NCERT SOLUTIONS

CLASS - 12th

aglasem.com

Class : 12th

Subject : Geography

Chapter : 2

Chapter Name : Data processing

Q1 Choose the correct answer from the four alternatives given below :

(i) The measure of central tendency that does not get affected by extreme values:

(a) Mean

(b) Mean and Mode

(c) Mode

(d) Median

(ii) The measure of central tendency always coinciding with the hump of any distribution is:

(a) Median

(b) Median and Mode

(c) Mean

(d) Mode

(iii) A scatter plot represents a negative correlation if the plotted values run from:

(a) Upper left to lower right

(b) Lower left to upper right

(c) Left to right

(d) Upper right to lower left

Answer.

(i) (d) Median

(ii) (b) Median and Mode

(iii) (a) Upper left to lower right

Page : 31 , Block Name : Multiple Choice Questions

Q2 Answer the following questions in about 30 words :

(i) Define the mean.

(ii) What are the advantages of using mode?

(iii) What is dispersion?

(iv) Define correlation.

(v) What is perfect correlation?

(vi) What is the maximum extent of correlation?

Answer.

(i) The mean is the average value of the data set which is obtained by summing up the values of all observations and dividing it by the number of observations.

The mean may be computed as under using the given formula : $X = \sum fx / N$

Where f is frequency , x is the value of observation, N is the total number of observations

(ii) Mode is that value which has maximum occurring frequency in data set.

Advantages

1. It is not affected by extreme values.

2. It can be determined for open ended series.
3. It is easy to understand and calculate

(iii) It refers to the extent of scattering of scores about the central tendency. It measures to what extent the individual numerical data vary from the average value. It gives dispersion and variability i.e compactness or spreadness of data.

(iv) Correlation is defined as the relationship between two or more sets of data. It is the degree to which a pair of variables is linearly related. The value varies between -1 to 1.

(v) Perfect correlation can be negative or positive. It is proportional relationship of equal magnitude between two variables.

Eg.

1. Positive Correlation- If on doubling x, the value of y also gets doubled , it is a perfect positive correlation.
2. Negative Correlation - If on doubling the variable x, y gets halved, it is called perfect negative correlation.

(vi) The maximum extent of correlation is -1 to +1, through zero. This can never be greater than 1. When the correlation is 1 (be it negative or positive) it is known as perfect correlation.

Page : 31 , Block Name : Short Answer

Q3 Answer the following questions in about 125 words :

(i) Explain relative positions of mean, median and mode in a normal distribution and skewed distribution with the help of diagrams.

(ii) Comment on the applicability of mean, median and mode.

(iii) Explain the process of computing Standard Deviation with the help of an imaginary example.

(iv) Which measure of dispersion is the most unstable statistic and why?

(v) Write a detailed note on the degree of correlation.

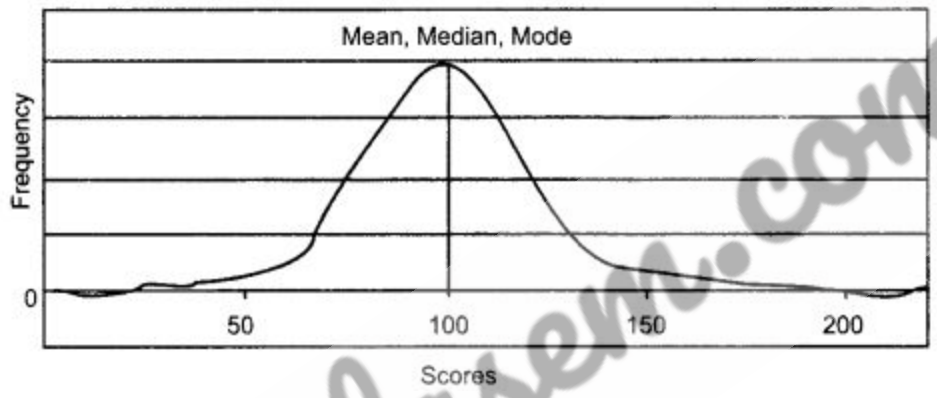
(vi) What are various steps for the calculation of rank order correlation?

Answer.

(i)

The three measures of the central tendency viz. mean, median and mode can be compared with the help of normal distribution curve

The characteristic feature of the normal curve is that it is bell-shaped and symmetrical. Maximum values in the middle and as one approaches the tails, the number of observations reduce. A normal curve can have high or low data variability.Example is given in the following figure:

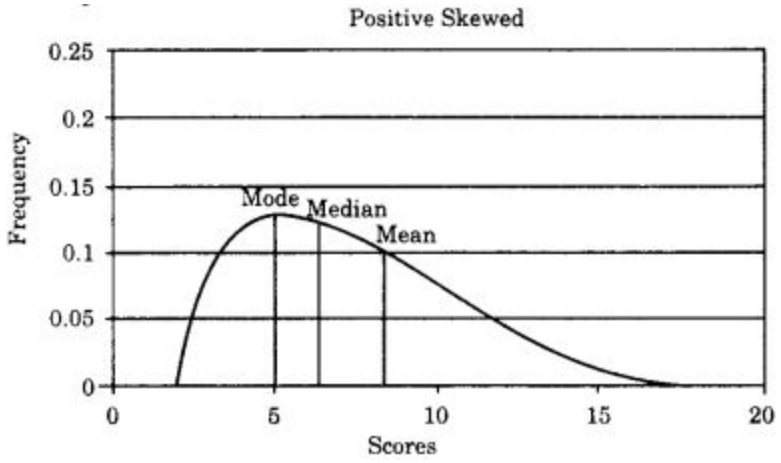


As seen the mean, median and mode coincide with the same score. This can be understood as:

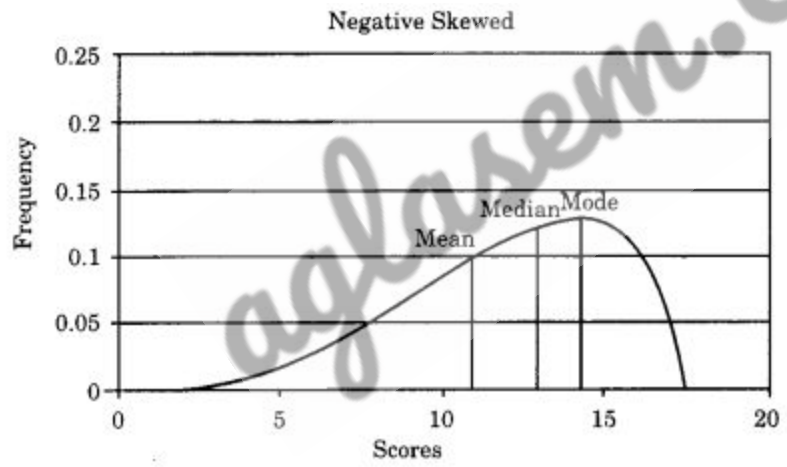
1. The score with the highest frequency occurs in the middle of the distribution (Mode) and exactly half of the scores occur above the middle and half of the scores occur below. (Median)
2. Most of the scores occur around the middle of the distribution or the mean.
3. Very high and very low scores do not occur frequently

However in case of skewness of data , the mean, median and mode will not coincide. This is represented with the help of following illustrations.

1. Positive-skew : Here, the right hand side tail is longer and mass of data is concentrated in the left which point to greater frequency of lower values.This is positively skewed normal curve.



2. Negative-skew : Here, the left hand side tail is longer and mass of data is concentrated in the right which point to greater frequency of higher values.This is a negatively skewed normal curve.



(ii) **Mean:**

Applicability grounds: Mean includes all the items of a dataset that are in a series and it gives equal importance to each of them. Therefore it is best used in a case where all data is important.

Applicability examples : Average runs of a player, Average income of employees , Average Score, Average number of children per family etc

It acts as a general representative value.It should be avoided when data is susceptible to outliers.

Median:

Applicability grounds: Median finds the middle values and is not affected by the extreme values of the series,it gives the mid point which divides data into two equal halves- Half bottom,Half top.It is best used when extreme values are less and need to be ignored.

Applicability Example: Runs scored by a leading cricketer are 50,50,51,52,53,54,260. Here median is 52 which more closely represents the reality.

Mode:

Applicability ground: Mode finds the most frequently occurring value.

Applicability Example: If we need to find the most favourite color of a batch of students,then the one chosen most will be mode.

(iii) Standard Deviation quantifies the amount of dispersion and variance of data about the average. It is defined as the square root of the average of the squares of deviations. To obtain SD, deviation of each score from the mean (x) is first squared.Then, all of the squared deviations are summed. This sum of the squared deviations is divided by the number of cases and then the square root is taken. Therefore, Standard Deviation is defined as the root mean square deviation. The formula is :

$$\sigma = \sqrt{\frac{\sum [x - \bar{x}]^2}{n}}$$

σ = standard deviation

$\sum$ = sum of

x = each value in the data set

$\bar{x}$ = mean of all values in the data set

n = number of value in the data set

Example :

For a given data set, standard deviation, Calculate the standard deviation for the following scores :

01, 03, 05, 07, 09

X	$x = (X - \bar{X})$	x^2
1	-4	16
3	-2	4
5	0	0
7	2	4
9	-16	16

$\sum X = 25$
$N = 5$

= Under Root(40/ 5) = under root(8) = 2.828

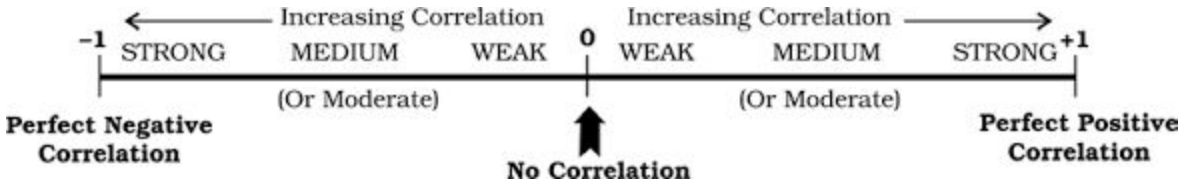
(iv) Range is the most unstable statistic. This is because:

1. Range is not based on all terms rather it is only based on the extreme values, all middle values are ignored.
2. Range is not representative of the data, it is a very crude measure.
3. Therefore it is also not a reliable measure of dispersion.
4. Ranges may not change even if all other values in between are changed

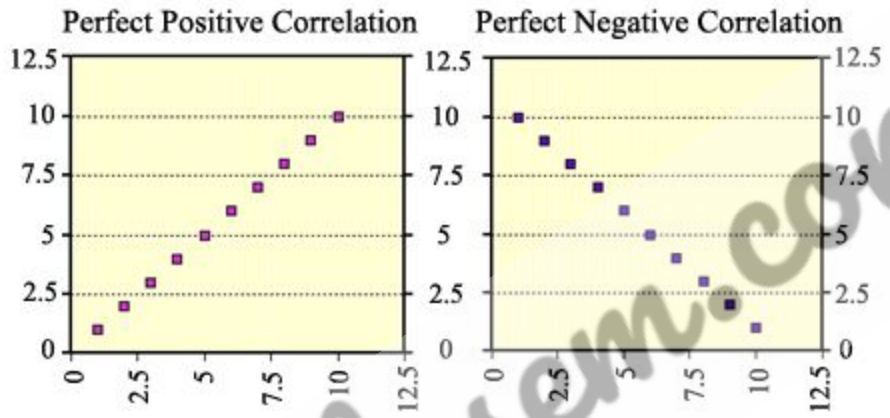
(v) Correlation is the degree to which a pair of variables is linearly related. The value varies between -1 to 1. It can be negatively or positively correlated. The maximum degree of correspondence goes upto 1.It can never be more than one.

Correlation of 1 or -1 is known as perfect correlation (positive and negative respectively). Between the two points of divergent, perfect correlations lies 0 (zero) correlation, a point of no correlation or absence of any correlation between the variables.

Spread of Direction and Degree of Correlation

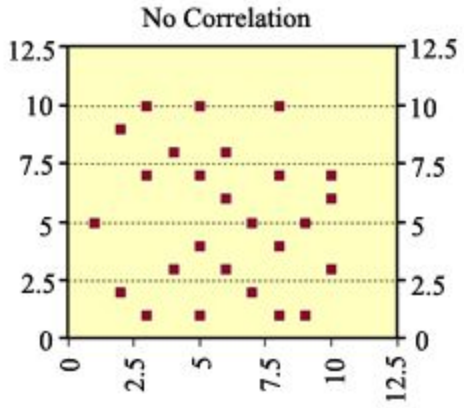


Perfect Correlations



Here, The pairs of values, when plotted, fall along a straight line and when this straight line runs from the lower left of the scatter plot to the upper right, it is an example of a perfect positive correlation (1.00) and when it runs from the upper left-hand part of the scatter gram to its lower right. It is an example of a perfect negative correlation (– 1.00).

No Correlation is one when any of the variables in the pair does not respond to the changes in the other as shown in the following figure.



Other Correlations : Between the perfect correlations (±1) and zero correlation lies weak, moderate and strong correlations. Larger the scattering, the weaker is the correlation. Smaller the scattering, stronger is the correlation.

This gives an overview of what is the degree of correlation, how it affects the data graphically and what it means.

(vi) Step 1: Rank both sets of data. Give the largest value rank 1, the second largest value rank 2, etc.

Step 2: Calculate the differences in the ranks, d.

Step 3: Work out the squares of the differences (d2).

Step 4: Calculate the sum of these squared differences,

Step 5: Spearman's Rank Correlation Coefficient is found by substituting this sum into the following formula:

$$1-\frac{6\times\sum d^2}{n\left(n^2-1\right)},$$

Where,

r = rank correlation

$\sum d^2$ = sum of the squares of the differences between the two sets of ranks

N = the number of pairs of X-Y

where n is how many pairs of data you have.

Page : 31 , Block Name : Long Answer