



GIZAPEDIA

gizapedia.hirusta.io

HOW TO CALCULATE THE MODE FOR DATA GROUPED IN INTERVALS

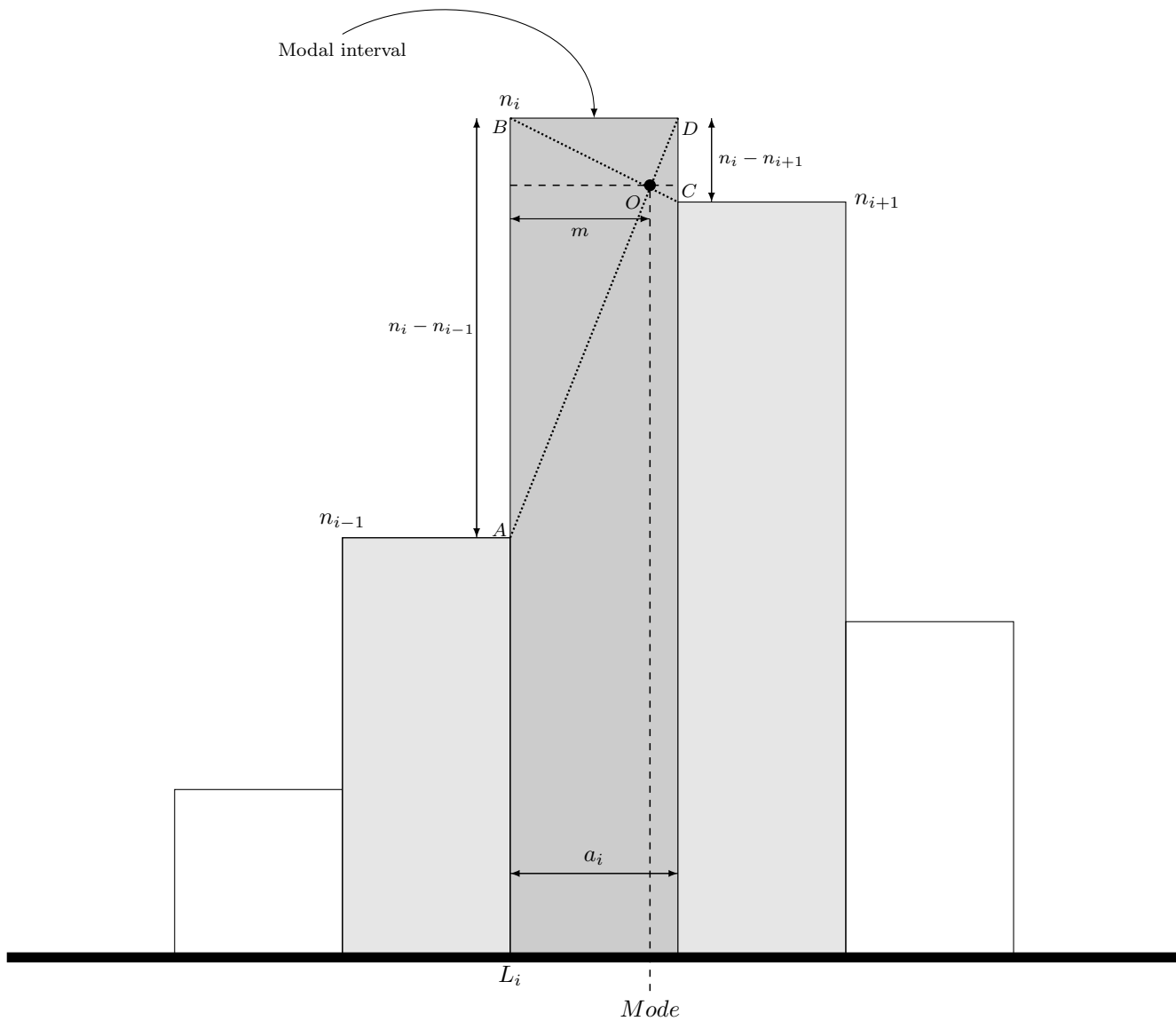
Author: Josemari Sarasola

You can get the document here: <https://gizapedia.hirusta.io/calculating-mode-grouped-data/>
All rights reserved.

How to calculate the mode for grouped data

For grouped data, it's not possible to calculate the mode exactly. The only way to have a value for the mode will an approximation.

The first step is to know the modal interval, that is the interval with the highest frequency. From then on, we get the frequencies of both the adjacent intervals, and we assume that the mode is closer to the limit of the adjacent interval with the highest frequency, in proportion to the difference with the frequency of the modal interval.



So, the approximated value for the mode is $L_i + m$. On the other hand, m is the height of OAB triangle. OAB and OCD are similar triangles, so they have the same base/height parameter:

$$\begin{aligned} \frac{n_i - n_{i-1}}{m} &= \frac{n_i - n_{i+1}}{a_i - m} \\ (a_i - m)(n_i - n_{i-1}) &= m(n_i - n_{i+1}) \\ a_i(n_i - n_{i-1}) &= m \cdot [(n_i - n_{i-1}) + (n_i - n_{i+1})] \\ m &= \frac{(n_i - n_{i-1}) + (n_i - n_{i+1})}{(n_i - n_{i-1})} \end{aligned}$$

$$\boxed{Mo = L_i + \frac{(n_i - n_{i-1}) + (n_i - n_{i+1})}{(n_i - n_{i-1})} a_i}$$

Example

We have got the distribution for weights of 18 years old males in Usurbil, Gipuzkoa, Basque Country (2018):

Weight (kg)	Frequency
50-60	4
60-70	18
70-80	12
80-90	4
90-100	2

- Modal interval 60-70.
- $n_i - n_{i-1} = 18 - 4 = 14$
- $n_i - n_{i+1} = 18 - 12 = 6$

$$Mo = 60 + \frac{14}{14 + 6} \times 10 = 67$$

Remarks

- The result doesn't change if instead the absolute frequencies (n) we get the relative frequencies (f).
- When the width of intervals is different, instead of the frequencies we get the h_i densities (f_i : relative frequency for each interval, a_i : width for each interval):

$$h_i = \frac{f_i}{a_i}$$