

The Mean

The mean is the sum of a set of values, divided by the number of values, i.e.:

$$\begin{aligned}\bar{X} &= (x_1 + x_2 \dots + x_n) / n \\ &= \left(\sum_{i=1}^n x_i \right) / n \\ &= \frac{1}{n} \sum_{i=1}^n x_i\end{aligned}$$

The mean is also known as the average, or sometimes the “center of mass” of a set of numbers. The symbol that conventionally stands for the mean, \bar{X} , is pronounced “X-bar.”

Sometimes a distinction is made between a) the “population” mean, μ , which is assumed to be some underlying “true” mean, or the mean of a set of observations that represents every instance of some phenomenon, or one that is a consequence of a particular physical model, and b) the “sample” mean, \bar{X} , which is the mean obtained using a specific set of observations, or a sample. The formulas for the sample and population means are identical.

In the analysis of spatial data, a *weighted mean* (or *locally weighted mean*) is often used:

$$\bar{X}_* = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i},$$

where the w 's are determined by some function of distance, e.g. $w_i = (1/d_i^2)$, where d_i is the distance between the i -th observation and some reference point, usually the center of the distribution of points, or the intensity or density of some process at each site.

Properties of the Mean

There are properties of the mean that prove useful in practice. First, the sum of the differences between individual observations and the mean is zero, i.e.

$$\sum_{i=1}^n (x_i - \bar{X}) = 0.$$

This follows from some of the rules that the summation operator follows, i.e.

$$\sum_{i=1}^n (x_i - \bar{X}) = \sum_{i=1}^n x_i - (n \times \bar{X}),$$

because $\sum_{i=1}^n (x_i + b) = \sum_{i=1}^n x_i + (n \times b)$. Then, because $\sum_{i=1}^n x_i = (n \times \bar{X})$ by definition,

$$\sum_{i=1}^n x_i - (n \times \bar{X}) = 0, \text{ and this implies } \sum_{i=1}^n (x_i - \bar{X}) = 0.$$

It is also possible to show that

$$\sum_{i=1}^n (x_i - \bar{X})^2 < \sum_{i=1}^n x_i - (x_i - A)^2,$$

for some value A . This implies that the mean is the “balance point” of the set of numbers x_i .