What is the method for finding the mean and standard deviation of grouped data?

Authored by

stats writer

April 26, 2024

The method for finding the mean and standard deviation of grouped data involves first determining the class intervals and frequencies of the data. Then, the midpoints of each class interval are calculated and multiplied by their respective frequencies. The sum of these products is divided by the total frequency to find the mean. To find the standard deviation, the difference between each data point and the mean is squared, multiplied by their respective frequencies, and then added together. The sum is divided by the total frequency and the square root is taken to find the standard deviation. This method takes into account the grouping of data and provides a more accurate measure of central tendency and variability.

Find Mean & Standard Deviation of Grouped Data

Often we may want to calculate the mean and standard deviation of data that is grouped in some way. For example, suppose we have the following grouped data:

Range	Frequency
1-10	2
11-20	7
21-30	10
31-40	3
41-50	1

While it's not possible to calculate the exact mean and standard deviation since we don't know the , it is possible to estimate the mean and standard deviation.

The following steps explain how to do so.

Calculate the Mean of Grouped Data

We can use the following formula to estimate the mean of grouped data:

Mean: Σmini / N

where:

mi: The midpoint of the ith group
ni: The frequency of the ith group

N: The total sample size

Here's how we would apply this formula to our dataset from earlier:

Range	Frequency (n_i)	Midpoint (m_i)	$m_i * n_i$
1-10	2	5.5	11
11-20	7	15.5	108.5
21-30	10	25.5	255
31-40	3	35.5	106.5
41-50	1	45.5	45.5

Mean = (11 + 108.5 + 255 + 106.5 + 45.5) / 23 = **22.89**

The mean of the dataset turns out to be **22.89**.

Note: The for each group can be found by taking the average of the lower and upper value in the range. For example, the midpoint for the first group is calculated as: (1+10) / 2 = 5.5.

Calculate the Standard Deviation of Grouped Data

We can use the following formula to estimate the standard deviation of grouped data:

Standard Deviation: $\sqrt{\Sigma ni(mi-\mu)2 / (N-1)}$

ni: The frequency of the ith group
mi: The midpoint of the ith group
μ: The mean
N: The total sample size

Here's how we would apply this formula to our dataset:

Range	Frequency (n_i)	Midpoint (m_i)	$m_i * n_i$	μ	$m_i - \mu$	$(m_i - \mu)^2$	$n_i(m_i - \mu)^2$
1-10	2	5.5	11	22.89	-17.39	302.41	604.82
11-20	7	15.5	108.5	22.89	-7.39	54.61	382.28
21-30	10	25.5	255	22.89	2.61	6.81	68.12
31-40	3	35.5	106.5	22.89	12.61	159.01	477.04
41-50	1	45.5	45.5	22.89	22.61	511.21	511.21

Standard Deviation = √((604.82 + 382.28 + 68.12 + 477.04 + 511.21) / 22) = **9.6377**

The standard deviation of the dataset turns out to be **9.6377**.