Question Bank for Statistics Part – 1

Chapter # 1: Introduction to Statistics

1. **Define statistics.**

Ans: It is a branch of science that deals in collection, processing, presentation, analysis and interpretation of numerical data in order to make decision.

2. **Define population and sample.**

Ans:
Population:
The total number of objects having some common characteristics is called as population.
e.g. (number of cars, number of trees, number of chairs)

Sample:
Any small part of population showing some common characteristics is called as sample.

3. **Write down characteristics of statistics.**

There are some important characteristics of statistics:

- Statistics are aggregate of facts.
- Statistics are numerically expressed.
- Statistics are collected in a systematic manner.
- Statistics are collected with a definite purpose.

4. **Define data.**

Ans: Data are the individual pieces of factual information recorded and used for the purpose of analysis. It is the raw information from which statistics are being created.

5. **What is primary data?**

Ans: The data which has just been collected from the source and has not gone through any kind of statistical treatment like sorting and tabulation is called as primary data.

Example: The data in the population census reports are primary because these ae collected, compiled and published by the population census organization.

6. **Define secondary data.**

Ans: The data which has already been collected by someone, that has undergone a statistical treatment like sorting and tabulation is called as secondary data.

Example: The data in economic survey of Pakistan is secondary because these are originally collected by the Federal Bureau of statistics, the State Bank of Pakistan.

7. **Sources of Primary data.**

Ans: Primary data is collected by the following sources:

i. Direct personal observation.
ii. Registration.
iii. Investigation through enumerators.
iv. Information through mailed questionnaire.

v. Through local correspondents.
vi. Through telephone.

8. Sources of secondary data.

Ans: Secondary data is collected by the following sources:

i. Through government organizations.
ii. Through semi-government organizations.
iii. Through teaching and research organizations.
iv. Through newspapers.
v. Through internet.

9. Describe any two uses of statistics.

Ans: the following are the uses of statistics:

i. It helps in collection of data.
ii. It is used for presentation of data.
iii. It helps in processing of data.
iv. It is used for comparison of data.

10. Define variable.

Ans: A measureable quantity which varies from one individual or object to another is called variable. For example: Weight, height, time etc.

11. What is constant?

Ans: The value which remains the same from person to person is called as constant. For example, value of pi.

12. How many types of variable?

Ans: There are two types of variables:

1- Quantitative variable 2- Qualitative variable

Quantitative variable: The variables which can be expressed numerically with or without units are known as quantitative variables. For example: Time, Heights, weights, etc.

Qualitative variable: The variable which can be expressed in the form of qualities like: eye color, hair color, IQ level etc.

13. Define discrete variable.

Ans: The variable which can be countable is called discrete variable. For example, number of cars, number of chairs in the classroom, number of houses in the street etc.

14. Define continuous variable.

Ans: A measureable variable is known as continuous variable. for example, height, weight, length etc.

15. Define inferential statistics.

Ans: The phase of statistics that is concerned with the procedures and methodology for obtaining valid conclusions is called inferential statistics.

16. Define descriptive statistics.

Ans: descriptive statistics deals with collection of data, its presentation in various forms, such as tables, graphs, diagrams, averages and other measures which would describe the data. For example, businessmen make use of descriptive statistics in presenting their annual reports.

Chapter # 2: **Presentation of data**

1. **What do you meant by the term 'classification'?**

Ans: The process of arranging data into classes or categories according to some common characteristics present in the data is called classification.

Examples:

1- Sorting of letters in a post office, the letters are classified according to the cities

2- The students of the college are classified according to their hair color.

3- The students of the university are classified according to their heights.

2. **Write down the types of classification.**

Ans: The data may be classified according to one, two or many characteristics:

i. One-way Classification: When the data is classified by one characteristic, it is called one way classification.
ii. Two-way classification: When the data is classified by two characteristics, it is called two-way classification.
iii. Many-ways classification: When the data is classified by many characteristics, it is called many-way classification.

3. **How many forms of classification?**

Ans: There are four main forms of classification:

i. Quantitative Classification.
ii. Qualitative Classification.
iii. Geographical Classification.
iv. Chronological or temporal Classification.

Quantitative Classification: When the data is classified by quantitative characteristics, it is called quantitative classification. For example, weight, height, income, etc.
Qualitative Classification: When the data is classified by qualitative characteristics, it is called qualitative classification. For example, sex, religion, color, intelligence, etc.
Geographical Classification: When the data is classified by geographical regions or locations, it is called geographical or spatial classification. For example, provinces, divisions, districts, cities, etc.

Chronological Classification: When the data is classified according to its time of occurrence, it is called chronological or temporal classification. For example, years, months, weeks, days, etc.

4. **Define tabulation.**

Ans: The process of arranging data into rows and columns is called tabulation.

5. **Differentiate between class limits and class boundaries.**

Ans:

Class Limit: Each class start from a lower limit and ends at an upper limit.

For Example:

Class	20 ----30	30 ------40	40 --------50	50 --------60

In the above example 20 is lower limit and 30 is upper limit of the first class and same goes for the others.

Class Boundary: Class limits are called class boundaries if the upper limit of 1st class equals to the lower limit of 2nd class and so on. Hence if classes progress without break, class limits are called class boundaries. In case we don't have equal classes then it can be converted easily by increasing the upper class limit and decreasing the lower class limit by the same amount so that there are no gaps left among the classes. For example add 0.5 in upper class and 0.5 in lower class.

Class	20 ----29	30 ------39	40 --------49	50 --------59
Class Boundaries	19.5 ------29.5	29.5 -------39.5	39.5 ------- 49.5	49.5 ---------59.5

6. **What is meant by relative frequency of a class?**

Ans: The frequency of a class divided by total frequency of the class is called relative frequency.

7. **Define frequency distribution.**

Ans: A frequency distribution is a tabular arrangement of data in which various items are arranged into classes and the number of items falling in each class (Called class frequency).

8. **Define class mark.**

Ans: The class mark or the midpoint is that value which divides a class into two equal parts. It is obtained by adding the lower and upper class limits or class boundaries of a class and dividing the resulting total by 2.

9. **Define histogram.**

Ans: A histogram consists of a set of adjacent rectangles having class boundaries along the x-axis and frequencies along y-axis.

10. **What is the table?**

Ans: A systematic arrangement of data into rows and columns is called table.

11. **What is grouped data?**

Ans: Data presented in the form of a frequency distribution is called grouped data.

12. **Write down the main parts of the table.**

Ans: Following are the different parts of a table out of which first four are main part:

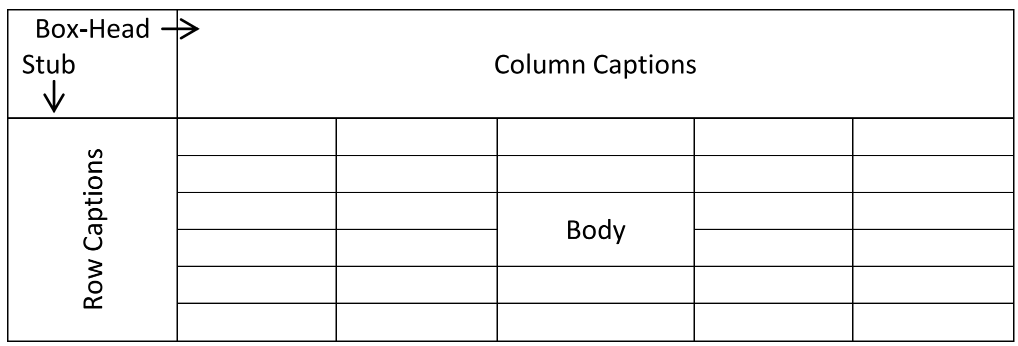
i. Title
ii. Column caption & box head
iii. Row caption & stub
iv. Body of the table
v. Prefatory note

vi. Foot note

vii. Source note

(………………. Title ……………….)

Prefatory Note



●●Footnote

●●Sourcenote

13. What is simple classification?

Ans: When the data is classify according to one characteristic and then it is known as simple classification.

Chapter # 3: **Measures of central location**

1. **What is meant by measures of central tendency?**

Ans: The averages tend to lie in the center of a distribution they are called measures of central tendency. They are also called measures of location because they locate the center of a distribution.

2. **Write the types of averages.**

Ans: The most commonly used averages are:

i. Arithmetic mean
ii. Geometric mean
iii. Harmonic mean
iv. Median
v. Mode

3. **What are two qualities of a good average?**

Ans: Properties of a good average are given below:

i. It is well defined
ii. It is easy to calculate
iii. It is easy to understand
iv. It is based on all the values
v. It is capable of mathematical treatment

4. **Mean of 5 values is 70. Find the sum of values.**

Ans: $\bar{X} = \frac{\sum x}{n}$

$\sum x = 350$ **Ans.**

5. **In a moderately skewed distribution, the values of mean and median are 120 and 110 respectively, find the value of mode.**

Ans: $Mode = 3Median - 2Mean$

$Mode = 3(110) - 2(120)$

$Mode = 90$ **Ans.**

6. **Given *u = (x – 170)/5, ∑fu = 100, ∑f = 200*, find arithmetic mean.**

Ans: By coding method:

$\bar{X} = A + \frac{\sum fU}{\sum f} \times h$ Where; $\left(U = \frac{x-A}{h}\right)$

$\bar{X} = 172.5$ **Ans.**

7. **Write down any two mathematical properties of arithmetic mean.**

Ans: There are following mathematical properties of Arithmetic Mean:

i. The sum of deviations of all observations from their mean is zero. i.e. $\sum(X - \bar{X}) = 0$
ii. The sum of squares of deviations of all observations from their mean is minimum. i.e. $\sum(X - \bar{X})^2$ is minimum

iii. The mean of a constant is constant itself. i.e. If $X = a$ then $\bar{X} = a$

iv. The mean is affected by change of origion and scale. If we add or subtract a constant from all the values or multiply or divide all the values by a constant, the mean is affected by the respective change.

i.e. If $Y = X \pm a$ then $\bar{Y} = \bar{X} \pm a$, If $Y = a \pm bX$ then $\bar{Y} = a \pm b\bar{X}$, If $Y = \frac{X}{a}$ then $\bar{Y} = \frac{\bar{X}}{a}$

8. Define mode and give its formula in case of grouped data.

Ans: The most repeated value in a data is called mode. It is denoted by $\hat{X}$

Formulas;

For ungrouped data: $\hat{X} = \text{The most frequent value in a data}$

For grouped data: $\hat{X} = l + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$

9. Find the mode of 3, 3, 7, 8, 10, 11, 10, 12, and 10.

Ans: 10 is most frequent value in the given data and then called the mode.

10. Define the median with formula.

Ans: The value which divides the ordered data into two equal parts is called as median. It is denoted by $\tilde{X}$.

11. Write down the properties of median.

Ans:

1. a constant "a" is added to each of the n observations y_1, y_2, ……. Y_n having median M, then the median of y_1+a, y_2+a, ……. y_n+a would be "a+M".
2. The sum of the absolute deviations of the observations from their median is minimum i.e.,

∑|y – median| is minimum

3. For a symmetrical distribution median is equidistant from the first and third quartiles i.e.,

Q_3 – Median = Median = Median – Q_1

12. What are the advantages and disadvantages of median?

Ans:

Merits of Median:

1. It is quick to find.
2. It is not much affected by exceptionally large or small values in a data.
3. It is suitable for skewed distribution.

Demerits of Median:

1. It is not rigidly defined.
2. It is not readily suitable for algebraic development.
3. It is less stable in repeated sampling experiments than the mean.
4. It is not based on all the observations.

13. **Define quartiles; also write down its formulas.**

Ans: The values which divide the ordered data into four equal parts are called quartiles. There are 3 quartiles; Q_1 is called first quartile or lower quartile.

For ungrouped data: $Q_1 = \left(\frac{n+1}{4}\right)^{th}\ value$

For grouped data: $Q_1 = l + \frac{h}{f}\left[\frac{n}{4} - c\right]$

Q_2 is called second quartile or median.

For ungrouped data: $Q_2 = 2\left(\frac{n+1}{4}\right)^{th}\ value$

For grouped data: $Q_2 = l + \frac{h}{f}\left[\frac{2n}{4} - c\right]$

Q_3 is called third quartile or upper quartile.

For ungrouped data: $Q_3 = 3\left(\frac{n+1}{4}\right)^{th}\ value$

For grouped data: $Q_3 = l + \frac{h}{f}\left[\frac{3n}{4} - c\right]$

14. **If the value of Q_2, D_5 and P_{50} are equal to 72.32 then find the median of the distribution.**

Ans: Median = 72.32

15. **Define geometric mean with formula.**

Ans: The geometric mean of 'n' positive values is defined as the n^{th} root of the product.
If $X_1, X_2, X_3, \ldots\ldots\ldots, X_n$ are '**n**' values of a variable **"x"** and none of them being Zero then the geometric defined as:

$$G = \sqrt[n]{X_1.X_2.X_3.......X_n}$$

16. **If geometric mean of 3 items is 7, find the product of all items?**

Ans:

$$G.M. = \sqrt[n]{x_1.x_2.x_3.\ \ldots.x_n}$$

$$G.M. = \sqrt[3]{a.b.c}$$

$$7 = (a.b.c)^{1/3}$$

$$\boldsymbol{a.b.c = 343}$$

17. **Find harmonic mean of 5, 10, and 20.**

Ans: By definition;

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

X	**1/x**
5	0.2
10	0.1
20	0.05

Total	0.35

H.M. = 8.571 Ans.

18. Define weighted mean.

Ans: When all the values in the data are not of equal importance then we assign them certain numerical values to express their relative importance. These assigned values are called weights. The average of these weights with values is called weighted mean. The weights may be the quantities consumed or the numerical coefficient and are generally denoted by ω. The weighted mean denoted by "$\bar{y}_w$" of a set of 'n' values say

$y_1, y_2, ...y_n$ With weights $\omega_1, \omega_2, ...\omega_n$ is then given by:

$$\bar{y}_w = \frac{w_1 y_1 + w_1 y_1 + ... + w_n y_n}{w_1 + w_2 + ... + w_n}$$

$$\bar{y}_w = \frac{\Sigma w_i y_i}{\Sigma w_i}$$

Where; i = 1, 2, 3, 4 …n

19. Define harmonic mean.

Ans: Harmonic mean is defined as the reciprocal of the mean of the reciprocals of the items in a series. It is the ratio of the number of items and the sum of reciprocal of items.

$$H = \frac{\Sigma f}{\Sigma f \frac{1}{x}}$$

20. Calculate geometric mean of X = 1, 1, 27.

Ans: by definition;

$$G.M. = \sqrt[n]{x_{1.} x_{2.} x_{3.} \dots . x_n}$$

$$G.M. = \sqrt[3]{1.1.27}$$

$G.M. = 3$ **Ans.**

21. Write down properties of geometric mean.

Ans:

1. G.M is always less then A.M. i.e GM < A.M.
2. Geometric mean of constant variable is always constant.

22. Write down merits of geometric mean.

Ans:

1. It is rigidly defined by a mathematical formula.
2. It is based on all values.
3. It is less affected by extremely large values.

23. **Write down demerits of geometric mean.**

Ans:

1. It is not calculated if any of the observations is zero or negative.
2. In case of negative values, it cannot be computed at all.
3. It is not easy to understand.

24. **Write down the merits of A.M.**

Ans:

1. It is rigidly defined by mathematical formula.
2. It is easy to calculate.
3. It is easy to understand.
4. It is based upon all the values.
5. It is stable statistics in repeated sampling experiments.

25. **Write demerits of A.M.**

Ans:

1. It is greatly affected by extreme value.
2. It cannot be calculated for open-end classes without assuming open ends.
3. It gives fallacious and misleading conclusions when there is too much variation in data.

26. **Find A.M. if $\sum fx = 500$ and $\sum f = 50$.**

Ans: By definition;

$$\bar{X} = \frac{\sum fx}{\sum f}$$

$\bar{X} = 10$ **Ans.**

27. **If G.M. of two values is 3. Find the product of two values.**

Ans: by definition;

$$G.M. = \sqrt[n]{x_{1.}x_{2.}x_{3.} \dots . x_n}$$

$$G.M. = \sqrt[2]{a.b}$$

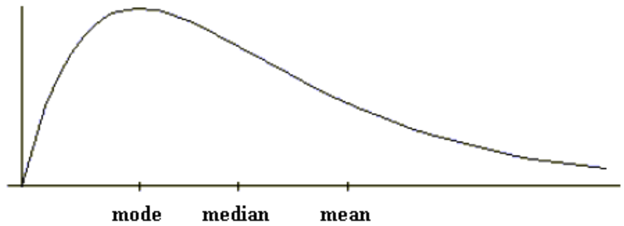
$$G.M. = (a.b)^{1/2}$$

$$\boldsymbol{a.b = 9}$$

28. **Illustrate the graphically positions of mean, median and mode for frequency curve which are skewed to the right and left.**

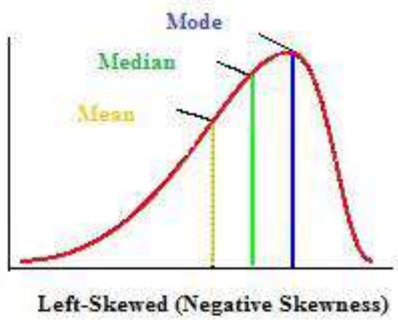
Ans.

a) For moderately positively skewed distributions, the following empirical relation holds.
Mean > Median > Mode



b) For moderately negatively skewed distributions, the following empirical relation holds.

Mean < Median < Mode



29. What are the properties of mode?

Ans:

1. If a constant a is added to each of the n observations $y_1 . y_2 \dots\ y_n$ having mode m, then the mode of $a + y_1, a + y_2 \dots .\ a + y_n$ would be $a + m$.
2. If a is multiplied with each of the n observations $y_1 . y_2 \dots\ y_n$ $h$aving mode m then the mode of $ay_1 . ay_2 \dots\ ay_n$ would be am.

30. Write down the merits of mode.

Ans:

1. It is easily located.
2. It is not affected by extreme values.
3. It can be located in case of open end classes.
4. It is simple to understand.

31. What are the demerits of mode?

Ans:

1. It is ill defined.
2. It is not based on all the values.
3. It is not capable of further algebraic treatment.
4. When the distribution has more than one mode then mode should not be calculated.

32. If $\overline{y}_1 = 3$ with $n_{1=}$ 3and $\overline{y}_2 = 4$ with $n_{2=}2$, then find $\overline{y}_c$.

Ans: By definition;

$$\bar{y}_c = \frac{n_1\overline{y_1} + n_2\overline{y_2}}{n_1 + n_2}$$

$\bar{y}_c = 17/5$ **Ans.**

33. Define average.

Ans: An Average is a single value which represents all values of data in some definite way.

For Example: The average income of middle class families is Rs.17000/per month.

34. For a certain frequency distribution, the mean was 40.5 and median was 36. Find mode by using of empirical relation.

Ans: Mode = 3Median – 2 mean

Mode = 3(36) – 2(40.5)

Mode = 27 **Ans.**

Chapter # 4: **Measures of Dispersion**

1. **Write any two advantages of the range.**

Ans:

1. It is easy to calculate.
2. It is useful measure in small samples.

2. **What are the demerits of range?**

Ans:

1. It is not based on all observations.
2. It depends only upon the extremes observations.

3. **Define relative dispersion.**

Ans: Relative measures of dispersion are calculated as ratios or percentages; for example, one relative measure of dispersion is the ratio of the standard deviation to the mean. Relative measures of dispersion are always dimensionless, and they are particularly useful for making comparisons between separate data sets or different experiments that might use different units. They are sometimes called coefficients of dispersion.

4. **What is quartile deviation?**

Ans: Half of the difference between the upper and the lower quartiles is called as quartile deviation or semi-inter quartile range. It is denoted by 'Q.D'

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

5. **Write down the merits of quartile deviation.**

Ans:

1. It is easy to calculate.
2. It is not affected by extreme values.

6. **Write down the demerits of quartile deviation.**

Ans:

1. If is not based on all the observations.
2. Q.D. will be the same values for all the distribution having the same quartiles.

7. **Define mean deviation**

Ans: It is defined as the mean of the absolute deviation of observations from mean, median or mode. By absolute deviations we mean that we consider all the deviations as positive. It is denoted by M.D. and calculated as:

$$M.D. = \frac{\sum |Y - M|}{n}$$ (Here, M is mean/median/mode)

8. Properties of mean deviation.

Ans:

1. M.D. from median is less than any other values i.e., $M.D. = \frac{\sum|Y - median\ |}{n}$ is least.
2. It is always greater than or equal to zero. $M.D. \geq 0$
3. For symmetrical distribution, the following relation holds; $M.D. = \frac{4}{5}\sigma$

9. Write down the merits of mean deviation.

Ans:

1. It is easy to calculate.
2. It is based on all the observations.

10. Write down the demerits of mean deviation.

Ans:

1. It is affected by the extreme values.
2. It is not readily capable of mathematical development.
3. It does not take into account the negative signs of the deviations from some average.

11. Write any two properties of standard deviation.

Ans:

1. The standard deviation of a constant is zero. $S.D.(a) = 0$
2. The standard deviation is independent of origin. $S.D.(y \pm a) = S.D.(y)$
3. When all the values are multiplied with a constant the standard deviation is multiplied by the constant i.e., $S.D.(ay) = |a|S.D.(y) \quad and \quad S.D.\left({}^{y}/_{a}\right) = |1/a|S.D.(y)$
4. The standard deviation of the sum or difference of two independent variables is the sum of their respective standard deviation for independent variables x and y.

$$S.D.(y \pm x) = S.D.(y) + S.D.(x)$$

12. Give two properties of variance.

Ans:

1. The variance of a constant is zero. Var$(a) = 0$
2. The variance is independent of origin. $Var\ (y \pm a) = Var(y)$
3. When all the values are multiplied with a constant the variance is multiplied by the square of the constant i.e., $var(ay) = a^2 var(y) \quad and \quad var\left({}^{y}/_{a}\right) = \frac{1}{a^2 var\ (y)}$

13. Given that mean = 156.17, median = 153.50 and standard deviation = 19.03. Calculate coefficient of skewness.

Ans: Karl pearson's second co-efficient of skewness:

$$= \frac{3(Mean - Median)}{S.D.}$$

$$= \frac{3(156.17 - 153.50)}{19.03}$$

$= 0.421$ **Ans.**

14. What is the use of coefficient of variation?

Ans: Co-efficient of variation is a relative measure of dispersion and independent of units of measurement and expressed in percentage. It is used to compare the variability of different sets of data. The group which has lower value of coefficient, coefficient of variation is comparatively more consistent.

15. What do you say about the relative dispersion of 5, 5, 5 and 5?

Ans: Relative dispersion would be zero, because the absolute dispersion of constant is zero.

16. If S^2 = 36 and $\overline{X}$ = 18, what is coefficient of variation?

Ans: By definition:

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

Solution;

$S^2 = 36$; $S = 6$ and $\overline{X} = 18$ put in above equation

$$\boldsymbol{C.V. = 33.333}$$

17. If variance of the value of '*X*' is 25, what is the standard deviation of *X*?

Ans: By taking square root of variance of x we get standard deviation of x. S.D.(x) = 5.

18. If S.D(X) = 10, then find the standard deviation of 5X?

Ans: By property of Standard deviation: $S.D.(ax) = |a|S.D.(x)$

$$S.D.(5x) = |5|S.D.(x)$$

$$S.D.(5x) = (5)(10)$$

$$\boldsymbol{S.D.(5x) = 50}$$

19. What is meant by symmetry?

Ans: In a symmetrical distribution, a deviation below the mean is equal to the corresponding deviation above the mean. This is called symmetry.

20. Define skewness.

Ans: Skewness is the lack of symmetry in a distribution around central (men, median or mode).

21. What are the types of measures of dispersion?

Ans: There are two types of measures of dispersion:

- Absolute Measure
- Relative Measure

22. Define the term variance.

Ans: Mean of squares of deviation of all the observations from their mean is called as variance. It is denoted by S^2.

$$S^2 = \frac{\sum(x-\bar{x})^2}{n}$$

OR

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$$

23. For a symmetrical distribution S.D. = 2. What is value of 4th moment about mean for mesokurtic data?

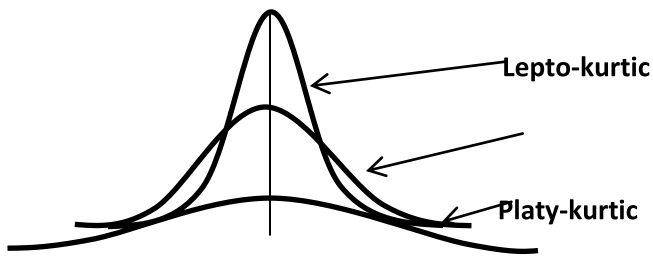
Ans: The dimensionless measure of kurtosis based on the moments is $\beta_2 = \frac{\mu_4}{\mu_2^2}$. If $\beta_2 = 3$, the distribution is mesokurtic (normal).

Variance = 4 by using S.D. = 2;

Forth moment about mean = $\mu_4 = 48$ **Ans.**

24. What do you know about kurtosis?

Ans: The word kurtosis is used to indicate the length of the tails and peakedness of symmetrical distributions. Symmetrical distribution may be platykurtic (more peaked), mesokurtic (normal) or leptokurtic (bit flat).



Meso-kurtic

25. If $b_2 = 3$ and $m_4 = 1875$, then what will be the standard deviation?

Ans: By using formula;

$\beta_2 = \frac{\mu_4}{\mu_2^2}$

$\mu_2 = 25$ (Variance)

S.D. = 5 **Ans.**

26. What is meant by absolute dispersion?

Ans: It can be defined as such a way that they have units (meters, grams) same as those of original measurements. There are following measures of absolute dispersion:

1. Range
2. Quartile Deviation
3. Mean Deviation
4. Variance
5. Standard Deviation

27. Write four measures of relative dispersion.

Ans: The relative measures of dispersion are given below:

1. Co-efficient of Range = $\frac{X_{max} - X_{min}}{X_{max} + X_{min}}$

2. Co-efficient of Quartile deviation $= \frac{Q_3-Q_1}{Q_3+Q_1}$
3. Mean Co-efficient of Dispersion $= \frac{M.D_{(\bar{X})}}{\bar{X}}$
4. Median Co-efficient of Dispersion $= \frac{\sum|x-\tilde{x}|}{n}$
5. Co-efficient of Standard Deviation $= \frac{S.D}{\bar{X}} \times 100$

28. If $s_x^2 = 5$, and y =2x, then what will be the value of variance of y?

Ans:

$$y = 2x$$

$$var(y) = var(2x)$$

$$var(y) = 2var(x)$$

$$\boldsymbol{var(y) = 10}$$

29. Write down the Bowley's and Karl Pearson's formula of coefficient of skewness.

Ans:

Bowley's coefficient of skeweness based on quartiles:

$$S_k = \frac{Q_3 + Q_1 - 2median}{Q_3 - Q_1}$$

Karl Pearson's second coefficient of skeweness:

$$S_k = \frac{3(Mean - Median)}{S.D}$$

Karl Pearson's first coefficient of skeweness:

$$S_k = \frac{Mean - Mode}{S.D}$$

30. Calculate lower quartile from the given data: 13, 3, 7, 15, 17, 5, 23

Ans: First step is to arrange the data into ascending order.

3, 5, 7, 13, 15, 17 and 23 ; n = 7

For ungrouped data: $Q_1 = \left(\frac{n+1}{4}\right)^{th}$ $value$

$Q_1 = 2^{nd}$ $value$

$Q_1 = 5$ **Ans.**

31. Compute coefficient of quartile deviation if $Q_1 = 10.20$ and $Q_3 = 58.29$

Ans: By using formula as we know;

Co-efficient of Quartile deviation $= \frac{Q_3-Q_1}{Q_3+Q_1}$

Co-efficient of Q.D. = 0.7022 **Ans.**

32. **Define standard deviation.**

Ans: The standard deviation is defined as the positive square root of the mean of the squares of the deviations of values from their mean. In other words, standard deviation is a positive square root of variance. It is denoted by 'S'.

$$S = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$$

33. **What is meant by dispersion?**

Ans: The degree to which numerical data tend to spread about an average value of the data is called as dispersion.

34. **Define measure of dispersion.**

Ans: A numerical quantity called measure of dispersion that describes the spread of the values in a set of data.

35. **Define range.**

Ans: The difference between the largest and the smallest observation is called as range. It is denoted by 'R'.

$$R = X_{max} - X_{min}$$

36. **What is the range of Bowley's coefficient of skewness?**

Ans: It lies between -1 to +1.

37. **If $u_2 = 4$ and $u_4 = 56$, find β_2.**

Ans: By using given formula;

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$\beta_2 = \frac{56}{16}$$

$$\boldsymbol{\beta_2 = 3.5}$$

38. **If var(X) = 4, then find var(3X).**

Ans: By using of variance property:

$$var(3x) = 9var(x)$$

$$var(y) = 9(4)$$

$$\boldsymbol{var(y) = 36}$$

39. **What are the measures of absolute dispersion?**

Ans: Some absolute measures of dispersion are:

1. Range
2. Quartile Deviation
3. Mean Deviation
4. Variance
5. Standard Deviation

40. If var(x) = 10 and y = 5x + 20, then find var(y).

Ans: By property of S.D. ;

$$y = 5x + 20$$

$$var(y) = var(5x + 20)$$

$$var(y) = 25var(x)$$

$$var(y) = 25(10)$$

$$\boldsymbol{var(y) = 250}$$

41. S.D. of a distribution is 4. Find second moment about mean.

Ans: As we know that second moment a bout mean is equals to variance then; variance = 16 (if S.D. = 4)

Chapter # 5: **Index Number**

1. **Define price index number.**

Ans: Price index number is a number that measures overall relative change in price of one or more commodities at a current period with respect to a standard base period. Consumer price index is an example of price index.

2. **Define index number.**

Ans: It is statistical tool which measures the relative change in a commodity or in a group of commodities with respect to time or locations

3. **Give any two uses of index numbers.**

Ans:

1- Index numbers are used for the comparison of prices.
2- Index numbers are used for forecasting.
3- Index numbers are used to measure the buying power of the money.

4. **Define base period.**

Ans: The period in which prices are compared with other period's price is called base period or reference period. Base period must be normal period without any irregular events like floods, strikes and wars etc.
There are two methods of selecting the base period:
1- Fixed base method
2- Chain base method

5. **Define link relative.**

Ans: Link relative is the percentage of ratio of the current year price and the preceding year price. In this case the base is not fixed. It is given by;

$$\textbf{Link Relative} = \frac{p_n}{p_{n-1}} \times 100$$

Where p_n is the pric in a given year and p_{n-1} is the price in the preceding year.

6. **Define composite index number**

Ans: **Composite Index numbers:** An index number is called a composite (aggregate) index number when it measures a relative change in two or more variables with respect to a base year.
For example index numbers for comparing two sets of prices from a wide variety of commodities, index numbers for comparing two sets of quantities from a wide variety of commodities.

7. **Define price relative.**

Ans: Price relative is the percentage of ratio of the current year price and the base year price.

8. **Define un-weighted index number.**

Ans: An index number that measures the change in prices of a group of commodities when the relative importance of commodities is not taken into account is called un-weighted index number.

9. **What is the relationship between Laspeyre's, Paasche's and Fisher's ideal index number?**

Ans: Fisher's ideal index number is the geometric mean of Laspeyre's and Paasche's index number.

10. **Given $\sum P_0$ =2550 and $\sum P_n$ = 2590. Find price index number using simple aggregative method.**

Ans: By formula;

Simple aggregate price index number:

$$p_{on} = \frac{\Sigma p_n}{\Sigma p_o} \times 100$$

p_{on} = 101.5686 **Ans.**

11. **Given $\sum p_0 q_n$ = 1000 and $\sum p_n q_n$ = 1360, find current year weighted index.**

Ans: By formula Paache's index number (Current year weighted index):

Paasche's Index Number:

$$p_{on} = \frac{\sum \mathrm{P}_n q_n}{\sum \mathrm{P}_o q_n} \times 100$$

p_{on} = 136 **Ans.**

12. **Given $\sum p_0 q_0$ = 850, $\sum p_n q_0$ = 1170. Find Laspayre's price index number.**

Ans: By formula Laspayer's index number (base year weighted index):

Laspeyre's index number (Aggregative Expenditure Method**):**

$$p_{on} = \frac{\sum \mathrm{P}_n q_0}{\sum \mathrm{P}_o q_0} \times 100$$

p_{on} = 137.65 **Ans.**

13. **If $\sum p_1 q_1$ = 480, $\sum p_0 q_1$ = 410, find current year weighted index number?**

Ans: By formula Paache's index number (Current year weighted index):

Paasche's Index Number:

$$p_{on} = \frac{\sum \mathrm{P}_1 q_1}{\sum \mathrm{P}_o q_1} \times 100$$

p_{on} = 117.07 **Ans.**

14. **If $\sum p_0 q_0$ = 322; $\sum p_1 q_0$ = 340; $\sum p_1 q_1$ = 362 and $\sum p_0 q_1$ = 326, find Fisher's price index number?**

Ans: By the given formula;

Fisher's Index Number:

$$p_{on} = \sqrt{Laspeyre \times Paasche}$$

$$P_{on} = \sqrt{\left(\frac{\sum P_1 q_o}{\sum P_o\, q_o} \times 100\right)\left(\frac{\sum P_1 q_1}{\sum P_o\, q_1} \times 100\right)}$$

p_{on} = 114.104 **Ans.**

15. **Which averages are used in index numbers? Name any two.**

Ans: Mean and Median

16. **Define consumer price index number.**

Ans: This index number measures the changes in the cost of living. By cost of living, we mean the cost of goods and services like food, rent, clothing, fuel and light, education, washing, etc. which are used and purchased by particular class of people.

17. **Differentiate between fixed base and chain base method.**

Ans:
Fixed base method: In fixed base method, the average price of a particular year or the average of the prices of a number of years is used as base.
Chain base method: In chain base method, the price of preceding year is taken as base.

18. **Given Laspayre's price index number = 120 and Paasche's price index number = 119.6, then find Fisher's index number.**

Ans: By formula;

Fisher's Index Number:

$$p_{on} = \sqrt{Laspeyre \times Paasche}$$

p_{on} = 119.78 **Ans.**

19. **If Paasche's index number is 105.72 and Laspeyre's index number is 107.22, then find Fisher's index number?**

Ans: By formula;

Fisher's Index Number:

$$p_{on} = \sqrt{Laspeyre \times Paasche}$$

p_{on} = 106.47 **Ans.**

20. **Given ∑w = 20, ∑wI = 1800. Find the cost of living index number by weighted average of relatives method.**

Ans: By using weighted index number;

$$P_{on} = \frac{\sum WI}{\sum W}$$

$P_{on} = 90$ **Ans.**

21. **Define simple and composite index numbers.**

Ans: **Simple Index Number:** Simple index number measures relative change in price or quantity or volume of one commodity at a current period or place with respect to base period.
For example: An index studying change in price of wheat over the last five years.

Composite Index numbers: An index number is called a composite (aggregate) index number when it measures a relative change in two or more variables with respect to a base year.
For example index numbers for comparing two sets of prices from a wide variety of commodities, index numbers for comparing two sets of quantities from a wide variety of commodities.

22. Define paasche's index numbers.

Ans: The index uses the current or given year quantities as weights. For this reason it is called current year weighted index. It is defined as mathematically:

$$P_{on} = \frac{\sum P_n q_n}{\sum P_o q_n} \times 100$$

23. Given $\sum p_0 = 660, \sum p_1 = 924$ and $\sum p_2 = 1056$, then compute simple aggregative price index number.

Ans: By using simple aggregative price index;

$$P_{o1} = \frac{\sum P_1}{\sum P_o} \times 100$$

$P_{o1} = 140$ **Ans.**

$$P_{o2} = \frac{\sum P_2}{\sum P_o} \times 100$$

$P_{o2} = 160$ **Ans.**

24. Given $\sum p_1 q_o = 1250$ and $\sum p_o q_o = 1200$, find base year weighted index number.

Ans: By using Laspayre's index number;

$$P_{o1} = \frac{\sum P_1 q_o}{\sum P_o q_o} \times 100$$

$P_{o1} = 104.167$ **Ans.**

25. Define weighted index number.

Ans: An index number that measures the change in prices of a group of commodities when the relative importance of commodities has been taken into account is called weighted index number.

26. Write down two advantages of chain base method.

Ans:

Advantages of chain base method:
1- Link relative are useful to make year to year comparison.
2- Changes in the geographical coverage can be accommodating.
3- New items can be substitute for old items provided the number of items remains the same.

27. Find C.P.I. if ∑w = 70, ∑wI = 800

Ans: By family budget:

$$P_{on} = \frac{\sum WI}{\sum W}$$

P_{on} =11.42857 **Ans.**

28. **Why fisher index number is called ideal?**

Ans: Fisher's ideal index satisfies both the time reversal and factor reversal tests.

29. **What are limitations of index numbers?**

Ans:

Limitations of index number:
1- It is not possible to take into account all changes in product.
2- There may be errors in the choice of base periods.
3- These are simply rough indications of the relative changes.

30. **Define whole price index.**

Ans: An index number that is designed to measure changes in the goods and services produced in different sectors of the economy and traded in wholesale markets is called the wholesale price index number.

31. **Write down some names of consumer items, considered in CPI?**

Ans:

a) Wheat
b) Rice Basmati
c) Sugar
d) Gram pulse
e) Cooling Oil
f) Potatoes
g) Bath soap
h) Electricity charges

32. **What is a market basket?**

Ans: The goods and the services are called the market basket. It consists of food, house rent, clothing, fuel and light, education and miscellaneous items.

Chapter # 6: **Probability**

1. **What is meant by equally likely events?**

Ans: when each outcome of a sample space is as likely to occur as any other, the outcomes are said to be equally likely. For example, if we toss a fair coin, the head is as likely to occur as the tail.

2. **What is meant by exhaustive events?**

Ans: outcomes are said to be exhaustive if they constitute the entire sample space. For example, if we toss a coin, the possible outcomes are a head and a tail. There is no other possibility, the coin will not stand on the edge.

3. **What is meant by permutation?**

Ans: An arrangement of all or some of a set of objects ina definite order is called permutation. Suppose we have different objects marked A, B, C, and D. with two objects A and B, the arrangements AB and BA are different permutation. With three objects A, B, and C, the permutations ABC, ACB, BCA, CBA and CAB are different permutations.

4. **Give classical definition of probability.**

Ans: Classical probability is the statistical concept that measures the likelihood (probability) of something happening. In a classic sense, it means that every statistical experiment will contain elements that are equally likely to happen (equal chances of occurrence of something).
(The ratio of the number of favorable outcomes to the number of all possible outcomes is called probability.)

$$P\ (A)\ = \frac{no.\ of\ favourable\ outcomes}{no. of\ possible\ outcomes} = \frac{m}{n}$$

5. **Define sample space.**

Ans: The set or collection of all possible outcomes of an experiment is called the sample space. It is denoted by "S". For example, when we toss a coin the sample space will be: S = {H, T} and for a rolling die: S = {1, 2, 3, 4, 5, 6}

6. **What is sample point?**

Ans: Each element of a sample space is called a sample point.

7. **Define composite event.**

Ans: If the event consists of more than one sample points, it is called a composite/compound event. For example, when two coins are tossed, the event A = {HH} that two heads appear is simple but the event B = {HH, HT, TH} that at least one head appears is a compound event.

8. **What are independent events?**

Ans: Two events A and B are said to be independent, if the occurrence of one event does not affect the occurrence of the other event.

9. **What is meant by dependent events?**

Ans: Two events A and B are said to be dependent, if the occurrence of one event affects the occurrence of other event.

10. **Define simple event.**

Ans: If an event consists of one sample points, it is called simple event.

11. What do you mean by not-mutually exclusive events?

Ans: Two events A and B are said to be not mutually exclusive, if they can occur together.

12. Explain mutually exclusive events.

Ans: Two events A and B are said to be mutually exclusive events, if they cannot occur together. Suppose we toss a coin and the head occurs, but the tail cannot. This type of events is known as mutually exclusive events.

13. If *P (A)* = 1/3, *P (AUB)* = ½ and *P (A∩B)* =1/10, find *P (B)*.

Ans: By addition law of not mutually exclusive events:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(B) = 0.933 \qquad \textbf{Ans.}$$

14. What is the probability of getting both sixes when two fair dice are thrown simultaneously?

Ans: Sample of two dice:

	1	2	3	4	5	6
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Let A be the event of getting both sixes.

$P(A) = 1/36$ **Ans.**

15. A fair die is rolled, what is the probability of getting an odd number?

Ans: Sample space of a rolling die: S = {1, 2, 3, 4, 5, 6}

Let A be the event of getting odd numbers: A = {1, 3, 5}

$$P(A) = \frac{m}{n}$$

$$P(A) = \frac{3}{6}$$

$$P(A) = 0.5 \qquad \textbf{Ans.}$$

16. What is a Venn diagram?

Ans: A simple and instructive way of representing the relationship between sets is by means of diagrams, called Venn Euler diagram or simply Venn diagram. In this diagram, the universal set U is represented by a rectangule, and subsets are represented by circles inside the rectangle.

17. For two mutually exclusive events A and B if P(A) = 0.25 and P(B) = 0.40, then find P(AUB).

Ans: P(AUB) = P(A) +P(B)

P(AUB) = 0.65 **Ans.**

18. Define event.

Ans: The possible outcome of an experiment is called an event. Thus an event is a subset of the sample space. Events are usually denoted by few capital letters A, B, C….

19. What will be the sample space if two coins are tossed?

Ans: S = {HH, HT, TH, TT}

20. A card is selected from 52 playing cards. What is probability that the card is king?

Ans: P(A) = 4/52 **Ans.**

21. If P(A) = 0.5 and P(B) = 0.2, find P(AUB) when 'A' and 'B' are mutually exclusive events.

Ans: P(AUB) = P(A) + P(B)

P(AUB) = 0.70 **Ans.**

22. What is the range of probability?

Ans: The range of probability is between 0 to 1.

23. Define the term combination.

Ans: When a selection of objects is made without paying regard to the order of selection, it is called combination."
It is denoted by ^nC_r.

$$^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Chapter – 7, 8

1. **What is a random variable?**

Ans: A variable whose values are determined by the outcomes of a random experiment is called random variable.

2. **Define continuous and discrete random variable.**

Ans:

Continuous Random Variable:

A random variable which takes on only a infinite number of values on a continuous scale in a given interval is called continuous variable. For example, the distance covered by a car between two points.

Discrete Random Variable: A random variable which takes on only a infinite number of values or a sequence of whole numbers is called a discrete random variable. For example, Number of people in a room, number of schools in a city Lahore.

3. **What is distribution function of a discrete random variable X?**

Ans: The distribution function of a X or the cumulative distribution function of X, denoted by F(x), is the probability that X will assume a value less than or equal to x, i.e.

$$F(x) = P(X \leq x)$$

4. **Define probability density function.**

Ans: The probability density function of a continuous random variable X is specified by a smooth curve such that the total area under the curve is one.

5. **What are the properties of discrete probability distribution?**

Ans: A discrete probability distribution has following two properties:

a) $0 \leq P(x_i) \leq 1$
b) $\sum P(x_i) = 1$

6. **What is meant by mathematical expectation of a random variable?**

Ans: If a discrete random variable *X* assumes the values 1, 2… *n* with respective probabilities *P*(1), *P*(2), … *P*(n) such that the sum of the probabilities is equal to 1, then the mathematical expectation or expected value of *X* denoted by *E(X)* is defined as, *E(X)* = $\Sigma xP(x)$. *E(X)* is also called mean of *X*, denoted by μ. Enlist two properties of expectation.

7. **Enlist properties of expectations.**

Ans: Properties of Expectation (Laws of Expectation)

1) $E(a) = a$
2) $E(aX + b) = aE(X) + b$
3) $E(XY) = E(X)E(Y)$
4) $E(X + Y) = E(X) + E(Y) \quad OR \quad E(X - Y) = E(X) - E(Y)$
5) $E[X - E(X)] = 0 \quad OR \quad E(X - \mu) = 0$

8. **Given *X* = 0, 1, 2 and *P(X)* = 9/16, 6/16, 1/16, find variance of *X*.**

Ans:

x	P(x)	x.P(x)	x^2.P(x)
0	0.56	0	0
1	0.38	0.375	0.375
2	0.06	0.125	0.25
Total	**1.00**	**0.50**	**0.63**

$$Mean = \mu = E(X) = \sum xP(x)$$

$$\boldsymbol{\mu = E(X) = 0.50}$$

$$Variance = \sigma_x^2 = Var(X) = \sum x^2\, P(x) - [E(x)]^2$$

$\boldsymbol{\sigma_x^2 = Var(X) = 0.38}$ **Ans.**

9. **Given the probability distribution. Find K.**
   X = 0, 1, 2, 3, 4 and *P(X)* = 1/210, 20/210, *K*, 70/210, 10/210.

Ans:

X	P(X)
0	1/210
1	20/210
2	K
3	70/210
4	10/210

By using the following property:

$$\sum P(x_i) = 1$$

1/210 + 20/210 + K +70/210 + 10/210 = 1

K = 109/210 **Ans.**

10. **Given that *f(x)* = *x*/10, x = 1, 2, 3, 4. Show that *f(x)* is a probability function.**

Ans: By using the Property:

$$\sum P(x_i) = 1$$

Putting the values of x into f(x);

1/10 + 2/10 + 3/10 + 4/10 = 1

10/10 = 1

1 = 1 Hence, proved that f(x) is a probability function.

11. **Given *f(x) = k/x²*, *x* = 1, 2, find *k*.**

Ans: putting the values of x into f(x) and then put equals to one;

$k/(1)^2 + k/(2)^2 = 1$

k/1 + k/4 = 1

5k/4 = 1

k = 4/5 **Ans.**

12. **Given *x* = 0, 2, 3 and *f(x)* = |1 – *X*|/4, find *E(X)*.**

Ans: Putting the values of x into f(x), we get…

x	f(x)	x.f(x)
0	1/4	0
2	1/4	1/2
3	2/4	3/2
Total		**2**

$$E(X) = \sum x.f(x)$$

$E(X) = 2$ **Ans.**

13. **Given *E(X)* = 0.63 and *Var(X)* = 0.2331 then find *E(X²)*.**

Ans:

$$Var(X) = E(X^2) - (E(X))^2$$

$E(X^2) = 0.63$ **Ans.**

14. **Given *X* = 1, 2, 3, 4, 5 and *P(X)* = 1/10, 3/10, *P*, 2/10, 1/10. Find the value of *P*.**

Ans:

X	P(X)
1	1/10
2	3/10
3	P
4	2/10
5	1/10

By using the following property:

$$\sum P(x_i) = 1$$

1/10 + 3/10 + P +2/10 + 1/10 = 1

P = 3/10 **Ans.**

15. **Find the probability distribution of the number of heads when two coins are tossed.**

Ans: Probability distribution of number of heads:

Sample space = {HH, HT, TH, TT}

X	f	P(x)
0	1	1/4
1	2	2/4
2	1	¼

16. **Define random experiment.**

Ans: An experiment in which outcomes vary from trial to trial is called random experiment.

17. **Enlist properties of probability mass function.**

Ans: Properties of Probability mass function (PMF):

a) $0 \leq P(x_i) \leq 1$
b) $\sum P(x_i) = 1$

18. **If E(X) = 1.4, then find E(5x – 4).**

Ans: By using the law of expectation:

$E(aX + b) = aE(X) + b$

$E(5x - 4) = 5E(x) - 4$

$E(5x - 4) = 3$ **Ans.**

19. **Given: E(x) = 0 and E(x^2) = 8/9. Find E($3x^2$ – 2x + 5).**

Ans: By using the laws of expectation:

$E(aX + b) = aE(X) + b$

$E(3x^2 - 2x + 5) = E(3x^2) - E(2x) + 5$

$E(3x^2 - 2x + 5) = 3E(x^2) - 2E(x) + 5$

$E(3x^2 - 2x + 5) = 7.667$ **Ans.**

20. **Given: E(x) = 0.56, var(x) = 1.36 and if y =2x + 1, then find E(y) and var(y).**

Ans: By using of Laws of expectation:

$Mean = E(y) = E(2x + 1)$

$E(y) = 2E(x) + 1$

$\boldsymbol{E(y) = 2.12}$

$Var(y) = Var(2x + 1)$

$Var(y) = Var(2x)$

$\boldsymbol{Var(y) = 5.44}$

Chapter # 9: **Binomial & Hypergeometric Distribution**

1. **Define Bernoulli trials.**

Ans: a Bernoulli trial two possible outcomes i.e., only success and failure is called Bernoulli trials. For each Bernoulli trial, the probability of success remains the same and the successive trials are independent.

2. **Define random experiment.**

Ans: An experiment in which outcomes vary from trial to trial is called random an experiment.

3. **Define Binomial Probability Distribution.**

Ans: if "P" is the probability of success in a single trial and "q" is the probability of failure, then the probabilityof exactly x successes in "n" trials of a binomial experiment is given by:

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

x = 0, 1, 2, 3, 4, …, n

4. **Define binomial experiment.**

Ans: An experiment in which the outcomes can be classified as success or failure and in which the probability of success remains constant from trial to trial is called Binomial experiment.

5. **What are the properties of binomial distribution?**

Ans: A binomial experiment possesses the following properties:

a) Each trial of experiment results in an outcome which can be classified into two categories i.e., success and failure.
b) The probability of success remains constant from one trial of the experiment to the next.
c) The repeated trials are independent.
d) The experiment is repeated a fixed number of times.

6. **A random variable X has a binomial distribution with *n* = 5 and *p* = 0.2, find *P(X = 2)*.**

Ans: By pdf of binomial distribution:

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$P(X = 2) = 0.0391 \quad \textbf{Ans.}$$

7. **Write down the formulas of mean, variance and standard deviation of binomial distribution.**

Ans:

$$\text{Mean} = np \quad , \quad \text{Variance } = npq \quad , \quad \text{Standard Deviation } = \sqrt{npq}$$

8. **A random variable *'X'* is binomially distributed when *n* = 15 and *p* = 0.4. Find mean and variance of *'X'*.**

Ans: By using formulas of mean and variance of binomial distribution:

$$\text{Mean} = np$$

$$\textbf{Mean} = \mathbf{6}$$

$$\text{Variance } = npq$$

$$q = 1 - p = 0.6$$

$$\mathbf{Variance} = \mathbf{3.6}$$

9. **In binomial distribution, mean = 6 and Variance = 2.4, find parameters of binomial distribution.**

Ans: As we know that;

$\text{Mean} = np$ and $\text{Variance} = npq$

$6 = np$ -------------------- 1

$2.4 = 6q$

$q = 0.4$ Then

$\boldsymbol{p = 0.6}$ (Put into 1)

$\boldsymbol{n = 10}$

Hence, 'n' and 'p' are the parameters of the binomial distribution.

10. **Find the number of trails of a binomial distribution which has mean = 12 and S.D = 2**

Ans: As we know that;

$\text{Mean} = np$ and $\text{Variance} = npq$

$12 = np$ (By using mean)-------------------- 1

$4 = 12q$ (By using variance)

$q = 0.333$ Then

$p = 0.666$ (Put into 1)

$\boldsymbol{n = 18}$

11. **A coin is tossed 5 times. What is the probability of getting exactly 3 heads?**

Ans: By using binomial probability distribution function;

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

Probability of head is our success; then,

$p = 0.5$, $q = 0.5$ and $n = 5$

$P(X = 3) = 0.3125$ **Ans.**

12. **What is Binomial frequency distribution.**

Ans: If the binomial probability distribution is multiplied by the number of experiments 'N' then the distribution is called binomial frequency distribution.

$$P(X = x) = N.\binom{n}{x} p^x q^{n-x}$$

13. **Define hyper geometric experiment.**

Ans: An experiment in which a random sample is selected without replacement from a finite population is called hyper geometric experiment.

14. **What are the properties of hyper geometric distribution?**

Ans: A hyper geometric experiment has the following properties:

a) Each trial of an experiment results in an outcome that can be classified into one of the two categories: success or failure.
b) The probability of success changes from one trial of the experiment to the next.
c) The repeated trials are dependent.
d) The experiment is repeated a fixed number of times.

15. **What is the Hypergeometric probability function?**

Ans: A population of 'N' items consists 'K' items of one kind (success) and N-K items are another kind (failure) and we are interested in the probability of getting 'x' success among 'n' items selected at random from the population.

$$P\ (X = x) = \binom{k}{x}\binom{N-k}{n-x}/\binom{k}{x}, \qquad x = 0, 1, 2, 3, 4, \ldots, n.$$

16. **What are the parameters of the hyper geometric distribution?**

Ans: n, k, and N.

17. What is mean and variance of hyper geometric distribution with parameters *N, n, K*?

Ans:

$$\text{Mean } = \frac{nk}{N} \quad , \quad \text{Variance } = \frac{nk}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right) \quad , \quad \text{Standard Deviation } = \sqrt{\frac{nk}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right)}$$

18. **Given *N* = 10, *n* = 4 and *K* = 5, find *E(X)*.**

Ans: As we know that expected value of x is known as the mean;

Then,

$$E(X) = \text{Mean } = \frac{nk}{N}$$

$E(X) = 2$ **Ans.**

19. **In a hyper geometric distribution *N* = 10, *n* = 2 and *k* = 3, find *P(X=0)***

Ans: By using Probability density function of Hypergeometric distribution:

$$P\ (X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$P\ (X = 0) = 0.467$ **Ans.**

20. **If *N* = 11, *n* = 5, *k* = 7, find variance of the hyper geometric distribution?**

Ans: The variance of hypergeometric distribution is :

$$\text{Variance } = \frac{nk}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right)$$

$$\text{Variance } = 0.6942$$ **Ans.**

21. If 'x' is a binomial random variable with n = 9 and p = 1/3, then find S.D.(3+2x)?

Ans: As we know that:

$$\text{Standard Deviation } = \sqrt{npq}$$

$$\text{Standard Deviation} = \text{S. D. (x) } = 1.4142$$

$$\text{S. D.}\,(3 + 2\text{x}) = 2\text{S. D.}\,(\text{x})$$

$$\text{S. D.}\,(3 + 2\text{x}) = 2.8284$$

22. In a binomial distribution mean = 36 and q = 0.83, find 'n' and 'p'?

Ans: As we know that;

$\text{Mean} = np$ and $\text{Variance } = npq$

$p = 1 - q$

$\boldsymbol{p = 0.17}$

$36 = n(0.17)$ (By using mean)-------------------- 1

$\boldsymbol{n = 212}$