Partition Values: Quartile, Deciles, Percentiles

Partition values or fractiles such a quartile, a decile, etc. are the different sides of the same story. In other words, these are values that divide the same set of observations in different ways. So, we can fragment these observations into several equal parts.

Quartile

Whenever we have an observation and we wish to divide it, there is a chance to do it in different ways. So, we use the median when a given observation is divided into two parts that are equal. Likewise, quartiles are values that divide a complete given set of observations into four equal parts.

Basically, there are three types of quartiles, first quartile, second quartile, and third quartile. The other name for the first quartile is lower quartile. The representation of the first quartile is 'Q_1.' The other name for the second quartile is median. The representation of the second quartile is by 'Q_2.' The other name for the third quartile is the upper quartile. The representation of the third quartile is by 'Q_3.'

- **First Quartile** is generally the one-fourth of any sort of observation. However, the point to note here is, this one-fourth value is always less than or equal to 'Q_1.' Similarly, it goes for the values of 'Q_2'and 'Q3.'

Deciles

Deciles are those values that divide any set of a given observation into a total of ten equal parts. Therefore, there are a total of nine deciles. These representations of these deciles are as follows – D_1, D2, D_3, D_4 … D_9.

D_1 is the typical peak value for which one-tenth (1/10) of any given observation is either less or equal to D_1. However, the remaining nine-tenths (9/10) of the same observation is either greater than or equal to the value of D_1.

Percentiles

Last but not the least, comes the percentiles. The other name for percentiles is centiles. A centile or a percentile basically divides any given observation into a total of 100 equal parts. The representation of these percentiles or centiles is given as – P_1, P2, P_3, P_4 … P_{99}.

P_1 is the typical peak value for which one-hundredth (1/100) of any given observation is either less or equal to P_1. However, the remaining ninety-nine-hundredth (99/100) of the same observation is either greater than or equal to the value of P_1. This takes place once all the given observations are arranged in a specific manner i.e. ascending order.

So, in case the data we have doesn't have a proper classification, then the representation of p^{th} quartile is **(n + 1) pth**

Here,

n = total number of observations.

p = 1/4, 2/4, 3/4 for different values of Q_1, Q_2, and Q_3 respectively.

p = 1/10, 2/10, …. 9/10 for different values of D_1, D_2, …… D_9 respectively.

p = 1/100, 2/100, ….. 99/100 for different values of P_1, P_2, ……… P_{99} respectively.

Formula

At times, the grouping of frequency distribution takes place. For which, we use the following formula during the computation:

$$\mathbf{Q = l_1 + [(N_p - N_i)/(N_u - N_i)] * C}$$

Here,

$\mathbf{l_1}$ = lower class boundary of the specific class that contains the median.

$\mathbf{N_i}$ = less than the cumulative frequency in correspondence to $\mathbf{l_1}$ (Post Median Class)

$\mathbf{N_u}$ = less than the cumulative frequency in correspondence to $\mathbf{l_2}$ (Pre Median Class)

C = Length of the median class ($\mathbf{l_2 - l_1}$)

The symbol 'p' has its usual value. The value of 'p' varies completely depending on the type of quartile. There are different ways to find values or quartiles. We use this way in a grouped frequency distribution. The best way to do it is by drawing an *ogive* for the present frequency distribution.

Hence, all that we need to do to find one specific quartile is, find the point and draw a horizontal axis through the same. This horizontal line must pass through N_p. The next step is to draw a perpendicular. The perpendicular comes up from the same point of intersection of the ogive and the horizontal line. Hence, the value of the quartile comes from the value of 'x' of the given perpendicular line.

Merits

1. They are easy to determine especially in case of the individual and discrete series.
2. They do not need all the data relating to a series like the mathematical averages viz. AM… G.M. and H.M.

3. They can be directly determined in case of an open end series without locating the lower limit of the lowest class, and the upper limit of the highest class.

4. They are useful in the computation of the measures of dispersion and skewness.

5. They give an idea about the character of a frequency distribution i.e. whether a series is symmetric, or asymmetric can be known by measuring their distance from the Median.

6. They are not affected very much by the extreme values of a series.

7. They can be located both graphically and tabularly.

Demerits

1. These averages are not easily understood by a common man.

2. The determination of their values in case of continuous series becomes cumbersome as it involves application of the formula of interpolation.

3. They are not based on all the observations of a series.

4. They need the rearrangement of series in the ascending order if given otherwise.

5. They do not study the entire data. For example, Q1, studies only, first 25%, Q2 only first 50%, and Q3 only first 75% of the data.

6. They are not capable of further algebraic treatment except in the computation of quartile deviation and coefficient of skewness.

7. They are affected very much by fluctuating of sampling.

8. They are influenced much by the number of items rather than their values.