Basic statistics on an array of data points

Physics 258 - DS Hamilton 2004

Enter an array of data to be analyzed:

$X :=$

0.4
0.17
1.16
-1.94

Click on the **Input Table** and expand it to see the array X used in this example. We could also read the data in from an external file.

Number of data points:

$n := \text{length}(X) \qquad n = 100$

$\text{rows}(X) = 100 \qquad \text{last}(X) = 99$

also check out "rows" and "last" in the help menu

Example 1: To find the mean value of the array of data, sum the values of the array and then divide that sum by the number of points n. This follows the definition of the mean of Bevington Eq. 1.1 or Taylor Eq 4.5.

$$\text{sum} := \left| \begin{array}{l} s \leftarrow 0 \\ \text{for } i \in 0..n-1 \\ \quad s \leftarrow s + X_i \end{array} \right.$$

Here we use a simple FOR loop to sum the values of the array from i=0 to n-1.

$\text{sum} = 6.2432$

$\text{meanval} := \frac{\text{sum}}{n} \qquad \text{meanval} = 0.0624$

Or we can use a Mathcad built-in function called "mean".

$\text{mean}(X) = 0.0624$

Mathcad also has several built-in summation operators.

$\frac{1}{n} \cdot \sum X = 0.0624 \qquad \frac{1}{n} \cdot \sum_{i=0}^{n-1} X_i = 0.0624$

Note the range of the index i.

Example 2: To determine the root mean square (RMS) deviation of the data set from the mean value, first sum the squares of the difference between each data point and the mean.

$$\text{sumsq} := \sum_{i=0}^{n-1} \left(X_i - \text{meanval}\right)^2 \qquad \text{sumsq} = 71.47$$

Then divide by n and take the square root

$$\text{rmsdev} := \sqrt{\frac{\text{sumsq}}{n}} \qquad \text{rmsdev} = 0.8454$$

This is also called the *population* standard deviation. There is a built in Mathcad function "stdev" that gives the same numerical result.

$$\text{stdev}(X) = 0.8454$$

The *sample* standard deviation is defined with a n-1 in the denominator by Taylor Eq. 4.9 and Bevington Eq. 1.9.

$$s_x := \sqrt{\frac{\text{sumsq}}{n-1}} \qquad s_x = 0.8497 \qquad \text{Stdev}(X) = 0.8497$$

The standard error (or deviation) of the mean (Taylor Eq 4.14 or Bevington Eq 4.14) is

$$\text{SEM} := \frac{s_x}{\sqrt{n}} \qquad \text{SEM} = 0.0850$$

There is another interesting but seldom used statistic, the "average" deviation (see footnote #2 in Taylor on page 99 or Eq 1.7 in Bevington) which involves summing the absolute values of the deviations.

$$\text{avedev} := \frac{1}{n} \cdot \sum_{i=0}^{n-1} \left| X_i - \text{meanval} \right| \qquad \text{avedev} = 0.689$$