## variance ( $\sigma^2$ , $s^2$ , var(X))

A measure of *dispersion* that is equal to the mean squared distance of a set of *data points* from their *mean*:

$$\sigma^2 = \frac{1}{n} \sum (X - \mu)^2$$

Distances are squared so that positive and negative values cannot cancel out each other. For example, the variance of the data set

$$x = \{1, 2, 3, 4, 5, 6\}$$

with mean  $\mu = 3.5$  equals

$$\frac{(1-3.5)^2 + \dots + (6-3.5)^2}{6} = 2.91\overline{6}$$

The notation  $\sigma^2$  indicates the *population* variance, while  $s^2$  denotes the *sample* variance. The sample variance is defined as:

$$s^2 = \frac{1}{n-1} \sum (x - \bar{x})^2$$

where the use of the factor of  $\frac{1}{n-1}$  instead of  $\frac{1}{n}$  is known as *Bessel's correction*. A more intuitive measure of dispersion is the *standard deviation*. *Probability distributions* have closed forms for computing their variance.