

Percentiles and Quartiles

$$\text{Percentage} = \frac{\text{Sum of all}}{\text{Total number}} \times 100$$

A percentile is a value below which a certain percentage of observation lies.

ex → 99 percentile means a person got better marks than 99% people.

$$\text{Percentile rank of } x = \frac{\text{No. of value below } x}{\text{Total values}}$$

It may tell us about the distribution of data, like how much data in how much percentile.

Two value can have same percentile rank.

$$\text{value index} = \frac{\text{Percentile}}{100} \times (\text{Total values} + 1)$$

* 5 Number theory

[A] Minimum

[B] First Quartile (25 percentile)

(θ_1)

[C] Median

[D] Third Quartile (75 percentile)

(θ_3)

[E] Maximum

→ Used to remove outliers

→ used to draw boxplot to remove outliers

→ To remove outliers we use fences.

→ we keep the values between the lower and higher fence and remove others.

$$\text{Lower fence} = \theta_1 - 1.5[\text{IQR}]$$

(LF)

$$\text{Higher fence} = \theta_3 + 1.5[\text{IQR}]$$

(HF)

where,

IQR is Inter Quartile range,

$$\text{IQR} = \theta_3 - \theta_1$$

→ To treat outliers make fence

→ To visualize outliers use Boxplot.

→ Drawing Boxplot

Ex - {1, 2, 2, 3, 3, 3, 4, 5, 5, 6, 6, 6, 6, 7, 8, 8, 9, 27}

$$\theta_1 = \frac{25}{100} \times (20+1) = 5.25$$

(ind)

$$\Rightarrow \theta_1 \text{ val} = \frac{3+2}{2} = 3$$

$$\theta_3 = \frac{75}{100} \times (20+1) = 15.75$$

(ind)

$$\Rightarrow \theta_3 \text{ val} = \frac{8+17}{2} = 12.5$$

$$\text{So } LF = 3 - 1.5 \times 3 = -3.75$$

$$HF = 3 + 1.5 \times 7.5 = 14.25$$

Our values must be within,

$$[-3.75 \text{ to } 14.25]$$

So, 27 is only outlier in our dataset.

Now,

minimum = 1, $\theta_1 = 3$, median = 5

$\theta_3 = 7.5$, maximum (without outliers) = 9.

