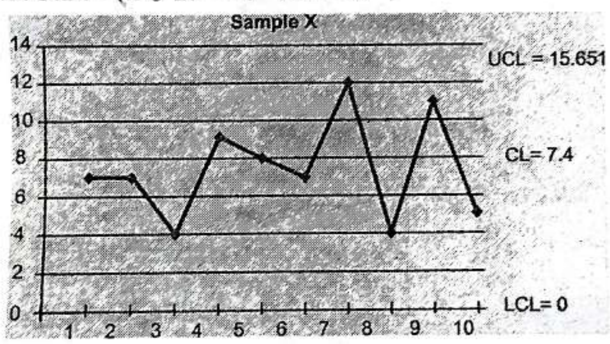

(Figure)

**(2) Control Limit for $R$ Chart :**

Control Limit (C.L.) = $\overline{C}$ = 7.4

U.C.L = $D_4\ R$ = 2.115*7.4 = 15.651

L.C.L. = $D_3\ \overline{R}$ = 0*7.4 = 0




(Figure)

**12. *In a production of certain rods, a process is said to be in control if outside diameters have a mean 2.5 and a S. D. of 0.004. Find the control limits for the mean of random samples of size 4.*** **(2019)**

## Control Limits

Control limits, also known as natural process limits, are horizontal lines drawn on a statistical process control chart, usually at a distance of ±3 standard deviations of the plotted statistic from the statistic's mean.

**Control Limits are Calculated by :**

Given $\mu_p = 2.5$ and $\sigma_p = 0.004; n = 4$

**Control Limits for $\overline{X}$ Chart :**

The Central Line (C.L)= $\mu_p = 2.5''$

Upper Control Limit (UCL) = $\mu_p + \frac{3\sigma_p}{\sqrt{n}} = 2.5 + \frac{3(0.004)}{\sqrt{4}} = 2.5 + 0.006 = 2.506''$

Upper Control Limit (UCL) $= 2.506''$

Lower Control Limit (LCL) = $\mu_p - \frac{3\sigma_p}{\sqrt{n}} = 2.5 - \frac{3(0.004)}{\sqrt{4}} = 2.5 - 0.006 = 2.494''$

Lower Control Limit (LCL) $= 2.494''$

**Therefore, the Control Limits for the mean of the Random Samples of Size 4 are :**

The central line (C.L)= $2.5''$

Upper Control Limit (UCL) = $2.506''$

Lower Control Limit (LCL) = $2.494''$ **(Ans.)**

13. ***The following table gives the result of inspection of 15 samples of 1000 items each taken of working days. Draw a np – chart :*** **(2019)**

| Sample No. | No. of Defectives |
|---|---|
| 1 | 9 |
| 2 | 10 |
| 3 | 12 |
| 4 | 8 |
| 5 | 7 |
| 6 | 15 |
| 7 | 12 |
| 8 | 10 |
| 9 | 8 |
| 10 | 10 |
| 11 | 7 |
| 12 | 13 |
| 13 | 14 |
| 14 | 15 |
| 15 | 16 |

A quality control chart is a graphic that depicts whether sampled products or processes are meeting their intended specifications and, if not, the degree by which they vary from those specifications. When each chart analyzes a specific attribute of the product it is called a univariate chart.

When a chart measures variances in several product attributes, it is called a multivariate chart. Randomly selected products are tested for the given attribute or attributes the chart is tracking.

In statistical quality control, the np-chart is a type of control chart used to monitor the number of nonconforming units in a sample. It is an adaptation of the p-chart and used in situations where personnel find it easier to interpret process performance in terms of concrete numbers of units rather than the somewhat more abstract proportion.

The np-chart differs from the p-chart in only the three following aspects :

**(1)** The control limits are $n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})}$ where $n$ is the sample size and $\bar{p}$ is the estimate of the long-term process mean established during control-chart setup.

**(2)** The number nonconforming ($np$), rather than the fraction nonconforming ($p$), is plotted against the control limits.

**(3)** The sample size, $n$, is constant.

## *Control Limits*

**Upper Control Limit :** $UCL = n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$

**Lower Control Limit :** $LCL = n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})}$

**Centre Line :** $CL = n\bar{p}$

Where, $n\bar{p} = \dfrac{\textit{Total number of defectives}}{\textit{No of Samples}}$

| Sample No | No of Defectives (np) |
|---|---|
| 1 | 9 |
| 2 | 10 |
| 3 | 12 |
| 4 | 8 |
| 5 | 7 |
| 6 | 15 |
| 7 | 12 |
| 8 | 10 |
| 9 | 8 |
| 10 | 10 |
| 11 | 7 |
| 12 | 13 |

| | |
|---|---|
| | 14 |
| 13 | 15 |
| 14 | 16 |
| 15 | $\sum np = 166$ |
| | |

$$\bar{np} = \frac{\sum np}{k} = \frac{166}{15} = 11.06 \cong 11$$

Upper Control Limit : $UCL = \bar{np} + 3\sqrt{\bar{np}(1-\bar{p})}$

$$UCL = \bar{np} + 3\sqrt{\bar{np}\left(1 - \frac{\bar{np}}{n}\right)} = 11 + 3\sqrt{11\left(1 - \frac{11}{1000}\right)}$$

$$= 11 + 3\sqrt{10.879}$$

$$UCL = 11 + (3 * 3.3) = 20.9$$

$$LCL = \bar{np} - 3\sqrt{\bar{np}\left(1 - \frac{\bar{np}}{n}\right)} = 11 - 3\sqrt{11\left(1 - \frac{11}{1000}\right)}$$

$$= 11 - 3\sqrt{10.879}$$

$$LCL = 11 - (3 * 3.3) = 1.1$$

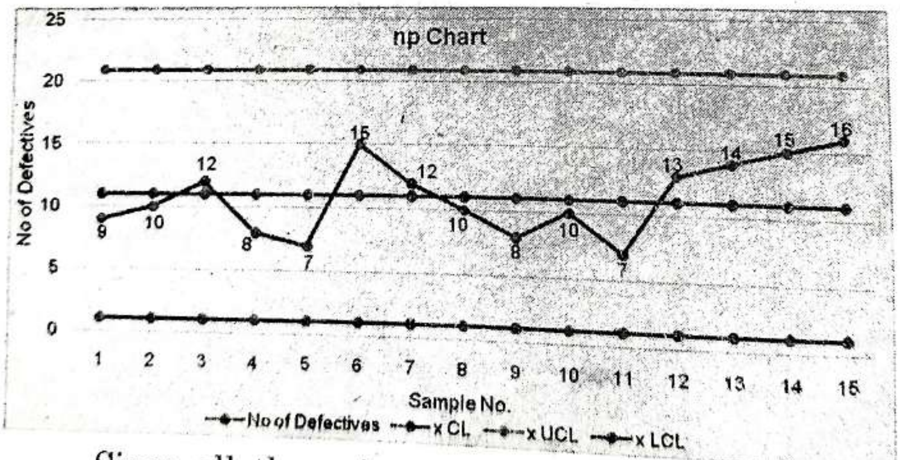
Control Lines :

$$CL = \bar{np} = 11,$$

$$UCL = 20.9$$

$$LCL = 1.1$$

**NP Chart :**




Since all the points lie between the Lower Control Limit and Upper Control Limit, the process is in control.

**14. *What are the advantages of statistical quality control?*** **(2018)**

### Advantages of Statistical Quality Control

(1) It provides a means of detecting error at inspection.

(2) It leads to more uniform quality of production.

(3) It improves the relationship with the customer.

(4) It reduces inspection costs.

(5) It reduces the number of rejects and saves the cost of material.

(6) It provides a basis for attainable specifications.

(7) It points out the bottlenecks and trouble spots.

(8) It provides a means of determining the capability of the manufacturing process.

(9) It promotes the understanding and appreciation of quality control.

**15. *The following table gives the result of inspection of 20 samples of 100 items each taken on working days. Draw a p-chart.*** **(2018)**

| Sample No. | No. of Defectives |
|---|---|
| 1 | 9 |
| 2 | 17 |
| 3 | 8 |
| 4 | 7 |
| 5 | 12 |
| 6 | 5 |
| 7 | 11 |
| 8 | 16 |
| 9 | 14 |
| 10 | 15 |
| 11 | 10 |
| 12 | 6 |
| 13 | 7 |
| 14 | 18 |
| 15 | 16 |
| 16 | 10 |
| 17 | 5 |
| 18 | 14 |
| 19 | 7 |
| 20 | 13 |

In statistical quality control, the p-chart is a type of control chart used to monitor the proportion of nonconforming (defective) units in a sample, where the sample proportion of defective units is defined as the ratio of the number of defective units to the sample size, $n$.

A standard deviation for a p-chart is calculated according to the following equation.:

$$\sigma_p = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$

Where $\overline{p} = \frac{\Sigma p}{n}$ are the average proportion defectives

The control limits are simply :

Control Limit $= CL = \overline{p} = \frac{\Sigma p}{n}$

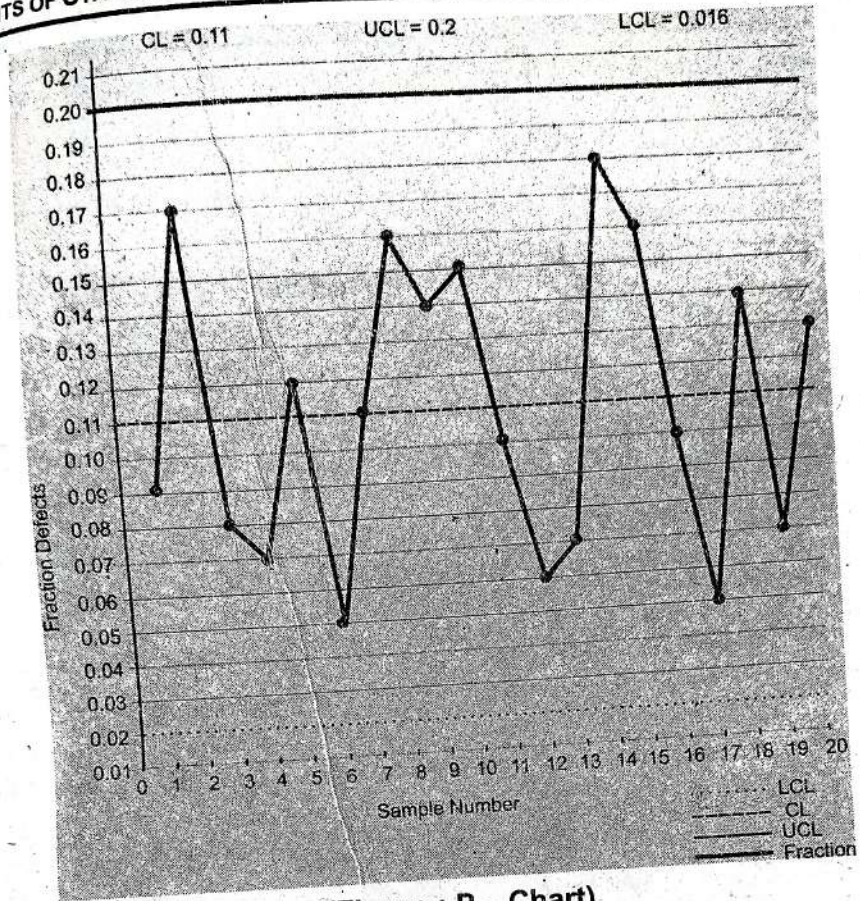
Upper Control Limit $= UCL = \overline{p} + 3 = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$

Lower Control Limit $= LCL = \overline{p} - 3 = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$

However, note that a $p$ value can never be negative, implying that the LCL should never be negative. If the LCL calculates to be negative then use LCL = 0.

Calculating the CL, UCL and LCL :

| Sample No. | No. of Defectives | Proportion Defective (p) p=d/n |
|---|---|---|
| 1 | 9 | 0.09 |
| 2 | 17 | 0.17 |
| 3 | 8 | 0.08 |
| 4 | 7 | 0.07 |
| 5 | 12 | 0.12 |
| 6 | 5 | 0.05 |
| 7 | 11 | 0.11 |
| 8 | 16 | 0.16 |
| 9 | 14 | 0.14 |
| 10 | 15 | 0.15 |
| 11 | 10 | 0.10 |
| 12 | 6 | 0.06 |
| 13 | 7 | 0.07 |
| 14 | 18 | 0.18 |
| 15 | 16 | 0.16 |
| 16 | 10 | 0.10 |
| 17 | 5 | 0.05 |
| 18 | 14 | 0.14 |
| 19 | 7 | 0.07 |
| 20 | 13 | 0.13 |




(Figure : P – Chart)

$$\sum p = 2.20$$

$$\bar{p} = \frac{2.20}{20} = \frac{220}{2000}$$

Therefore, Control limit : $CL = 0.11$

**Upper Control Limit :** $UCL = \bar{p} + 3\sigma_p = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

$$= 0.11 + 3\sqrt{\frac{0.11(1-0.11)}{100}}$$

$$= 0.11 + 3\sqrt{\frac{0.11(0.89)}{100}} = 0.11 + 3\sqrt{0.000979}$$

$$= 0.11 + 3 \times 0.0313$$

$$= 0.2039 = 0.2$$

$$LCL = \bar{p} - 3\sigma_p = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$= 0.11 - 0.0939$$

$$= 0.0161 = 0.016$$

**(Ans.)**

**16.** ◆ ***What are the limitations of statistical quality control?*** **(2017)**

◆ ***Write down the limitations of S.Q.C.***

**Limitations of S.Q.C.**

The readers should not be misguided into believing, that S.Q.C. is a panacea against any evil of production. Many more considerations are to be made in the production of units besides S.Q.C. The main consideration is the requirement of the consumer and the demand in the market. Many more factors are involved and overall responsibility of maintaining the quality and running of the process lies with the production manager. S.Q.C. is a part of the production process rather than the production being a part of S.Q.C.

**17.** ***The following table gives the average daily production figure for 20 months each of 25 working days. Given that the population standard deviation of daily production is 35 units draw a control chart for mean :*** **(2014, 2017)**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 210 | 205 | 210 | 212 | 211 | 209 | 219 | 204 | 212 | 209 |
| 212 | 215 | 208 | 214 | 210 | 204 | 211 | 211 | 203 | 211 |

Control limits for the $\overline{X}$ chart are given by :

$$UCL = \overline{\overline{X}} + 3\frac{\sigma}{\sqrt{n}}$$

$$LCL = \overline{\overline{X}} - 3\frac{\sigma}{\sqrt{n}}$$

Where, UCL and LCL are the upper and lower control limits, $n$ is the subgroup size, and $\sigma$ is the estimated standard deviation of the individual values.

Computing $\overline{\overline{X}}$, $UCL$ and $LCL$ as :

$$\overline{\overline{X}} = \frac{\sum \overline{x}}{n} = CL(\text{Control Limit})$$

$$\sum \overline{x} = 4200$$

Therefore,

$$\overline{\overline{X}} = \frac{4200}{20} \Rightarrow \overline{\overline{X}} = 210$$

$$UCL = \overline{\overline{X}} + \frac{3\sigma}{\sqrt{n}} \Rightarrow \quad UCL = 210 + \frac{3 \times 35}{\sqrt{25}}$$

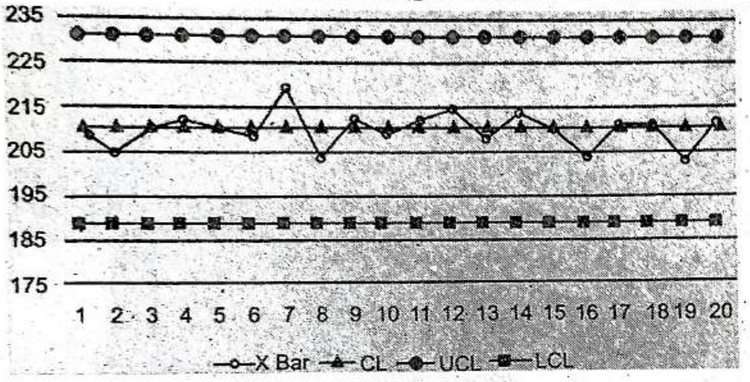
$$UCL = 231$$

$$LCL = \overline{\overline{X}} - \frac{35}{\sqrt{n}}$$

$$UCL = 210 - \frac{3 \times 35}{\sqrt{25}} \Rightarrow UCL = 189$$

Hence, $CL = 210, UCL = 231, LCL = 189.$

**Plotting the Control Chart :** Since all the points lie within the UCL and LCL the process is in Cóntrol.




**(Figure : X Bar Chart)**

18. ***20 Samples each of size 100, of glass vessels, were inspected. The results of inspection are given below :***

| Sample No. | No. of Defects |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 3 |
| 4 | 0 |
| 5 | 2 |
| 6 | 3 |
| 7 | 1 |
| 8 | 2 |
| 9 | 0 |
| 10 | 4 |
| 11 | 3 |
| 12 | 2 |
| 13 | 0 |
| 14 | 4 |
| 15 | 1 |
| 16 | 7 |
| 17 | 0 |

| | |
|---|---|
| 18 | 1 |
| 19 | 3 |
| 20 | 1 |

**(2017)**

***Draw a p chart.***

In statistical quality control, the p-chart is a type of control chart used to monitor the proportion of nonconforming (defective) units in a sample, where the sample proportion of defective units is defined as the ratio of the number of defective units to the sample size, *n*.

A standard deviation for a *p*-chart is calculated according to the following equation :

$$\sigma_p = \sqrt{\frac{p(1-\bar{p})}{n}}$$

Whare *p* is the average proportion defectives

The control limits are simply :

Control Limits : $\text{CL} = \bar{p} = \frac{\sum p}{n}$

Upper Control Limits : $\text{UCL} = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

Lower Control Limits : $\text{LCL} = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

However, note that a *p* value can never be negative, implying that the LCL should never be negative. If the LCL calculates to be negative then use LCL = 0.

Calculating the CL, UCL and LCL :

| Sample No. | No. of Defectives | Proportion Defective (p) |
|---|---|---|
| 1 | 2 | 0.02 |
| 2 | 1 | 0.01 |
| 3 | 3 | 0.03 |
| 4 | 0 | 0 |
| 5 | 2 | 0.02 |
| 6 | 3 | 0.03 |
| 7 | 1 | 0.01 |
| 8 | 2 | 0.02 |
| 9 | 0 | 0 |
| 10 | 4 | 0.04 |
| 11 | 3 | 0.03 |
| 12 | 2 | 0.02 |
| 13 | 0 | 0 |
| 14 | 4 | 0.04 |
| 15 | 1 | 0.01 |

| | | |
|---|---|---|
| 16 | 7 | 0.07 |
| 17 | 0 | 0 |
| 18 | 1 | 0.01 |
| 19 | 3 | 0.03 |
| 20 | 1 | 0.01 |

$\sum p = 0.4$

$$\bar{p} = \frac{0.40}{20} = 0.02$$

Therefore, Control : $CL = 0.02$

Upper Control Limits : $UCL = 0.02 + 3\sqrt{\frac{0.02(1-0.02)}{100}}$

$$= 0.02 + 3\sqrt{\frac{0.02(0.98)}{100}}$$

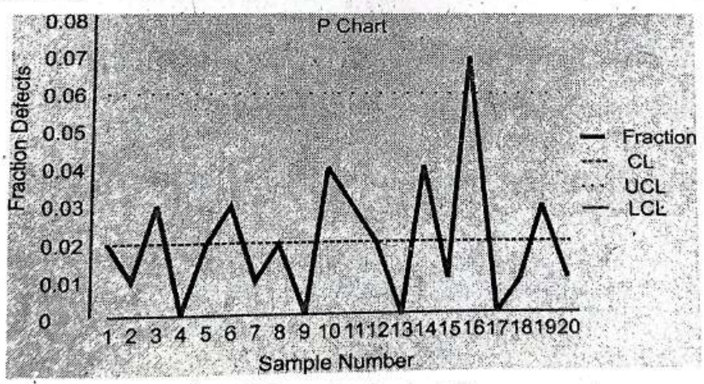
$UCL = 0.02 + 0.042 = 0.062$

Lower Control Limits : $LCL = 0.02 - 3\sqrt{\frac{0.02(1-0.02)}{100}}$

$$= 0.02 - 3\sqrt{\frac{0.02(0.98)}{100}}$$

$LCL = -0.022 = 0$ (because $LCL$ cannot be negative)

**Drawing the $p$ Chart with $CL = 0.02$, $UCL = 0.062$, $LCL = 0$**




**(Figure : P chart)**

The above chart shows that the process is Out of Control as the fraction defect of sample size 16 lies outside the UCL.

**19. *Write a short note on $\bar{X}$ chart.*** **(2015)**

In industrial statistics, the X-bar chart is a type of Shewhart control chart that is used to monitor the arithmetic means of successive samples of constant size, $n$. This type of control chart is used for characteristics that can be measured on a continuous scale, such as weight, temperature, thickness etc. For example, one might take a sample of 5 shafts from production every hour, measure the diameter of each, and then plot, for each sample, the average of the five diameter values on the chart.

For the purposes of control limit calculation, the sample means are assumed to be normally distributed, an assumption justified by the Central Limit Theorem.

The $X$-bar chart is always used in conjunction with a variation chart such as the $x$ and $R$ chart or and $S$ chart. The $R$-chart shows sample ranges (difference between the largest and the smallest values in the sample), while the $S$-chart shows the samples' standard deviation. The $R$-chart was preferred in times when calculations were performed manually, as the range is far easier to calculate than the standard deviation; with the advent of computers, ease of calculation ceased to be an issue, and the $S$-chart is preferred these days, as it is statistically more meaningful and efficient. Depending on the type of variation chart used, the average sample range or the average sample standard deviation is used to derive the $X$-bar chart's control limits.

An $X$ bar-$R$ chart plots the process mean ($X$ bar chart) and process range (R chart) over time for variables data in subgroups. This combination control chart is widely used to examine the stability of processes in many industries.

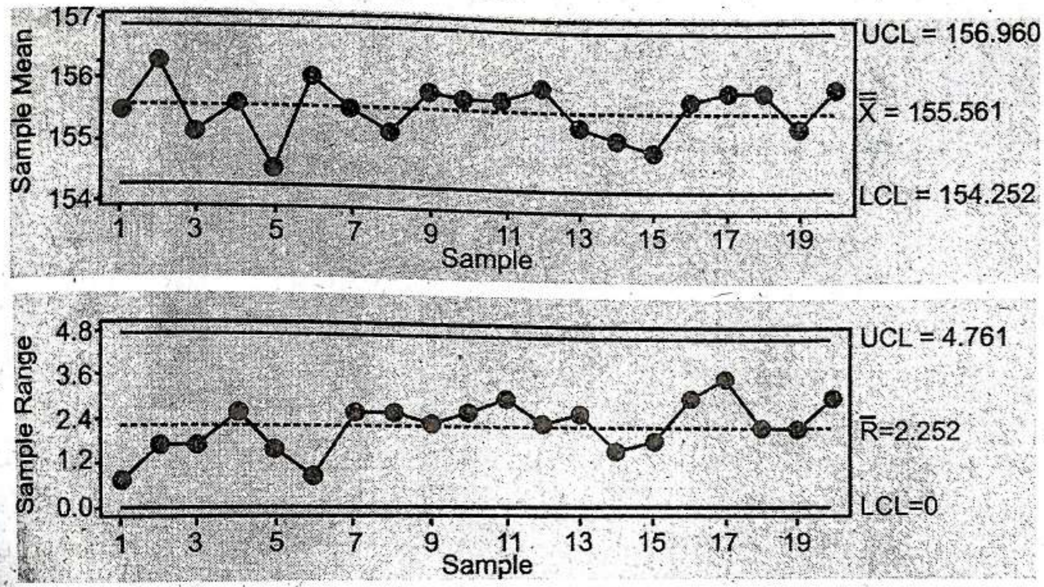
For example, you can use $X$ bar-$R$ charts to monitor the process mean and variation for subgroups of part lengths, call times, or hospital patients blood pressure over time.

The $X$ bar chart and the $R$ chart are displayed together because you should interpret both charts to determine whether your process is stable. Examine the $R$ chart first because the process variation must be in control to correctly interpret the $X$ bar chart. The control limits of the $X$ bar chart are calculated considering both process spread and center. If the $R$ chart is out of control, then the control limits on the $X$ bar chart may be inaccurate and may falsely indicate an out-of-control condition or fail to detect one.

You can use the X bar-$R$ chart when your subgroup size is 8 or less. Use the $X$ bar-$S$ chart when your subgroup size is 9 or more.

## *Example of an X bar-R Chart*

A plastics manufacturer wants to determine whether the production process for a new product is in control. Analysts sample 5 units every hour for 20 hours and assess the strength of the plastic.



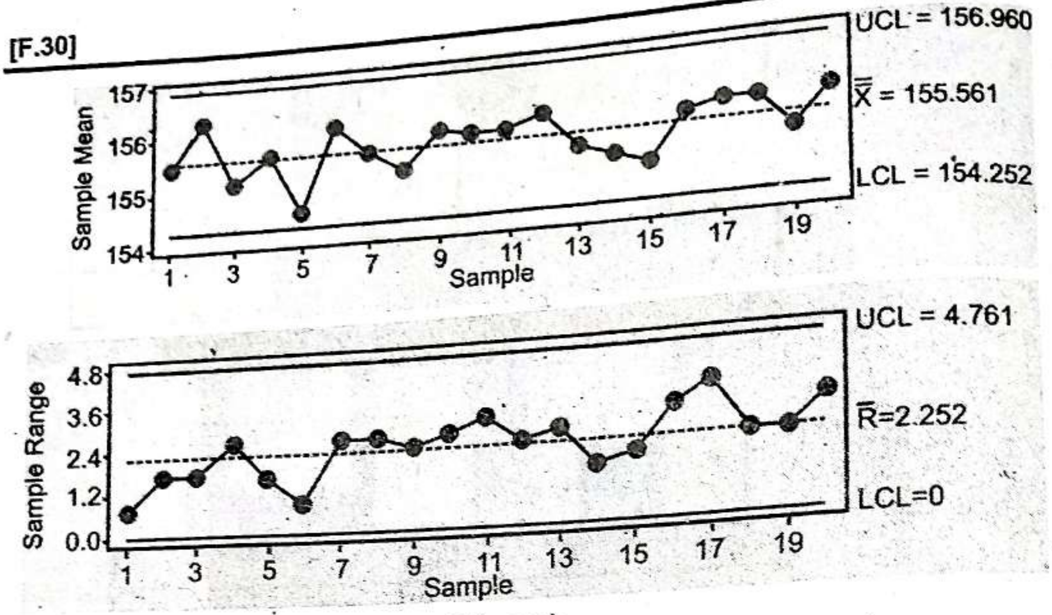

**(Figure)**

The $X$ bar chart and the $R$ chart are displayed together because you should interpret both charts to determine whether your process is stable. Examine the $R$ chart first because the process variation must be in control to correctly interpret the $X$ bar chart. The control limits of the $X$ bar chart are calculated considering both process spread and center. If the $R$ chart is out of control, then the control limits on the $X$ bar chart may be inaccurate and may falsely indicate an out-of-control condition or fail to detect one.

You can use the X bar-$R$ chart when your subgroup size is 8 or less. Use the $X$ bar-$S$ chart when your subgroup size is 9 or more.

**Example of an $X$ bar-$R$ chart :** A plastics manufacturer wants to determine whether the production process for a new product is in control. Analysts sample 5 units every hour for 20 hours and assess the strength of the plastic.

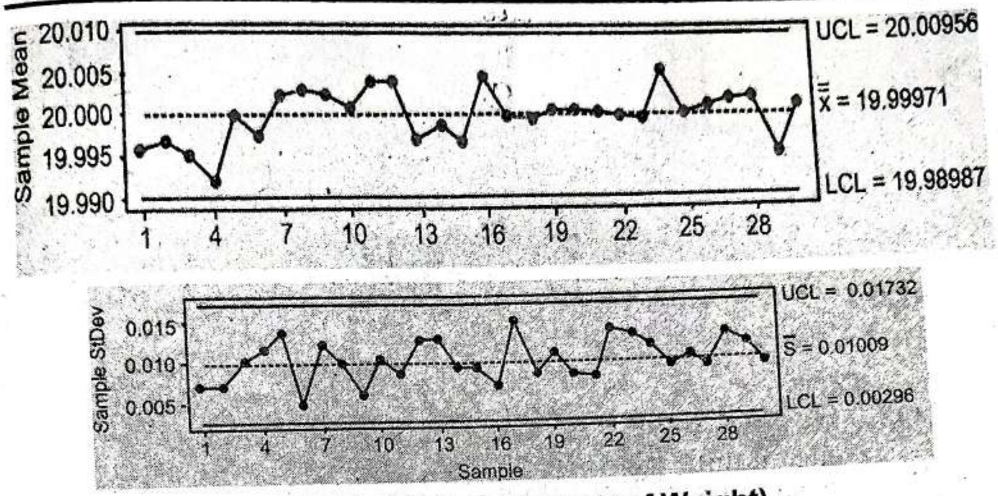
[F.30]




**(Figure)**

The points vary randomly around the center line and are within the control limits for both charts. No trends or patterns are present. The strength of the plastic product is stable across the 20 subgroups. An *X* bar-*S* chart plots the process mean (*X* bar chart) and process standard deviation (*S* chart) over time for variables data in subgroups. This combination control chart is widely used to examine the stability of processes in many industries. For example, you can use *X* bar-*S* charts to examine the process mean and variation for subgroups of part lengths, call times, or hospital patients' blood pressure over time. The *X* bar chart and the *S* chart are displayed together because you should interpret both charts to determine whether your process is stable. Examine the *S* chart first because the process variation must be in control to correctly interpret the *X* bar chart. The control limits of the *X* bar chart are calculated considering both process spread and center. If the S chart is out of control, then the control limits on the *X* bar chart may be inaccurate and may falsely indicate an out-of-control condition or fail to detect one. Use the *X* bar-*S* chart when your subgroup size is 9 or more. You can use the *X* bar-*R* chart when your subgroup size is 8 or less.

**Example of an *X* bar-*S* chart :** A paint manufacturer wants to assess the stability of the process used to fill paint cans. Analysts collect subgroups of 10 cans every hour and use an *X* bar-*S* chart to monitor the mean and variation of the filled paint cans.




**(Figure : X Bar – S Chart of Weight)**

The points vary randomly around the center line and are within the control limits. No trends or patterns are present. The variability in the fill weight is stable across the 30 subgroups.

**20. *Write a short note on the following :*** **(2015)**
**(1) *Histogram*** **(2) *Ogive***

**(1) Histogram :** The histogram is a graph that uses contiguous vertical bars to display the frequency of the data (unless the frequency equals 0) contained in each class. The heights of the bars equal the frequency (after certain scale has been chosen) and the bases of the bars lie on the corresponding class.

**Steps for Constructing a Histogram :**

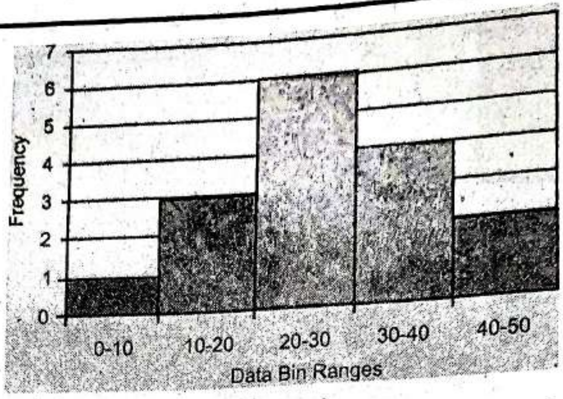
**(a)** Draw and label the $x$ (horizontal) and the $y$ (vertical) axes.

**(b)** Represent the frequencies on the $y$ axis and the class boundaries on the $x$ axis.

**(c)** Using the frequencies as the heights draw vertical bars for each class.

**Example :** Lets draw a Histogram from the following Distribution :

| Data Range | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 |
|---|---|---|---|---|---|
| Frequency | 1 | 3 | 6 | 4 | 2 |




(Figure)

(2) **Ogive Curve :** An ogive is a graph that represents the cumulative frequencies for the classes in a frequency distribution. It shows how many of values of the data are below certain boundary.

**Steps for Constructing an Ogive :**

(a) Draw and label the $x$ (horizontal) and the $y$ (vertical) axes.

(b) Represent the cumulative frequencies on the $y$ axis and the class boundaries on the $x$ axis.

(c) Plot the cumulative frequency at each upper class boundary with the height being the corresponding cumulative frequency.

(d) Connect the points with segments. Connect the first point on the left with the $x$ axis at the level of the lowest lower class boundary.

**Note :** For the ogive we need the class boundaries and the cumulative frequencies.

**Example :** Draw Ogive Curve from the following distribution :

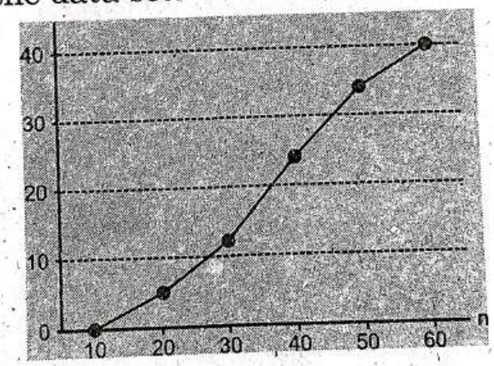
| Class | Frequency | Cumulative Frequency |
|---|---|---|
| 10 - 20 | 5 | 5 |
| 20 – 30 | 7 | 12 |
| 30 – 40 | 12 | 24 |
| 40 – 50 | 10 | 34 |
| 50 - 60 | 6 | 40 |

**Plot the Ogive :** The first coordinate in the plot always starts at a $y$ – value of 0 because we always start from a count of zero. So, the first coordinate is at (10; 0) – at the beginning of the first interval. The second coordinate is at the end of the first interval (which is also the beginning of the second interval) and at the first cumulative count, so (20; 5). The third coordinate is at the end of the second interval and at the second cumulative count, namely (30; 12), and so on.

Computing all the coordinates and connecting them with straight lines gives the following ogive.

Ogives do look similar to frequency polygons, which we saw earlier. The most important difference between them is that an ogive is a plot of cumulative values, whereas a frequency polygon is a plot of the values themselves. So, to get from a frequency polygon to an ogive, we would add up the counts as we move from left to right in the graph.

Ogives are useful for determining the median, percentiles and five number summaries of data. Remember that the median is simply the value in the middle when we order the data. A quartile is simply a quarter of the way from the beginning or the end of an ordered data set. With an ogive we already know how many data values are above or below a certain point, so it is easy to find the middle or a quarter of the data set.




**(Figure)**

**21.** ***The following data relate to faculty wise enrolment in college.***

| Faculty : | Science | Arts | Commerce | Total |
|---|---|---|---|---|
| No. of Students | 2,010 | 1,100 | 2,390 | 5,500 |

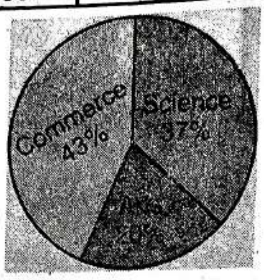
***Represent the data by a pie diagram.*** **(2015)**

To represent the data in pie chart we need to calculate the Percentage Frequency and then convert it to Degrees as :

Percentage Frequency = Relative Frequency × 100 = (Absolute Frequency/ Total Frequency) × 100

Degree = Relative Frequency × $360^0$

| Faculty | Students | Percentage Students | Degree |
|---|---|---|---|
| Science | 2010 | 36.6 | 131 |
| Arts | 1100 | 20.0 | 72 |
| Commerce | 2390 | 43.4 | 157 |
| **Total** | **5500** | **100** | **$360^0$** |




**(Figure : Faculty Wise Enrolment)**

❑❑

# BCA

## (SEM. III) EXAMINATION, DEC, 2019

## BCA – 305 (N) : ELEMENTS OF STATISTICS

*Time : Three Hours* *Maximum Marks : 75*

**Note :** Attempt questions from *all* Sections as directed.

**Inst. :** The candidates are required to answer only in serial order. If there are many parts of a question, answer them in continuation.

### SECTION – A
### (Short Answer Type Questions)

**Note : Attempt *All* questions. Each question carries 3 marks.**

1. (A) Write a short note on scope of Statistics.

(B) Write the objects of classification.

(C) The mean marks of 50 students in Section A is 40 and mean marks of 40 students in Section B is 45. Find the combined mean of 90 students in both the Sections.

(D) The arithmetic mean of two observations is 25 and the harmonic mean is 9. What is geometric mean of the series?

(E) Find the range and coefficient of range for the following data :

| Variable | 83 | 88 | 93 | 98 | 103 | 108 |
|---|---|---|---|---|---|---|
| Frequency | 6 | 8 | 9 | 13 | 5 | 3 |

(F) How many words can be formed from the letters of the word 'SUNDAY'? How many of these begin with D?

(G) If ${}^nP_r = 1680$ and ${}^nC_r = 70$, find r.

(H) In a production of certain rods, a process is said to be in control if outside diameters have a mean 2.5 and a S. D. of 0.004. Find the control limits for the mean of random samples of size 4.

(I) Explain the following :
(i) Mutually exclusive event
(ii) Conditional probability

### SECTION – B
### (Long Answer Type Questions)

**Note : Attempt *any two* questions. Each question carries 12 marks.**

2. Find less than and more than cumulative frequencies and draw Ogives' from the following data :

| Class Interval | Frequency |
|---|---|
| 100 – 110 | 2 |
| 110 – 120 | 3 |
| 120 – 130 | 7 |
| 130 – 140 | 11 |
| 140 – 150 | 15 |
| 150 – 160 | 7 |
| 160 – 170 | 2 |
| 170 – 180 | 3 |

3. Compute quartile deviation and its coefficient from the information given below :

| Mid Value | Frequency |
|---|---|
| 5 | 5 |
| 15 | 8 |
| 25 | 15 |
| 35 | 20 |
| 45 | 16 |
| 55 | 10 |
| 65 | 6 |

4. Calculate median and mode for the following frequency distribution :

| Marks | Frequency |
|---|---|
| 25 – 30 | 22 |
| 30 – 35 | 34 |
| 35 – 40 | 50 |
| 40 – 45 | 42 |
| 45 – 50 | 38 |
| 50 – 55 | 14 |

5. Find standard deviation and coefficient of variation of the following data :

| Class Interval | Frequency |
|---|---|
| 0 – 10 | 10 |
| 10 – 20 | 15 |
| 20 – 30 | 25 |
| 30 – 40 | 25 |
| 40 – 50 | 10 |
| 50 – 60 | 10 |
| 60 – 70 | 5 |

## SECTION – C
### (Long Answer Type Questions)

**Note : Attempt *any two* questions. Each question carries 12 marks.**

6. (a) Prove that :

$$P(n,r) = P(n-1,r) + r.P(n-1, r-1)$$

(b) Find the number of ways of which 5 boys and 3 girls can be seated in a row so that no two girls are together.

7. (a) A box contains 4 different red and 6 different white balls. In how many ways can 6 balls be selected so that there are at least two balls of each colour?

(b) What is the chance of throwing a total of 5 or 10 with two dice?

8. (a) A problem in statistics is given to the three students A, B and C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ respectively. What is the probability that the problem will be solved?

(b) A bag contains 3 red and 4 black balls and a second bag contains 5 black balls and 4 red balls. A ball is drawn from each bag. Find the probability that one is red and other is black.

9. The following table gives the result of inspection of 15 samples of 1000 items each taken of working days. Draw a np – chart :

| Sample No. | No. of Defectives |
|---|---|
| 1 | 9 |
| 2 | 10 |
| 3 | 12 |
| 4 | 8 |
| 5 | 7 |
| 6 | 15 |
| 7 | 12 |
| 8 | 10 |
| 9 | 8 |
| 10 | 10 |
| 11 | 7 |
| 12 | 13 |
| 13 | 14 |
| 14 | 15 |
| 15 | 16 |

❑❑

# BCA

## (SEM. III) EXAMINATION, MARCH, 2021

## BCA – 305 (N) : ELEMENTS OF STATISTICS

*Time : Three Hours* *Maximum Marks : 75*

**Note :** Attempt questions from *all* Sections as directed.

**Inst. :** The candidates are required to answer only in serial order. If there are many parts of a question, answer them in continuation.

### SECTION – A
### (Short Answer Type Questions)

**Note : Attempt *all* questions. Each question carries 3 marks.**

1. (A) What do you mean by cumulative frequency distribution.
   (B) Discuss the importance of classification in statistics.
   (C) The mean of 200 items was 50. Later on it was discovered that two items were misread as 92 and 8 instead of 192 and 88. Find out the correct mean.
   (D) Find the geometric mean of 4, 8, 16.
   (E) The smallest value in a set of observations is 10 with a range of 25. Find the largest observation and the co – efficient of range.
   (F) Find n, if :
   $$(n+1)! = 12 \times (n+1)!$$
   (G) Find the number of permutations of the letters of the word 'EXAMINATION'?
   (H) A bag contains tickets numbered from 1 to 25. Two tickets are drawn. Find probability that both the numbers are prime.
   (I) Write a short note of statistical quality control.

### SECTION – B

**Note : Attempt *any Two* questions. Each question carries 12 marks.**

2. Find less than and more than cumulative frequencies and draw 'Ogives' from the following data :

| Class | Frequency |
|---|---|
| 0-9 | 5 |
| 10-19 | 15 |
| 20-29 | 18 |
| 30-39 | 30 |
| 40-49 | 15 |
| 50-59 | 10 |
| 60-69 | 5 |
| 70-79 | 2 |

3. Find median and mode from the following distribution :

| Income (₹) | No. of Persons |
|---|---|
| 100-200 | 15 |
| 100-300 | 33 |
| 100-400 | 63 |
| 100-500 | 83 |
| 100-600 | 100 |

**4.** Find standard deviation and coefficient of variation from the following data :

| Marks | No. of Students |
|---|---|
| 0-10 | 10 |
| 10-20 | 15 |
| 20-30 | 25 |
| 30-40 | 30 |
| 40-50 | 15 |
| 50-60 | 10 |
| 60-70 | 5 |

**5.** Calculate mean deviation from median and its coefficient from the following data :

| Size | Frequency |
|---|---|
| 0-10 | 4 |
| 10-20 | 8 |
| 20-30 | 11 |
| 30-40 | 15 |
| 40-50 | 11 |
| 50-60 | 7 |
| 60-70 | 4 |

## SECTION – C

**Note : Attempt *any* 2 questions. Each question carries 12 marks.**

**6.** (a) In how many ways can 10 books be arranged on a shelf so that a particular pair of books shall be :

(i) Always together

(ii) Never together

(b) If $1 \le r \le n$, prove that

$$c(n,r)+c(n,r+1)=c(n+1,r+1)$$

**7.** (a) A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is1/7 and that of wife's selection is 1/5. What is the probability that

(i) Both of them will be selected

(ii) Only one of them will be selected

(b) A bag contains 4 white and 2 black balls and a second bag contains 3 of each colour. A bag is selected at random, and a ball is drawn at random from the bag chosen. What is the probability that the ball drawn is white?

8. The following table gives the result of inspection of 15 samples of 100 items each taken on working days. Draw a n. p. chart.

| Sample No. | No. of Detectives |
|---|---|
| 1 | 9 |
| 2 | 17 |
| 3 | 8 |
| 4 | 7 |
| 5 | 12 |
| 6 | 5 |
| 7 | 11 |
| 8 | 16 |
| 9 | 14 |
| 10 | 15 |
| 11 | 10 |
| 12 | 6 |
| 13 | 7 |
| 14 | 8 |
| 15 | 5 |

9. The following data shows the value of samples mean X and range R for 10 samples of size 5 each. Calculate the values for control line and control limits for mean chart and range chart and determine whether process is under control

| Sample No. | Mean (x) | Range (R) |
|---|---|---|
| 1 | 15 | 7 |
| 2 | 17 | 7 |
| 3 | 15 | 4 |
| 4 | 18 | 9 |
| 5 | 17 | 8 |
| 6 | 14 | 7 |
| 7 | 18 | 12 |
| 8 | 15 | 4 |
| 9 | 17 | 11 |
| 10 | 16 | 5 |

Give for

$$n = 5, A_2 = 0.577, D_3 = 0, D_4 = 2.115$$

□□

ELEMENTS

Time : 2

Note : T r p a p

Note : A as m

1. (A) (B) (C)

(D) (E)

(F)

(G)

(H)

(I)

Note : T i i

2. B

# BCA

## (SEM. III) EXAMINATION, 2022-23

## BCA – 3005 : ELEMENTS OF STATISTICS

*Time : 2 Hours* *Maximum Marks : 75*

**Note : This paper consists of three Sections A, B and C. Carefully read the instructions of each Section in solving the questions paper. Candidates have to write their answers in the given answer-copy only. No separate answer copy (B-Copy) Will be provided.**

### SECTION – A

### (Short Answer Type Questions)

**Note : All questions are compulsory. Answer the following questions as short answer type questions. Each question carries 5 marks.**

1. (A) Write about difference between primary and secondary data.
   (B) Discuss the importance of classification in statistics.
   (C) The average marks of 100 students were found to be 30. Later, it was discovered that a score of 32 was misread as 23. Find the corrected mean of the 100 students.
   (D) Find the harmonic mean of 8, 16, 24.
   (E) The smallest value in a set of observations is 20 with a range of 45. Find the largest observation and the co-efficient of range.
   (F) In how many ways can a student choose 5 courses out of 8 course if 2 courses are compulsory for every student?
   (G) How many different words can be formed with the letters of the word 'MISSISSIPPI'?
   (H) If two balls are drawn from a bag containing 2 white, 4 red and 5 black balls, what is the chance that both the balls, what is the chance that both the balls are black?
   (I) Write a short note on statistical quality control.

### SECTION – B

### (Long Answer Type Questions)

**Note : This section contains four questions from which one question is to be answered as long question. Each question carries 15 marks.**

2. By Grouping method locate mode from the following data :

| Class | Frequency |
|---|---|
| 0-5 | 5 |
| 5-10 | 7 |
| 10-15 | 9 |
| 15-20 | 18 |
| 20-25 | 16 |

| | |
|---|---|
| 25-30 | 15 |
| 30-35 | 6 |
| 35-40 | 3 |

Or

3. Prepare "Less than" and "More than" cumulative frequency distribution. Also draw ogives from the following data :

| Weight (in kg ) | Frequency |
|---|---|
| 30-35 | 12 |
| 34-40 | 18 |
| 40-45 | 27 |
| 45-50 | 20 |
| 50-55 | 17 |
| 55-60 | 6 |
| 60-65 | 5 |

Or

4. Calculate semi-interquartile range and coefficient of quartile deviation from the following data :

| Marks | No. of students |
|---|---|
| 0-10 | 6 |
| 10-20 | 5 |
| 20-30 | 8 |
| 30-40 | 15 |
| 40-50 | 7 |

Or

5. Calculate standard deviation and coefficient of variation from the following data giving the age distribution of 542 members :

| Age group (in years) | No. of members |
|---|---|
| 20-30 | 3 |
| 30-40 | 61 |
| 40-50 | 132 |
| 50-60 | 153 |
| 60-70 | 140 |
| 70-80 | 51 |
| 80-90 | 2 |

## SECTION – C
### (Long Answer Type Question)

**Note : This section contains four questions from which one question is to be answered as long question. Each question carries 15 marks.**

6. (a) There are 5 men and 4 ladies to dine at a round table. In how many ways can they seat themselves so that no two ladies are together?

(b) If $1 \le r \le n$, prove that :

$$C(n,r)+C(n,r-1)=C(n+1,\ r)$$

Or

7. (a) A pair of dice is tossed twice. Find the probability of scoring 7 points :
   (i) once
   (ii) at least once

   (b) A problem in statistics is given to the three students A, B and C Whose chance of solving it are $\frac{1}{2}, \frac{2}{3}, \frac{1}{4}$ respectively. What is the probability that the problem is solved?

Or

8. The number of defects in 18 rolls of cloth each of 150 meters length is given by 3, 5, 8, 9, 4, 2, 5, 9, 6, 4, 8, 12, 7, 5, 10, 10, 7 and 5. Draw C – chart and give your comment.

Or

9. The following table gives the result of inspection of 15 samples of 100 items each taken on working days. Draw a np-chart :

| Sample No. | No. of Defectives |
|---|---|
| 1 | 9 |
| 2 | 17 |
| 3 | 8 |
| 4 | 7 |
| 5 | 12 |
| 6 | 5 |
| 7 | 11 |
| 8 | 16 |
| 9 | 14 |
| 10 | 15 |
| 11 | 10 |
| 12 | 6 |
| 13 | 7 |
| 14 | 18 |
| 15 | 10 |

□□

[G.10]

# BCA

## (SEM. III) EXAMINATION, 2023-24

## BCA – 3005 : ELEMENTS OF STATISTICS

Time : Two Hours | Maximum Marks : 75

**Note :** This paper consists of three Section A, B and C. Carefully read the instructions of each Section in solving the question paper. Candidates have to write their answers in the given answer-copy only. No separate answer – copy (B copy) will be provided.

### SECTION – A
### (Short Answer Type Questions)

**Note :** ***All*** **questions are compulsory. Answer the following questions as short answer type questions. Each question carries 5 marks.**

1. (A) Write a short note on discrete and continuous distribution?
(B) What do mean by distrust of Statistics?
(C) The average marks of 100 students were found to be 50. Later, it was discovered that a score of 64 was misread as 84. Find the corrected mean of the 100 students.
(D) Find the geometric mean of 4, 9, 12 and 48.
(E) Compare between mean deviation and standard deviation.
(F) Find the number of Permutations of the letters of the word 'MATHEMATICS'. How many of these begin with H and end with S.
(G) If P(A) = 0.3 and $P(A \cup B) = 0.7$. Find P(B), If A and B are mutually exclusive.
(H) A bag contains 8 balls of which 5 are white and 3 are black. Two balls are drawn at random. What is the probability that both are white?
(I) In the production of certain rods a process is said to be in control if the outside diameters have a mean 2.532 and standard deviation of 0.002. Find the central limits for the mean of random samples of size 4.

### SECTION – B
### (Long Answer Type Questions)

**Note : This section contains four questions from which one question is to be answered as long question. Each question carries 15 marks.**

2. By using grouping method, locate Mode from the following data.

| Mid value | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|---|---|---|---|---|---|---|---|
| Frequency | 7 | 12 | 17 | 29 | 31 | 5 | 3 |

(Or)

3. Find 'less than' and 'more than' cumulative frequencies and draw 'Ogives' from the following data :

| Weight (in kg) | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 |
|---|---|---|---|---|---|---|---|
| Frequency | 3 | 5 | 12 | 18 | 14 | 6 | 2 |

(Or)

4. Calculate interquartile range. Quartile Deviation and coefficient of quartile deviation from the following data :

| Marks | 4-8 | 9-13 | 14-18 | 19-23 | 24-28 |
|---|---|---|---|---|---|
| No. of Students | 3 | 4 | 3 | 2 | 4 |

(Or)

5. Calculate Standard deviation and coefficient of variation from the following data :

| Class Interval | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 |
|---|---|---|---|---|---|---|
| Frequency | 6 | 12 | 15 | 28 | 20 | 14 |

## SECTION – C

### (Long Answer Type Questions)

**Note : This section contains four questions from which one question is to be answered as long question. Each question carries 15 marks.**

6. (a) Find the number of ways in which 5 boys and 3 girls can be seated in a row so that no two girls are together.

   (b) Three group of children contain respectively 3 girls and 1 boy, 2 girls and 2 boys and 1 girl and 3 boys. One child selected at random from each group. Show that the chance that the three selected consist of 1 girl and 2 boys is $\frac{13}{32}$

(Or)

7. (a) Two dice are thrown. Find the probability that at least a sum of 10 occurs.

   (b) A box contains 5 different red and 6 different white balls. In how many ways can 6 balls be selected so that there are at least two balls of each colour.

(Or)

8. Mobile charger supplier drawn randomly constant sample size of 500 chargers every day for quality control test. Defects in each charger are recorded during testing. 12, 14, 16, 18, 16, 14, 12, 12, 32, 16, 18, 16, 12, 14, 16, 18, 12, 19, 18, and 21.
Draw control chart for the number of defects (C-chart) and give your comments.

(Or)

9. The data shows the sample mean and range for 10 samples for size 5 each. Draw mean chart, range chart and comment on the state of control of the process.

| Sample No. | Mean $(\overline{X})$ | Range (R) |
|---|---|---|
| 1 | 21 | 5 |
| 2 | 26 | 6 |
| 3 | 23 | 9 |
| 4 | 18 | 7 |
| 5 | 19 | 4 |
| 6 | 15 | 7 |
| 7 | 14 | 4 |
| 8 | 20 | 9 |
| 9 | 16 | 8 |
| 10 | 16 | 6 |

Assume for n = 5, A2 = 0.58, D3 = 0 and D4 = 2.115. ❑❑