Answer on Question #62250 – Math – Statistics and Probability

Question

The set of data a={ 12,6,7,3,15,10,18,5 } and b={ 9,3,8,8,9,8,9,18 }.

i. Measure the central tendency (mean, median, mode, GM, HM) of a given data set and explain which one is the poor measure

ii. Measure the dispersion (range, mean deviation, standard deviation, variance) of a given data set and explain which one is the poor measure

iii. Measure the skewness and kurtosis of a given dataset and explain

Solution

i. For a={ 12,6,7,3,15,10,18,5 }:

$$mean = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{12+6+7+3+15+10+18+5}{8} = 9.5$$

Rearrange elements of the set a as follows:

a={ 3, 5,6,7, 10, 12, 15,18 }.

The sample size $n = 8$ is even, hence the median is the mean of the two middle values in the sorted list:

$median = \frac{7+10}{2} = 8.5.$

There is no mode, because there is no value that appears most often in the set of data.

$$GM = \sqrt[n]{x_1 x_2 x_3 \dots x_n} = \sqrt[8]{12 \cdot 6 \cdot 7 \cdot 3 \cdot 15 \cdot 10 \cdot 18 \cdot 5} = 8.2$$

$$HM = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{8}{\frac{1}{12}+\frac{1}{6}+\frac{1}{7}+\frac{1}{3}+\frac{1}{15}+\frac{1}{10}+\frac{1}{18}+\frac{1}{5}} = 7.0$$

We do not have the mode for this data. So, the mode is the poor measure of central tendency for this data.

For b={ 9,3,8,8,9,8,9,18 }:

$$mean = \frac{\sum_{i=1}^{m} y_i}{m} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

Rearrange elements of the set b as follows:

b={ 3,8,8, 8,9, 9, 9,18 }.

The sample size $m = 8$ is even, hence the median is the mean of the two middle values in the sorted list:

$median = \frac{8+9}{2} = 8.5.$

There are two modes: 8 and 9, because these are the values that appears most often in the set b of data

$$GM = \sqrt[m]{y_1 y_2 y_3 \dots y_m} = \sqrt[8]{9 \cdot 3 \cdot 8 \cdot 8 \cdot 9 \cdot 8 \cdot 9 \cdot 18} = 8.2$$

$$HM = \frac{m}{\sum_{i=1}^{m} \frac{1}{y_i}} = \frac{8}{\frac{1}{9} + \frac{1}{3} + \frac{1}{8} + \frac{1}{8} + \frac{1}{9} + \frac{1}{8} + \frac{1}{18} + \frac{1}{9}} = 7.3$$

We have an outlier in this data (18). So, the mean is the poor measure of central tendency for this data.

ii. For a={ 12,6,7,3,15,10,18,5 }:

$$range = x_{max} - x_{min} = 18 - 3 = 15$$

Mean deviation

$MD = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}| = \frac{1}{8}(|3 - 9.5| + |5 - 9.5| + |6 - 9.5| + |7 - 9.5| + |10 - 9.5| + |12 - 9.5| +$
$+|15 - 9.5| + +|18 - 9.5|) = 4.25.$

Variance

$V = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{8}(|3 - 9.5|^2 + |5 - 9.5|^2 + |6 - 9.5|^2 + |7 - 9.5|^2 + |10 - 9.5|^2 + |12 - 9.5|^2 +$
$|15 - 9.5|^2 + +|18 - 9.5|^2) = 23.75.$

Standard deviation

$SD = \sqrt{V} = \sqrt{23.75} = 4.87.$

For b={ 9,3,8,8,9,8,9,18 }

$range = y_{max} - y_{min} = 18 - 3 = 15.$

Mean deviation

$MD = \frac{1}{m}\sum_{i=1}^{m}|y_i - \bar{y}| = \frac{1}{8}(|3 - 9| + |8 - 9| + |8 - 9| + |8 - 9| + |9 - 9| + |9 - 9| + |9 - 9| + +|18 -$
$9|) = 2.25.$

Variance

$$V = \frac{1}{m}\sum_{i=1}^{m}(y_i - \bar{y})^2 =$$
$$= \frac{1}{8}(|3 - 9|^2 + |8 - 9|^2 + |8 - 9|^2 + |8 - 9|^2 + |9 - 9|^2 + |9 - 9|^2 + |9 - 9|^2 + |18 - 9|^2) = 15.$$

Standard deviation

$SD = \sqrt{15} = 3.87.$

The range is a poor measure of dispersion, because we have an outlier in this data (18).

iii. For a={ 12,6,7,3,15,10,18,5 }:

$$\text{skewness} = \frac{1}{V^{3/2}} \cdot \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^3 = \frac{1}{23.75^{\frac{3}{2}}} \cdot \frac{1}{8} \left((3-9.5)^3 + (5-9.5)^3 + (6-9.5)^3 + (7-9.5)^3 + (10-9.5)^3 + (12-9.5)^3 + (15-9.5)^3 + (18-9.5)^3 \right) = 0.40$$

$$kurtosis = \frac{1}{V^2} \cdot \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^4 = \frac{1}{23.75^2} \cdot \frac{1}{8} \left(|3-9.5|^4 + |5-9.5|^4 + |6-9.5|^4 + |7-9.5|^4 + |10-9.5|^4 + |12-9.5|^4 + |15-9.5|^4 + |18-9.5|^4 \right) - 3 = -1.10$$

For this data set, the skewness is 0.40 and the kurtosis is -1.10, which indicates small skewness and kurtosis.

For b={ 9,3,8,8,9,8,9,18 }:

$$\text{skewness} = \frac{1}{V^{3/2}} \cdot \frac{1}{m} \sum_{i=1}^{m} (y_i - \bar{y})^3 = \frac{1}{15^{\frac{3}{2}}} \cdot \frac{1}{8} \left((3-9)^3 + (8-9)^3 + (8-9)^3 + (8-9)^3 + (9-9)^3 + (9-9)^3 + (9-9)^3 + (18-9)^3 \right) = 1.10$$

$$kurtosis = \frac{1}{V^2} \cdot \frac{1}{m} \sum_{i=1}^{m} (y_i - \bar{y})^4 = \frac{1}{15^2} \cdot \frac{1}{8} \left((3-9)^4 + (8-9)^4 + (8-9)^4 + (8-9)^4 + (9-9)^4 + (9-9)^4 + (9-9)^4 + (18-9)^4 \right) - 3 = 1.36$$

For this data set, the skewness is 1.10 and the kurtosis is 1.36, which indicates moderate skewness and kurtosis.