STATISTICS FORMULA SHEET

CA Nishant Kumar

Stay updated with the latest FREE lectures: https://www.youtube.com/@ekagrataca



Chapter 13 – Statistical Description of Data CA NISHANT KUMAR

1. Frequency Density of a Class Interval = $\frac{Class Frequency}{Class Length}$ 2. Relative Frequency of a Class Interval = $\frac{Class Frequency}{Total Frequency}$ 3. Percentage Frequency of a Class Interval = $\frac{Class Frequency}{Total Frequency} \times 100$



Chapter 14 – Measures of Central Tendency and Dispersion



Topic 1 – Arithmetic Mean

1. Mean of Individual Series $\overline{x} = \frac{\overline{x} - \overline{x}}{n}$

2. Mean of Discrete/Continuous Series $\overline{x} = \frac{1}{\nabla x}$

(In continuous series, the mid-point of the class interval is taken to be *x*) 3. If two variables *x* and *y* are related as y = a + bx, and \overline{x} is known, then $\overline{y} = a + b\overline{x}$. 4. Combined Mean $\overline{x} = \frac{n_1 \overline{x}_1 + n_2 \overline{x}_2}{n_1 + n_2}$



Topic 2 – Geometric Mean

- 1. GM of Individual Series = n^{th} root of the product of *n* numbers
- 2. GM of Discrete/Continuous Series = $(x_1^{f_1} \times x_2^{f_2} \times ... \times x_n^{f_n})^{\overline{N}}$

(In continuous series, the mid-point of the class interval is taken to be *x*) 3. If z = xy, then *GM* of z = GM of $x \times GM$ of y

4. If $z = \frac{x}{y}$, then $GM \text{ of } z = \frac{GM \text{ of } x}{GM \text{ of } y}$

5. For a set of *r* observations, $\log G = \frac{1}{r} \sum \log x$

Topic 3 – Harmonic Mean 1. HM of individual series HM = -2. HM of Discrete/Continuous Series HM =(In continuous series, the mid-point of the class interval is taken to be x) 3. Combined *HM* n_2 CA NISHANT KUMAR

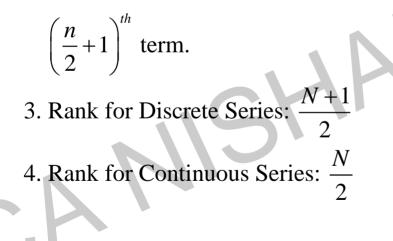
- 4. The harmonic mean of 1, $\frac{1}{2}$, $\frac{1}{3}$, ..., $\frac{1}{n}$ is given by: $\frac{2}{n+1}$
- 5. Average Speed = $\frac{2xy}{x+y}$

6. Relationship between *AM*, *GM*, *HM* for constant observations: AM = GM = HM7. Relationship between *AM*, *GM*, *HM* for distinct observations: AM > GM > HM8. For positive observations, $GM^2 = AM \times HM$

Topic 4 – Median

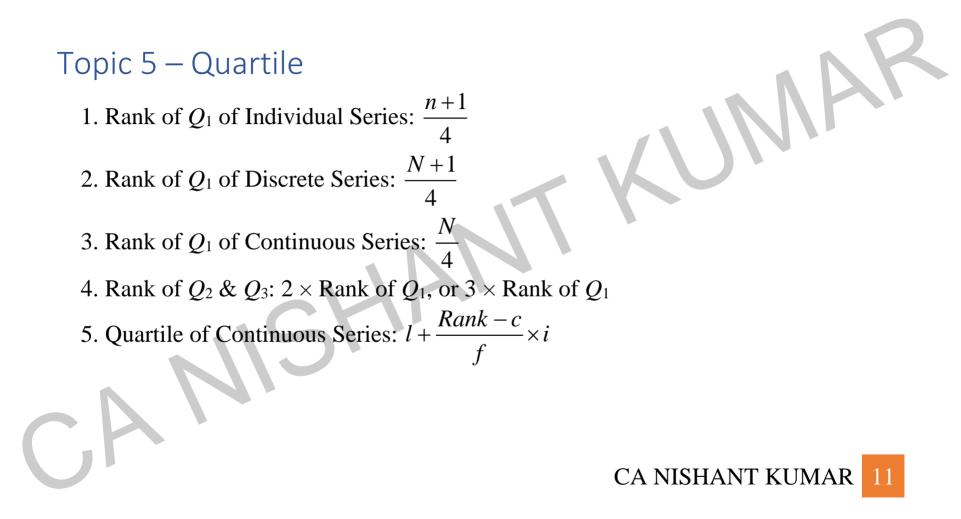
1. For odd number of items, Rank of Median of Individual Series: $\frac{n+1}{2}$

2. For even number of items, Median is given by the average of $\left(\frac{n}{2}\right)^m$ term and





5. Median of Continuous Series: $l + \frac{Rank - c}{f} \times i$ 6. If two variables x and y are related as y = a + bx, and x_{me} is known, then $y_{me} = a + bx_{me}$. CA NISHANT KUMAR 10



Topic 6 – Decile 1. Rank of D_1 of Individual Series: $\frac{n+1}{2}$ 10 2. Rank of D_1 of Discrete Series: $\frac{N+1}{2}$ 10 3. Rank of D_1 of Continuous Series: 10 4. Rank of D_2 & D_3 : 2 × Rank of D_1 , or 3 × Rank of D_1 5. Decile of Continuous Series: $l + \frac{Rank - c}{c} \times i$ CA NISHANT KUMAR Topic 7 – Percentile 1. Rank of P_1 of Individual Series: $\frac{n+1}{n}$ 100 N+12. Rank of P_1 of Discrete Series: 100 3. Rank of P_1 of Continuous Series: 1004. Rank of P_2 & P_3 : 2 × Rank of P_1 , or 3 × Rank of P_1 5. Percentile of Continuous Series: $l + \frac{Rank - c}{c} \times i$ CA NISHANT KUMAR 13

Topic 8 – Mode

- 1. Mode of Continuous Series = $l + \frac{f_1 f_0}{2f_1 f_0 f_2} \times i$
- 2. If two variables x and y are related as y = a + bx, and x_{mo} is known, then $y_{mo} = a + bx_{mo}$.
- 3. Relationship between Mean, Median, Mode
 - a. For Symmetric Data: Mean = Median = Mode
 - b. For Skew Symmetric Data: Mode = 3Median 2Mean
 - c. For Positively Skewed Data: Mean > Median > Mode
 - d. For Negatively Skewed Data: Mean < Median < Mode



Topic 9 – Range

- 1. Range of Individual Series = Largest Observation Smallest Observation
- 2. Coefficient of Range = $\frac{Largest \ Observation Smallest \ Observation}{Largest \ Observation + Smallest \ Observation}$
- 3. Range of Continuous Series = Upper Most Class Boundary Lower Most Class Boundary
- 4. Coefficient of Range = $\frac{UMCB LMCB}{UMCB + LMCB}$

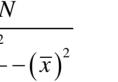
5. If two variables x and y are related as y = a + bx, and R_x is known, then $R_y = |b| \times R_x$.



Topic 10 – Mean Deviation $\underline{\sum} |x - A|$ 1. Mean Deviation of Individual Series = n 2. Mean Deviation of Discrete/Continuous Series = NMean Deviation About A 3. Coefficient of Mean Deviation $\times 100$ 4. If two variables x and y are related as y = a + bx, and MD_x is known, then $MD_y = |b| \times MD_x.$ CA NISHANT KUMAR

Topic 11 – Standard Deviation

- 1. Standard Deviation of Individual Series = $\sqrt{\frac{\sum(x-\overline{x})}{\sum(x-\overline{x})}}$
- 2. Standard Deviation of Individual Series = $\sqrt{\frac{2}{1}}$
- 3. Standard Deviation of Discrete/Continuous Series = $\sqrt{\frac{\sum f(x-\overline{x})^2}{NT}}$
- 4. Standard Deviation of Discrete/Continuous Series = $\sqrt{\frac{\sum fx^2}{NT} (\bar{x})^2}$
- 5. Coefficient of Variation = $\frac{SD}{AM} \times 100$



- 6. Variance = Square of Standard Deviation
- 7. Standard Deviation of first *n* natural numbers = $\sqrt{1}$
- 8. Standard Deviation of only two numbers *a* and *b* = $\frac{|a-b|}{2}$
- 9. If two variables x and y are related as y = a + bx, and SD_x is known, then $SD_y = |b| \times SD_x$.
- 10. Combined Standard Deviation = $\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}}$ Here, $d_1 = \overline{x_1} - \overline{x}$; $d_2 = \overline{x_2} - \overline{x}$



Topic 12 – Quartile Deviation 1. Quartile Deviation = $\frac{Q_3 - Q_1}{2}$ 2. Coefficient of Quartile Deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$ 3. If two variables x and y are related as y = a + bx, and QD_x is known, then $QD_{v} = |b| \times QD_{x}$. CA NISHANT KUMAR 19

Chapter 15 – Probability

No. of Favourable Cases/Events/Outcomes Total No. of Cases/Events/Outcomes 1. Probability = 2. When an experiment with total number of events a is repeated b number of times, the total number of outcomes is given by a^b . 3. Odds in Favour of Event $A = \frac{\text{Number of Favourable Outcomes}}{\text{Number of Unfavourable Outcomes}}$ Number of Unfavourable Outcomes 4. Odds Against Event A =Number of Favourable Outcomes and *B* 5. If exclusive. two events are not mutually then. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



- 6. Two events *A* and *B* are mutually exclusive, if $A \cap B = \phi$. Therefore, $P(A \cap B) = 0$, or $P(A \cup B) = P(A) + P(B)$.
- 7. If three events A, B, and C are not mutually exclusive, then $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
- 8. Three events A, B, and C are mutually exclusive, if $P(A \cup B \cup C) = P(A) + P(B) + P(C)$.
- 9. Probability that only event *A* occurs: $P(A-B) = P(A) P(A \cap B)$
- 10. Probability that only event *B* occurs: $P(B-A) = P(B) P(A \cap B)$
- 11. $P(A \cap B) = P(A) \times P(B)$
- 12. Probability of event *A* given that event *B* has already occurred is given by P(A/B):



$$P(A / B) = \frac{P(A \cap B)}{P(B)}$$

13. Probability of event *B* given that event *A* has already occurred is given by P(B|A):

$$P(B/A) = \frac{P(B \cap A)}{P(A)}$$
14.
$$P(A'/B) = \frac{P(A' \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)}$$
15.
$$P(A/B') = \frac{P(A \cap B')}{P(B')} = \frac{P(A) - P(A \cap B)}{1 - P(B)}$$



16.
$$P(A'/B') = \frac{P(A' \cap B')}{P(B')} = \frac{P(A \cup B)'}{1 - P(B)} = \frac{1 - P(A \cup B)'}{1 - P(B)}$$

- 17. Probability that only event *A* or only event *B* occurs: $P(A) + P(B) - 2P(A \cap B)$
- 18. Expected value (μ) of a random variable (x) is given by: $\mu = E(x) = \sum p_i x_i$
- 19. Expected value of (x^2) is given by: $E(x^2) = \sum \left[p_i(x_i^2) \right]$
- 20. Variance (σ^2) of a random variable (x) is given by: $V(x) = \sigma^2 = E(x - \mu)^2 = E(x^2) - \mu^2$
- 21. If *a* and *b* are two constants related with two random variables *x* and *y* as y = a + bx, then the mean, i.e., the expected value of *y* is given by: $\mu_y = a + b\mu_x$.

- 22. If *a* and *b* are two constants related with two random variables *x* and *y* as y = a + bx, then the standard deviation of *y* is given by: $\sigma_y = |b| \times \sigma_x$.
- 23. Expectation of sum of two random variables is the sum of their expectations, i.e., E(x+y) = E(x) + E(y), for any two random variables *x* and *y*.
- 24. Expectation of the product of two random variables is the product of the expectation of the two random variables, provided the two variables are independent, i.e., $E(x \times y) = E(x) \times E(y)$.



Chapter 16 Distributions

Topic 1 – Binomial Distributions

1. $P(x) = {}^{n}C_{x}p^{x}q^{n-x}$, for x = 0, 1, 2, 3, ..., n

2. The mean of the binomial distribution is given by $\mu = np$.

3. A binominal distribution is symmetrical when p = q.

4. Mode of a Binomial Distribution is given by largest integer contained in $\mu_0 = (n+1)p$, OR $\mu_0 = (n+1)p$ and [(n+1)p]-1



Theoretical

5. The variance of the binomial distribution is given by $\sigma^2 = npq$.

6. If p = q = 0.5, variance is the maximum, and is given by $\frac{n}{4}$.

7. Standard Deviation of a binomial distribution is given by $\sigma = \sqrt{npq}$.

8. Let x and y be two independent binomial distributions where x has the parameters n_1 and p, and y has the parameters n_2 and p. Then (x+y) will be a binomial distribution with parameters $(n_1 + n_2)$ and p.



Topic 2 – Poisson Distribution

1.
$$P(x) = \frac{e^{-m} \times m^x}{x!}$$
, for $x = 0, 1, 2, 3, ..., n$

- 2. The mean of Poisson distribution is given by *m*, i.e., $\mu = m = np$.
- 3. The variance of Poisson distribution is given by $\sigma^2 = m = np$.
- 4. The standard deviation of Poisson distribution is given by $\sigma = \sqrt{m} = \sqrt{np}$.
- 5. Mode of a Poisson Distribution is given by largest integer contained in m, OR m and m-1
- 6. Let x and y be two independent poisson distributions where x has the parameter m_1 , and y has the parameter m_2 . Then (x + y) will be a poisson distribution with parameter $(m_1 + m_2)$.



Topic 3 – Normal Distribution

1.
$$P(x) = f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}$$
, for $-\infty < x < \infty$

2. If $\mu = 0$, and $\sigma = 1$, the variable is known as the standard normal variate (or variable).

It is given by
$$z = \frac{x - \mu}{\sigma}$$
.

- 3. Relationship between MD, SD, and $QD \rightarrow 4SD = 5MD = 6QD$
- 4. Mean Deviation = 0.8σ .
- 5. Quartile Deviation = 0.675σ .
- 6. $Q_1 = \mu 0.675\sigma$
- 7. $Q_3 = \mu + 0.675\sigma$
- 8. Median $-Q_1 = Q_3 Median$.



9. Points of inflexion are given by $x = \mu - \sigma$ and $x = \mu + \sigma$ 10. If there are two Independent Normal Distributions $x \sim N(\mu_1, \sigma_1^2)$ and $y \sim N(\mu_2, \sigma_2^2)$, then z = x + y follows normal distribution with mean $(\mu_1 + \mu_2)$ and $SD = \sqrt{\sigma_1^2 + \sigma_2^2}$ respectively.



Correlation and Chapter 17 – Regression CA NISHANT KUMAR 30

Topic 1 – Karl Pearson's Product Moment Correlation Coefficient

1.
$$r = r_{xy} = \frac{Cov(x, y)}{S_x \times S_y} = \frac{Cov(x, y)}{\sigma_x \times \sigma_y}$$

a. $Cov(x, y) = \frac{\sum (x - \overline{x})(y - \overline{y})}{n} = \frac{\sum xy}{n} - \overline{x}.\overline{y}$
b. $S_x = \sqrt{\frac{\sum (x - \overline{x})^2}{n}} = \sqrt{\frac{\sum x^2}{n} - \overline{x}^2}$
c. $S_y = \sqrt{\frac{\sum (y - \overline{y})^2}{n}} = \sqrt{\frac{\sum y^2}{n} - \overline{y}^2}$



2.
$$r = \frac{n\sum xy - \sum x \times \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \sqrt{n\sum y^2 - (\sum y)^2}}$$

- 3. Let there be two variables x and y. Let the correlation coefficient between them be
 - r_{xv} . Now, if they are changed to another set of variables, say, u and v, then,

$$r_{uv} = r_{xy}$$
, if *b* and *d* have the same sign, or
 $r_{uv} = -r_{xy}$, if *b* and *d* have opposite signs.
Here,

Here,

$$b = \frac{-\text{Coefficient of } u}{\text{Coefficient of } x}, \text{ and } d = \frac{-\text{Coefficient of } v}{\text{Coefficient of } y}$$
4. $r_{xy} = \frac{bd}{|b||d|} \cdot r_{uv}$



Topic 2 – Spearman's Rank Correlation Coefficient 1. $r_R = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$ 6 12 2. When ranks repeat, then $r_R = 1$ CA NISHANT KUMAR 33

Topic 3 – Coefficient of Concurrent Deviations 1. If (2c - m) > 0, then $r_c = \sqrt{\frac{(2c - m)}{m}}$ (2c-m)2. If (2c - m) < 0, then $r_c = -\sqrt{2}$ т CA NISHANT KUMAR 34

Topic 4 – Regression

1. if y depends on x, then the regression line of y on x is given by: a. either y = a + bx, where,

i.
$$b = b_{yx} = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$
, or, $b_{yx} = r\frac{\sigma_y}{\sigma_x}$, or, $b_{yx} = \frac{Cov(x, y)}{(\sigma_x)^2}$, and
ii. $a = a_{yx} = \overline{y} - (\overline{x} \times b_{yx})$

 b_{yx} is known as the regression coefficient.

b. or,
$$(y - \overline{y}) = b_{yx}(x - \overline{x})$$

2. if x depends on y, then the regression line of x on y is given by:

a. either x = a + by, where,



i.
$$b = b_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum y^2 - (\sum y)^2}$$
, or, $b_{xy} = r\frac{\sigma_x}{\sigma_y}$, or, $b_{xy} = \frac{Cov(x, y)}{(\sigma_y)^2}$, and
ii. $a = a_{xy} = \overline{x} - (\overline{y} \times b_{xy})$
 b_{xy} is known as the regression coefficient.
or, $(x - \overline{x}) = b_{xy}(y - \overline{y})$
3. $b_{yu} = \frac{Scale \ of \ v}{Scale \ of \ u} \times b_{yx}$
4. $r = \sqrt{b_{yx} \times b_{xy}}$
CA NISHANT KUMAR 36

Topic 5 – Probable Error and Standard Error

1. Probable Error (P.E.) is given by

a.
$$P.E. = 0.674 \times \frac{1 - r^2}{\sqrt{N}}$$
, or
b. $P.E. = 0.6745 \times \frac{1 - r^2}{\sqrt{N}}$, or
c. $P.E. = 0.675 \times \frac{1 - r^2}{\sqrt{N}}$

2. Limits of the correlation coefficient of the population is given by $p = r \pm P.E$.

3. Standard Error (S.E.) is given by $S.E. = \frac{1-r^2}{\sqrt{N}}$





Topic 6 – Coefficient of Determination and Non-Determination

- 1. Coefficient of Determination (Also known as "Percentage of Variation Accounted for") $(r^2) = \frac{\text{Explained Variance}}{\text{Total Variance}}$
- 2. Coefficient of Non-Determination (Also known as "Percentage of Variation Unaccounted for") = $1 r^2$



Chapter 18 – Index Numbers

 $\frac{P_{1}}{-1} \times 100^{-1}$

 P_0

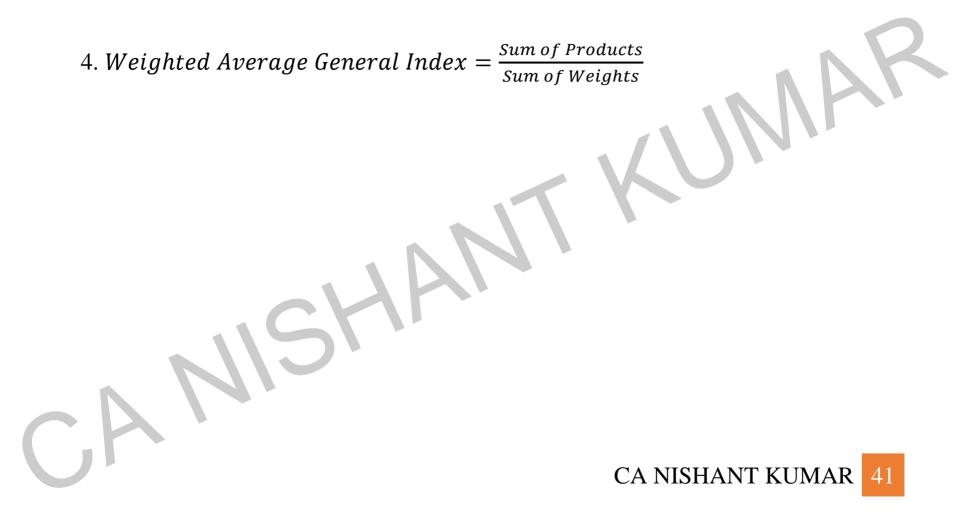
N

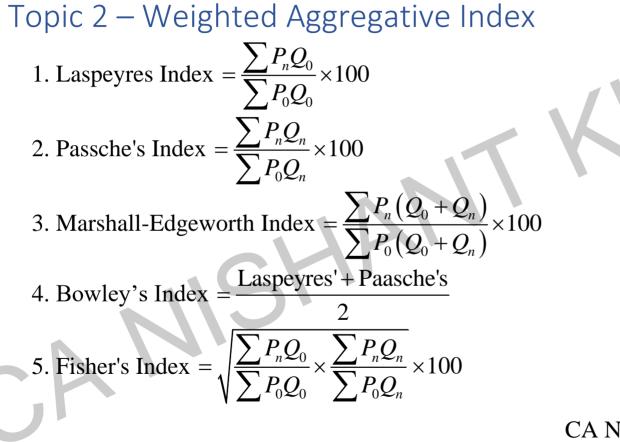
Topic 1 – Price Relative

- 1. Price Relative = $\frac{P_n}{P_0} \times 100$.
- 2. Simple Aggregative Price Index = $\frac{\sum r_n}{\sum P} \times 100$

3. Simple Average of Price Relatives = —

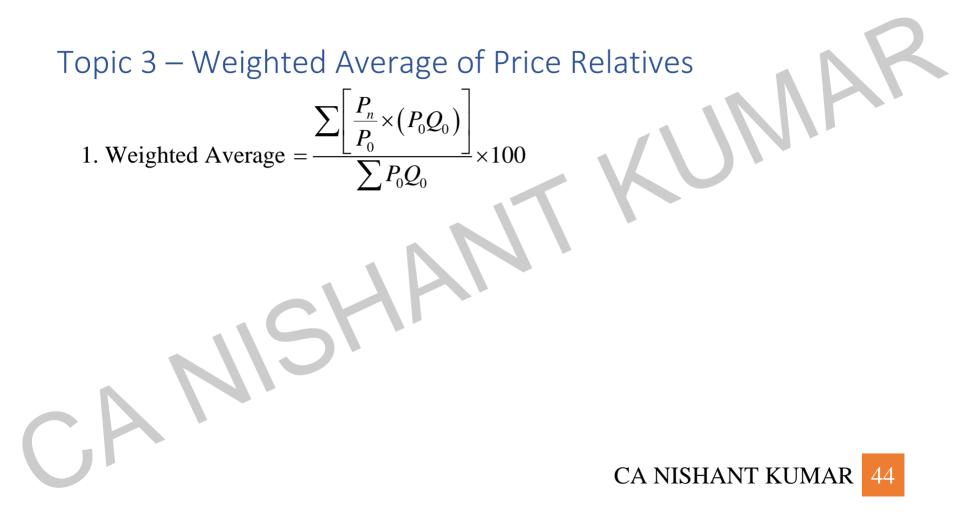






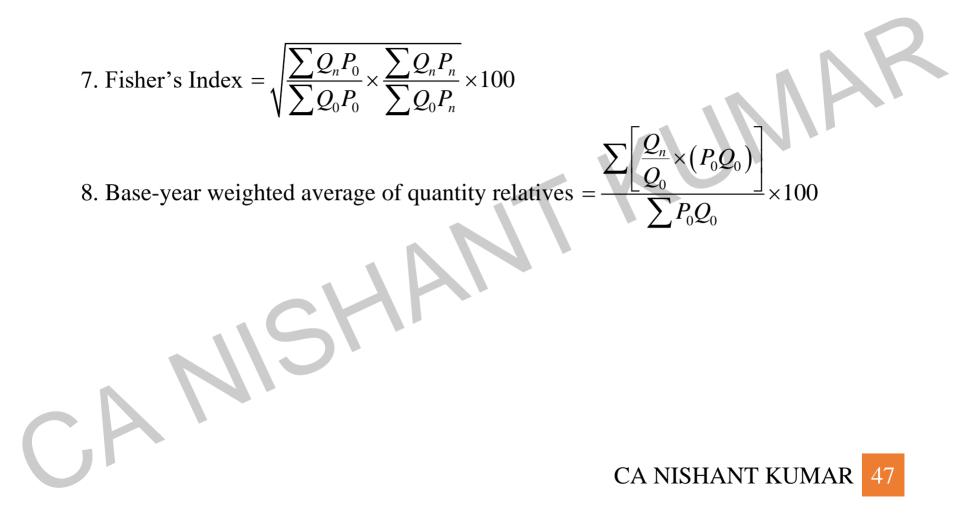


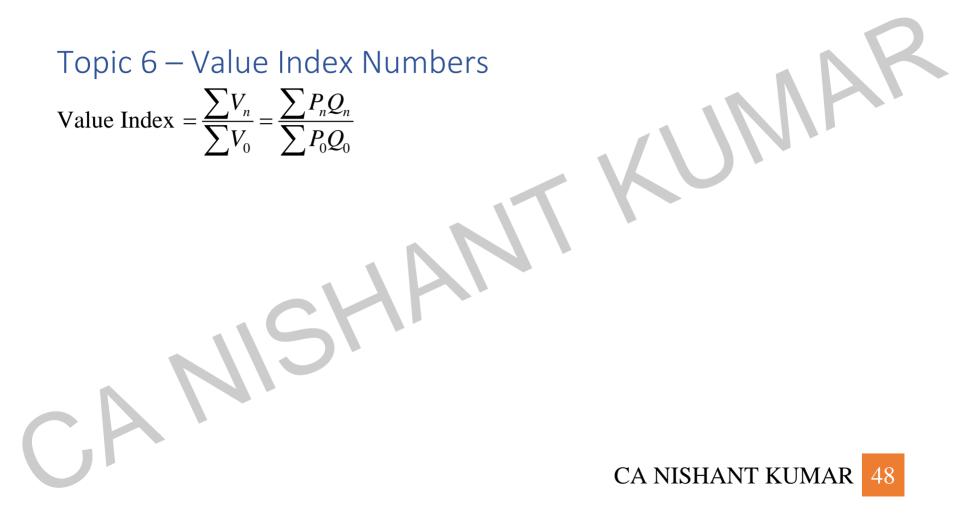




Topic 4 – The Chain Index Numbers 1. Chain Index = $\frac{\text{Link Relative of the Current Year } \times \text{Chain Index of the Previous Year}}{1}$ 100 CA NISHANT KUMAR 45

Topic 5 – Quantity Index Numbers 1. Simple Aggregate of Quantities = $\frac{\sum Q_n}{\sum Q_0} \times 100$ 2. Simple Average of Quantity Relatives = - $-\times 100$ n 3. Laspeyre's Index = 6. Paasche's Index $\frac{n}{2} \times 100$ CA NISHANT KUMAR 46





Topic 7 – Deflating Time Series Using Index Numbers

1. Deflated Value =	Current Value
	Price Index of the Current Year
2 ($\mu rrent Value = -$	Base Price (P ₀)
	Current Price (P _n)
3. Real Wages = $\frac{A}{Cost}$	Actual Wages × 100
	t of Living Index × 100



Topic 8 – Shifting and Splicing of Index Numbers

1. Shifted Price Index = $\frac{\text{Original Price Index}}{\text{Price Index of the Year on which it has to be shifted}}$



× 100

Topic 9 – Test of Adequacy 1. Time Reversal Test $P_{01} \times P_{10} = 1$ 2. Factor Reversal Test $P_{01} \times Q_{01} = V_{01}$ 3. Circular Test $P_{01} \times P_{12} \times P_{20} = 1$ CA NISHANT KUMAR

51