IMPORTANT CONCEPTS Part-I

1. **The word Statistiics :** is used in three meanings:-

   i) The aggregate of numerical facts in any field of enquiry. Word 'data' is meant in this sense.

   ii) The Science of systamatic collection, presentation, analysis and interpretation of data. In other words the "subject of statistics" is meant by it.

   iii) It is used as the plural of the word 'Statistic' a term used to define any value extracted from the sample, a part of the population.

2. **Primary data :** are the most original data which are collected for the first time and have not undugone any statistical treatment.

3. **Secondary data :** are the data which have undergone at least once any statistical treatment.

4. **Grouped data :** are data in arranged form in classes or groups along with their corresponding frequencies.

5. **Variable :** is any quantitative characteristic which varies from one individual or object to another.

6. **Constant :** is any value which does not change but remains fixed.

7. **Discrete Variable :** is a variable which takes countable or countably infinite number of values.

8. **An Array :** is an arrangement of data in ascending or descending order.

9. **Classification :** is the process of arranging data in classes according to their similarities.

10. **Tabulation :** is the process of arranging data in rows and columns.

11. **Freq. Distribution :** is the tabular arrangement of data in to groups according to size or magnitude.

12. **Class Limits :** are the values defining the classes or groups the smaller value is called lower class limit and the greater one is called upper class limit.

13. **Class-boundaries (C.Bs) :** are precise or true class limits. These values separate the class limits. The smaller value is called lower class boundary and the greater one is called upper class boundary.

14. **Class mark/Mid point/Middle Value :** is the mean of lower and upper class limits or boundaries.

15. **Size of Class Interval :** is the difference between the upper and lower class boundaries of a class.

16. **Frequency :** is the no. of individuals or objects in a class.

17. **Cummulative freq. :** is the total frequency of all less then the upper C.B. of a given class.

18. **Simple Bar Diagram :** is a chart consisting of vertical or horizontal bars of equal width with heights equal to the corresponding frequencies.

19. **Multiple Bar Diagram :** is a grouped bar diagram used to the represent two or more sets of inter-related data.

20. **Pie Chart :** is the sub-division of a circle into sectors whose areas are proportional to the different angles of the total area of a circle i.e. 360^{O}.

21. **Historigram :** is the graphic representation of a time-series.

22. **Histogram :** is the graphic representation of freq. dist. by a set of adjacent rectangles in which area of each rectangle is proportional to the corresponding class frequencies.

23. **Frequency Polygon :** (A closed shape with many corners) is the graphic form of a frrequency dist. obtained by joining the mid-points of the types of each rectangle of the histogram and the end points are joined to the mid-points of arbitrary classes taken at the start and at the end on x-axis.

24. **Cummulative frequency curve or ogive :** is a graphic representation obtained by plotting cummulative frequencies against the upper C.B. and by joining the points smoothly.

25. **Average :** is a balancing point of a set of data or it is a central value representing the entire data.

26. **Measures of Central Tendency or Measures of Location :** are measures which tend to lie in centre or measures used to locate the centre of the data.

27. **Arithmatic Mean (A.M.) :** is a value obtained by dividing the sum of all the values by their total number.

28. **G.M. :** is the ηth positive root of the product of non-zero and non-negative values.

29. **H.M. :** is the raciprocal of the A.M.. of raciprocals of the non-zero values.

30. **Mode :** is a value which occurs most of the times the data.

31. **Median :** is the most middle value of the data.

32. **<u>When an average should be used</u> :**

A.M. should be used when data are relatively homogeneous.

G.M. should be used when data are ratios of change.

H.M. should be used when data are rates of change.

Mode should be used when data have a predominated figure.

Median should be used when data are hetrogeneous.

33. **Quartiles :** are the values that divide the data into four equal parts.

34. **Deciles :** are the values that divide the data into ten equal parts.

35. **Percentiles :** are the values that divide the data into hundred equal parts.

36. **Dispersion :** is the scatterness of values about some central value of the data.

37. **Measure of dispersion :** is a value that measures the dispersion or scatterness of data about some central value.

38. **Absolute measure of dispersion :** is measure of disperson expressed in the same units as of the data.

39. **Relative measure of disperson :** is measure of dispersion expressed in the form of co-efficient and is unitless.

40. **Range :** is the difference between maximum and minimum values of the data.

41. **Mean Deviation :** is the mean of the absolute deviations from any measure of central tendency.

42. **Variance :** is the mean of squared deviations of values from their arithmatic mean.

43. **Co-efficient of variation (C.V.) :** is the percentage of S.D. with respect to mean.

44. **Co-efficient of S.D. :** is the ratio of S.D to mean.

45. **Moments :** are the means of rth powered-deviations.

46. **Skewness :** is the lack of symmetry.

47. **Kurtosis :** is the peakedness or flatness of the uni modal curve.

48. **Index Nos. :** are the economic barometers to know relative change in a variable or a group of related variables with respect to time or space.

49. **Simple I/No. :** is a number that measures relative change in a single variable with respect to time or space.

50. Composite I/No. : is a number that measures a ralative change in a group of related variables with respect to time or space.

51. The main steps used in the construction of I/No. are : (I) Purpose & Scope (ii) Selection of Commodities (iii) Collection of prices (iv) Selection of base period (v) Choice of average (vi) Selection of appropriate weights.

52. Fixed base method : is a method in which a particular economically stable year or average of some years is taken as base.

53. Chain base Method : is a method in which a particular year is not fixed but goes on changing with the given year i.e. the relatives are computed with the immediately preceding year as a base.

54. η Factorial : is product of first n natural numbers.

55. Permutaion : is an arrangement of all or some of a set of objects in a definite order.

56. Combination : is an arrangement of objects without regard to their order.

57. Experiment : is any process of obtaining an observation.

58. Random Experiment : is an experiment which produces different outcomes even thought it is repeated under the similar conditions.

59. Trial : is the single performance of an experiment.

60. Outcomes : are the results obtained from an experiment.

61. Sample Space : is a complete list of all possible outcomes obtained from random experiment.

62. Sample Points : are the elements of a sample Space.

63. Event : is a subset of a Sample Space.

64. Simple Event : is the event that consists of only one sample point of the Sample Space.

65. Composite Event : is the event that consists of more than one sample point of the Sample Space.

66. Equaly likely events : are the events that have the same no. of sample points of the sample space.

67. Mutually Exclusive Events : are the events that have no common sample points of the sample space.

68. Independent Event : are those events which do not effect each other with respect to their occurance.

69. **Dependent Event :** are those events which effect each other with respect to their occurance.

70. **Exhaustive Events :** are the disjoint events that constitute the entire sample space.

71. **The Classical OR "a priori" definition of probability :** "If a random experiment can be performed in 'n' mutually exclusive, equally likely, exhaustive number of sample points and "m" of these are favourable to the occurance of a certain event A, then the probability of the occurance of the event A (P(A)) is the ratio m/n. i.e. $P(A) = m/n$

72. **Relative frequency OR "a posteriori" OR Statistical OR empirical definition of probability :** If a random experiment is repeated a large number of times, say, 'n' under the similar condition, if an event A is observed to occur 'm' times, then the probability of occurance of event A P(A) is defined as the limit of the relative frequency m/n as tends to infinity. i.e. $P(A) = \lim_{n \to \infty} \frac{m}{n}$

73. **Axiomatic definition of probability :** Let the Sample Sapce S has the Sample Points E_1, E_2, E_3 .....E_i, ... E_n . Each Sample point is assigned a real number $P(E_1), P(E_2), P(E_3)$ .... $P(E_i)$ $P(E_n)$ which are called the probabilities of the points. The probability of any point denoted by P(Ei) must obey the following the axioms.

i) $o < P(Ei) < 1$: $(i = 1,2,3 \ldots n)$

ii) $P(S) = 1$

iii) For the mutually exclusive events E_1 and E_2, $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

74. **Random Numbers :** are the nos. obtained by some random process (manual or mechanical) from a discrete uniform distribution of basic numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

75. **Random Variable (r.v) :** is the set of numerical values obtained from the outcomes of a random experiment.

76. **Discrete r.v. :** are the variables that can assume funite or countably infinite values in a sample space.

77. **Continuous r.v. :** are the variables that can assume value in an interval [a,b], where a<b.

78. **Probability Distribution OR Probability function :** is an arrangement of values of r.v. along with their corresponding probabilities in the form of pairs, table, graph or formula.

79. **Discrete Probability Distribution is defined as :** If X is a discrete r.v. assuming the values $x_1, x_2, x_3, \ldots x_i \ldots x_n$, with their corresponding probabilities $P(X = x_i) = f(x_i) = P(x_i)$ $i = 1,2,3, \ldots n$ is called discrete probability distribution

80. **Continuous probability distribution function (p.d.f) is defined as :** If X is a continuous r.v. defined in an interval [a,b] then the p.d.f of X is defined as $P(a \le x \le b) = \int_a^b f(x)\,dx$

81. **Binomial Experiment:** if in an experiment

i) Each trial results either in success or failure.

ii) The trials are independent.

iii) The probability of success remains constant from trial to trial, and

iv) The experiment is repeated a fixed number of times.

82. **Binominal probability distribution:** if p is the probability of success of an event in a single trial then the probability of exactly x success in 'n' independent trials of a 'Binominal Experiment' is given by

$$f(x) = \binom{n}{x} p^x q^{n-x} \qquad \text{for } x = 0,1,2 \ldots\ldots,n$$

83. **Binominal random variable:** is a random variable X that follows binominal probability distribution.

84. **Binominal Frequency distribution:** if N, the number of sets or experiments multiplies basic binominal probability function, the result is Binominal Frequency distribution.

85. **Hypergeometric Experiment:** if in an experiment

i) Each trial results either in success or failure.

ii) The successive trials are dependent.

iii) The probability of success changes from trial to trial, and

iv) The experiment is repeated a fixed number of times.

86. **Hypergeometric probability distribution:**

If 'N' is number of elements in a population, from which 'K' are success and a sample of size 'n' is selected from the population then the probability of 'x' successes at X= x given by

$$f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \qquad \begin{array}{l} \text{for } x = 0,1,2 \ldots\ldots,n \\ N \geqslant n \\ K \geqslant x \\ n \geqslant x \end{array}$$

87. **Hypergeometric random variable:** is a random 'x' that follows the hypergeometric probability distribution:

IMPORTANT FORMULAE Part-I

1. MEASURES OF CENTRAL TENDENCY

(a) Arithmatic Mean (Mean) ($\overline{X}$)

	Ungrouped data	Grouped data
Direct Method	$\overline{X} = \frac{\Sigma X}{n}$	$\overline{x} = \frac{\Sigma fX}{\Sigma f}$
Short cut Method	$\overline{X} = a + \frac{\Sigma D}{n}$	$\overline{x} = a + \frac{\Sigma fD}{\Sigma f}$
Coding Method	___	$\overline{x} = a + \frac{\Sigma fU}{\Sigma f} \times h$

Where 'a' an assumed value (Provisional Mean)

'D' Deviation of 'x' from the value 'a' i.e. $D = x - a$

'h' Class Interval Size (Uniform)

$U = \frac{X - a}{h}$ or $U = \frac{D}{h}$

Sum of deviationsfrom the A.M. is always Zero i.e. $\Sigma(X - \overline{X}) = 0$ and $\Sigma f(X - \overline{X}) = 0$

(b) Weighted Mean ($\overline{X}w$)

Here use weights instead of frequencies, $\overline{X}w = \frac{\Sigma XW}{\Sigma W}$

(c) Combined Mean ($\overline{Xc}$) /3

If $\overline{X}_1, \overline{X}_2, \ldots \overline{X}_k$, are the Means of the numbers in each

k - the group $n_1, n_2, \ldots n_k$ then $\overline{X}c = \frac{\Sigma n_i \overline{X}_i}{\Sigma n_i}$

(d) Geometric Mean (G.M.)

Ungrouped data	Grouped data
$G.M. = (X_1 \times X_2 \ldots X_n)^{1/n}$	___
$G.M. = \text{Antilog}\left[\frac{\Sigma \log x}{n}\right]$	$G.M. = \text{Antilog}\left[\frac{\Sigma f \log x}{\Sigma f}\right]$

(e) Harmonic Mean (H.M.)

$H.M. = \frac{n}{\Sigma(1/x)}$ $H.M. = \frac{\Sigma f}{\Sigma(f/x)}$ or $H.M. = \frac{\Sigma f}{\Sigma f(1/x)}$

Relation between (AM, GM, and HM)

(i) GM, $-\sqrt{(A.M.)\,(H.M.)}$ (ii) A.M. $\geqslant$ G.M $\geqslant$ H.M

(e) Median :

Median = The value of $\frac{n+1}{2}$th item of (arranged data) (ungrouped data)

Median $= L + \frac{h}{f}\left(\frac{N}{2} - c\right)$ (Grouped data)

Where $N = \Sigma f_i$ $\frac{N}{2}$ = Median group (M.G.)

'L' = the lower boundary of the M.G.

'f' = the frequency of the M.G.

'c' = the preceding cumulative frequency of the M.G.

'h' = the class interval size of the M.G.

Quartiles : (Q)

Q_i = The value of $i\left(\frac{n+1}{4}\right)$ th item (Ungrouped data)

$Q_i = 1 + \frac{h}{f}\left(\frac{iN}{4} - c\right)$ (Grouped data)

Where $i = 1,2,3$

Deciles : (D)

D_J = The value of $J\left(\frac{n+1}{10}\right)$ th item (Ungrouped data)

$D_J = 1 + \frac{h}{f}\left(\frac{JN}{10} - c\right)$ (Grouped data)

Where $J = 1, 2, 3, \ldots\ldots, 9$

Percentiles : (P)

P_k = The value of $k\left(\frac{n+1}{100}\right)$ th item (Ungrouped data)

$P_k = L + \frac{h}{f}\left(\frac{KN}{100} - c\right)$ (Grouped data)

Where $K = 1, 2, 3, \ldots\ldots, 99$

Note: Median $= Q_2 = D_5 = P_{50}$ and $D_1 = P_{10}$ $D_2 = P_{20}$ ..... $D = P$

Also $Q_1 = P_{25}$ and $Q_3 = P_{75}$

Mode :

$$\text{Mode} = L + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$\text{Mode} = x + \frac{f_2 - f_1}{(f_m - f_1) + (f_m - f_2)} \times \frac{h}{2}$$

Mode = 3 Median - 2 Mean (approximately) (Empirical Relation)

Where,

f_m = the frequency of the Model Class or (group) (M.G.)

f_1 = the preceding frequency of the (M.G.)

f_2 = the proceding frequency of the (M.G.)

L = the lower boundary of the M.G.

h = the class interval size of M.G

x = the mid point of M.G.

If, Mean = Median = Mode, the Distribution will be symmetrical

If, Mean > Median > Mode, the Distribution will positively skewed

If, Mean < Median < Mode, the Distribution will negatively skewed

(2) MEASURES OF DISPERSION

(A) ABSOLUTE MEASURES OF DISPERSION :

(a) Range: (R)

$R = X_m - X_o$ Where X_m = Maximum Value, X_o = Minimum Value

(b) Interquartile Range (I.Q.R.)

$I.Q.R. = Q_3 - Q_1$ Where Q_3 = Upper quartile, Q_1 = Lower quartile

(c) Semi-Interquartile Range (S.I.Q.R.) OR Quartile Deviation (Q.D)

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

(d) Mean Deviation (Average Deviation) (M.D)

	Ungrouped data	Grouped data
M.D. (Mean) =	$\frac{\Sigma\lvert X - \overline{X}\rvert}{n}$	$M.D. = \frac{\Sigma f\lvert X - \overline{X}\rvert}{\Sigma f}$
M.D. (Median) =	$\frac{\Sigma\lvert X - \text{Median}\rvert}{n}$	$M.D. = \frac{\Sigma f\lvert X - \text{Median}\rvert}{\Sigma f}$
M.D. (Mode) =	$\frac{\Sigma\lvert X - \text{Mode}\rvert}{n}$	$M.D. = \frac{\Sigma f\lvert X - \text{Mode}\rvert}{\Sigma f}$

(e) Standard deviation (S) and variance (S^2)

	Ungrouped data	Grouped data
Direct Method	$S = \sqrt{\frac{\Sigma(X - \overline{X})^2}{n}}$	$S = \sqrt{\frac{\Sigma f(X - \overline{X})^2}{\Sigma f}}$
Indirect Method	$S = \sqrt{\frac{\Sigma X^2}{n} - \left(\frac{\Sigma X}{n}\right)^2}$	$S = \sqrt{\frac{\Sigma fX^2}{\Sigma f} - \left(\frac{\Sigma fX}{\Sigma f}\right)^2}$
Shortcut Method	$S = \sqrt{\frac{\Sigma D^2}{n} - \left(\frac{\Sigma D}{n}\right)^2}$	$S = \sqrt{\frac{\Sigma fD^2}{\Sigma f} - \left(\frac{\Sigma fD}{\Sigma f}\right)^2}$
Coding Method	——	$S = \sqrt{\frac{\Sigma fU^2}{\Sigma f} - \left(\frac{\Sigma fU}{\Sigma f}\right)^2} \times h$

(B) RELATIVE MESURES OF DISPERSION:

(i) Co-efficient of Dispersion $= \frac{X_m - X_o}{X_m + X_o}$

(i) Co-efficient of Q.D. $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$

(iii) Mean Co-efficient of Dispersion $= \frac{M.D.}{\text{Mean}}$

Median Co-efficient of Dispersion = $\frac{\text{M.D.}}{\text{Median}}$

Mode Co-efficient of Dispersion = $\frac{\text{M.D.}}{\text{Mode}}$

(iv) Co-efficient of S.D. = $\frac{\text{S.D.}}{\text{Mean}}$

(v) Co-efficient of variation (C.V.) = $\frac{\text{S.D.}}{\text{Mean}} \times 100$

(f) Combined Variance

If $n_1, n_2, \ldots n_k$, are the Nos. of k-th groups having Means $\overline{X}_1, \overline{X}_2, \ldots \overline{X}_k$, and variances, $S_1^2, S_2^2, \ldots, S_k^2$ then Combined

$$\text{Mean } \overline{X}c = \frac{n_1\overline{X}_1 + n_2\overline{X}_2 + \ldots + n_k\overline{X}_k}{n_1 + n_2 + \ldots + n_k} = \frac{\Sigma n_i \overline{X}_i}{\Sigma n_i}$$

and Combined Variance

$$S^2_{(comb)} = \frac{1}{n}\left[\sum_{i=1}^{k} ni\{Si^2 + (\overline{X}i - \overline{X}c)^2\}\right] \quad n = \Sigma n_i$$

Co-efficient of variation (comb) = $\frac{S_{(comb)}}{\overline{X}c} \times 100$

(g) Corrected variance

$$S^2_{(corrected)} = S^2 - \frac{h^2}{12}$$

Co-efficient of variation (corrected) = $\frac{S_{(corrected)}}{\overline{X}} \times 100$

(C) PROPERTIES OF VARIANCE

If a and b are any two constants, and x, y, be Independent variables, then

(i) Var (a) = 0

(ii) Var (x ± a) = Var (x)

(iii) Var (ax) = a^2 var (x)

(iv) Var (ax ± b) = a^2 var (x)

(v) Var (x ± y) = Var (x) + Var (y)

(3) MOMENTS

i) Moments about Arithmatic Mean:

Ungrouped data

$$m_1 = \frac{\Sigma(X - \overline{X})}{n} = o$$

$$m_2 = \frac{\Sigma(X - \overline{X})^2}{n} = S^2$$

$$m_3 = \frac{\Sigma(X - \overline{X})^3}{n}$$

$$m_4 = \frac{\Sigma(X - \overline{X})^4}{n}$$

Grouped data

$$m_1 = \frac{\Sigma f(X - \overline{X})}{\Sigma f} = o$$

$$m_2 = \frac{\Sigma f(X - \overline{X})^2}{\Sigma f} = S^2$$

$$m_3 = \frac{\Sigma f(X - \overline{X})^3}{\Sigma f}$$

$$m_4 = \frac{\Sigma f(X - \overline{X})^4}{\Sigma f}$$

ii) Moments about origin (x=a) P.M.

$$m_1' = \frac{\Sigma(x - a)}{n} = \frac{\Sigma D}{n} \qquad m_1' = \frac{\Sigma f(x - a)}{\Sigma f} = \frac{\Sigma fD}{\Sigma f}$$

$$m_2' = \frac{\Sigma D^2}{n} \qquad m_2' = \frac{\Sigma fD^2}{\Sigma f}$$

$$m_3' = \frac{\Sigma D^3}{n} \qquad m_3' = \frac{\Sigma fD^3}{\Sigma f}$$

$$m_4' = \frac{\Sigma D^4}{n} \qquad m_4' = \frac{\Sigma fD^4}{\Sigma f}$$

iii) Moments about Zero (x = o)

$$m_1' = \frac{\Sigma X}{n} = \overline{X} \qquad m_1' = \frac{\Sigma fX}{\Sigma f} = \overline{X}$$

$$m_2' = \frac{\Sigma X^2}{n} \qquad m_2' = \frac{\Sigma fX^2}{\Sigma f}$$

$$m_3' = \frac{\Sigma X^3}{n} \qquad m_3' = \frac{\Sigma fX^3}{\Sigma f}$$

$$m_4' = \frac{\Sigma X^4}{n} \qquad m_4' = \frac{\Sigma fX^4}{\Sigma f}$$

iv) Moments about Coding Method : $U = \frac{(x - a)}{h}$

$$\left.\begin{array}{ll} m_1' = \dfrac{\Sigma fU}{\Sigma f} \times h & m_2' = \dfrac{\Sigma fU^2}{\Sigma f} \times h^2 \\ m_3' = \dfrac{\Sigma fU^3}{\Sigma f} \times h^3 & m_4' = \dfrac{\Sigma fU^4}{\Sigma f} \times h^4 \end{array}\right] \text{Grouped data}$$

v) Moments about A.M. in terms of P.M.

$$m_1 = o$$

$$m_2 = m_2' - (m_1')^2 \longrightarrow \text{variance}$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$m_4 = m_4' - 4m_1'm_3' + 6m_1'^2m_2' - 3(m_1')^4$$

vi) Moment Ratios : β_1 β_2 or b_1, b_2

$$\beta_1 = b_1 = \frac{(m_3)^2}{(m_2)^3} \qquad \beta_2 = b_2 = \frac{m_4}{(m_2)^2}$$

If $\beta_1 = o$, the distribution will be symmetrical
other wise positively or negatively skewed

If $\beta_2 = 3$, the distribution will be Normal (Mesokurtic)

If $\beta_2 > 3$, the distribution will be Leptokurtic

If $\beta_2 < 3$, the distribution will be Playtykurtic

vii) Co-efficient of skewness : (Sk)

Karl Pearson Co-efficient of Skewness

$$Sk = \frac{\text{Mean - Mode}}{\text{S.D.}} \qquad \text{and} \qquad Sk = \frac{3(\text{Mean - Median})}{\text{S.D.}}$$

Bowleys (Quartiles) Co-efficient of Skewness

$$Sk = \frac{Q_3 + Q_1 - 2\,\text{Median}}{Q_3 - Q_1}$$

If $Sk = 0$, the distribution will be symmetrical

If Sk or $\sqrt{\beta_1} > o$, the distribution will be Positively Skewed

If Sk or $\sqrt{\beta_1} < o$, the distribution will be Negatively Skewed

viii) Sheppareds Correction : (Corrected moments)

$$\overline{m}_2 = m_2 - \frac{h^2}{12} \qquad \text{(Corrected Variance)}$$

$$\overline{m}_4 = m_4 - \frac{h^2}{2}\, m_2 + \frac{7}{240} \times h^4$$

ix) Relationship between different Measues of dispersion

$$\text{Q.D.} = \frac{2}{3}\,\text{(S.D)} \quad \text{app.} \qquad \text{M.D.} = \frac{4}{5}\,\text{(S.D)} \quad \text{app.}$$

$$\text{Q.D.} = \frac{5}{6}\,\text{(S.D)} \quad \text{app.}$$

(4) INDEX NUMBER

Price Relative $P_{on} = \frac{Pn}{Po} \times 100$

Link Relative $P(n-1, n) = \frac{Pn}{Pn - 1} \times 100$

Where P_n = Price for the current (given) year

P_o = Price for the base year

q_n = Quantity for the current (given) year

q_o = Quantity for the base year

a) **Unweighted Index Numbers :** **(Two Methods)**

i) Simple aggregate Method, $Pon = \frac{\Sigma Pn}{\Sigma Po} \times 100$

ii) Simple avergae of Relative Method

$Pon = \frac{1}{k} \Sigma \left[\left(\frac{Pn}{Po} \right) \times 100 \right]$ (A.M. as average)

$Pon = \left[\Pi \left(\frac{Pn}{Po} \right) \right]^{\frac{1}{k}} \times 100$ (G.M. as average)

Where Π = product sign K = No, of Commodities

Median as usual, the middle most relative for each year is taken as Median

a) **Weighted Index Numbers :** **(Two Methods)**

i) Weighted aggregative Index

ii) Weighted average of Relative Index

i) **Weighted aggregate Price Index Numbers :**

Computed by the formulae

Laspeyre's Index $Pon = \frac{\Sigma Pnqo}{\Sigma Poqo} \times 100$

Paasche's Index $Pon = \frac{\Sigma Pnqn}{\Sigma Poqn} \times 100$

Fisher Ideal Index $Pon = \sqrt{\frac{\Sigma Pnqo}{\Sigma Poqo} \times \frac{\Sigma Pnqn}{\Sigma Poqn}} \times 100$

OR

$Pon = \sqrt{\text{Lasp (Pon)} \times \text{Paash (Pon)}}$

Marshal Edgeworth Index $\quad Pon = \dfrac{\Sigma Pn\,(qo + qn)}{\Sigma Po\,(qo + qn)} \times 100$

OR $\quad Pon = \dfrac{\Sigma Pnqo + \Sigma Pnqn}{\Sigma Poqo + \Sigma Poqn} \times 100$

i) Weighted aggregate Quantity Index Numbers :

Laspeyre's Index $\quad Qon = \dfrac{\Sigma qnPo}{\Sigma qoPo} \times 100$

Paasche's Index $\quad Qon = \dfrac{\Sigma qnPn}{\Sigma qoPn} \times 100$

Fisher Ideal Index $\quad Qon = \sqrt{\dfrac{\Sigma qnPo}{\Sigma qoPo} \times \dfrac{\Sigma qnPo}{\Sigma qoPn}} \times 100$

OR

$$Qon = \sqrt{\text{Lasp }(Qon) \times \text{ Paash }(Qon)}$$

Marshal Edgeworth Index $\quad Qon = \left[\dfrac{\Sigma qn\,(Po + Pn)}{\Sigma qo\,(Po + Pn)}\right] \times 100$

$$Qon = \frac{\Sigma qnPo + \Sigma qnPn)}{\Sigma qoPo + \Sigma qoPn} \times 100$$

i) Weighted average of Relative Price Index :

$$Pon = \frac{\Sigma(\frac{Pn}{Po}) \times Poqo}{\Sigma Poqo} \times 100$$

OR

$$Pon = \frac{\Sigma(\frac{Pn}{Po}) \times w}{\Sigma w} \times 100$$

OR

$Pon = \dfrac{\Sigma IW}{\Sigma W}$ $\qquad$ Where $I = \dfrac{Pn}{Po} \times 100$, $\qquad W = Poqo$

Consumer's Price Index Numbers : (CPI)

Aggregative Expenditure Method. $\quad Pon = \dfrac{\Sigma Pnqo}{\Sigma Poqo} \times 100$

Family (Household) Budget Method. $\quad Pon = \dfrac{\Sigma IW}{\Sigma W} \times 100$

(5) PROBABILITY

Factorial: $n! = n(n-1)(n-2)....3.2.1 \quad (0! = 1)$

eg $6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$

Permutation: (P) $P^n_r = \dfrac{n!}{(n-r)!}$

$$P = \frac{n!}{n_1!\, n_2! \, ... n_k!} \qquad \sum_{i=1} nk = n$$

When all objects taken at a time having different types or kinds

Combination $c^n_r = \dfrac{n!}{(n-r)!\, r!}$

Laws of Probability :

If A and B are any two events, then

i) Addition Theorem : (Law)

$P(AUB) = P(A) + P(B)$ (Events are Mutually exclusives)

$P(AUB) = P(A) + P(B) - P(A \cap B)$ (Events are not Mutually exclusives)

ii) Multiplication Theorem : (Law)

$P(A \cap B) = P(A) \,.\, P(B)$ (Events are Independent)

$P(A \cap B) = P(A) \,.\, P(B/A)$ (Events are Dependent)

$P(A \cap B) = P(B) \,.\, P(A/B)$ (Events are Dependent)

iii) Conditional Probability : (Law)

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad : \quad P(B) > o$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad : \quad P(A) > o$$

iv) Complement Law of Probability :

$P(A) + P(\overline{A}) = 1$

(6) RANDOM VARIABLE AND PROBABILITY DISTRIBUTION

PROPERTIES OF : **(i) Discrete Probability Distribution:**

i) $f(x) > 0$ ii) $\Sigma P(xi) = 1$

(ii) Continuous Probability density function:

i) $f(x) > 0$ ii) $\int f(x)\,dx = 1$

OR Total Area $= \frac{1}{2}$[sum of ordinates $x = 1$

Properties of Expectation :

If a and b are two constants and x ,y are tow independent variables, then

i) $E(a) = a$ ii) $E(x + a) = E(x) + a$

iii) $E(ax) = aE(x)$ iv) $E(ax + b) = a\,E(x) + b$

v) $E(x + y) = E(x) + E(y)$ vi) $E(xy) = E(x)\,E(y)$

vii) $E[x - E(x)]) = E(x) - E(x)$

$= \text{Zero}$

BINOMIAL AND HYPERGEOMETRIC DISTRIBUTION

P.d.f of Binomial distribution:-

$$f(x) = \binom{n}{x} p^x q^{n-x} \qquad \text{for } x = 0,1,2 \ldots\ldots, n$$

Mean of the binomial distribution $= np$

Variance of the binomial distribution $= npq$

S.D. of the binomial distribution $= \sqrt{npq}$

P.d.f. of Hypergeometric distribution is :-

$$f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \qquad \text{for } x = 0,1,2 \ldots\ldots, n$$

Mean of the Hypergeometric distribution $= \frac{nK}{N}$

Variance of the Hypergeometric distribution $= \frac{nK}{N} \left(\frac{N-K}{N}\right) \left(\frac{N-n}{N-1}\right)$