Quantitative Tools - Level 1 - Fall 2015
Louis de Charsonville

Lecture 7

Values of dispersion and concentration

1 Values of dispersion

Central values give only a partial view of a set of values. They do not give any information about the statistical dispersion of the values. The dispersion of a variable is how squeezed or stretched is its distribution.

1.1 Interquantiles ranges, Deciles

Definition 1.1. q-Quantile
The q-Quantiles of a variable are the points that cut the distribution in q equal parts. There are $q-1$ q-Quantiles.

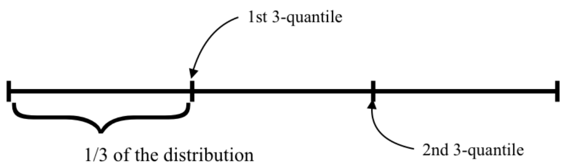

Figure 1: Example of a 3-quantile

The most quantile are :

- the 2-quantile is the point that cuts the distribution of a variable in two. It is known as the **median**.
- the 4-quantiles or **quartiles** are the three points that cut the distribution in 4 equal parts. Usually, we note them as Q_1, Q_2, Q_3 with $Q_1 < Q_2 < Q_3$.
- the 10-quantiles or **deciles** are the nine points that cut the distribution in 10 equal parts. Usually, we note them as D_1, D_2, ... , D_9 with $D_1 < D_2 < ... < D_9$
- the 100-quantiles or **percentiles** or **centiles** are the 99 points that cut the distribution in 100 equals parts.

If Q is even ($Q = 2k$) then the k^{th} quantile is the median.

Definition 1.2. Interquartile range
The interquartile range, or *midspread*, is the difference between the lower and the upper quartiles. Thus :

$$IQR = Q_3 - Q_1$$

The **relative interquartile range** equals to the interquartile divided by the unweighted arithmetic mean (or average) :

$$\text{Relative IQR} = \frac{Q_3 - Q_1}{\bar{X}}$$

The **midhinge** is the average of the lower and upper quartile. The midhinge is usually different from the median.

Definition 1.3. Interdecile range
The interdecile range is the difference between the lower and the upper decile. Thus :

$$IDR = D_9 - D_1$$

The **relative interdecile range** equals to the interdecile divided by the average :

$$\text{Relative IDR} = \frac{D_9 - D_1}{\bar{X}}$$

Definition 1.4. A measure of inequality : the ratio D9/D1
The ratio of the upper decile and the lower decile, that is D_9/D_1, is one of the measure of the inequality of a distribution. It evidences the difference between the top and the bottom of the distribution.

1.2 Absolute deviation

1.2.1 Definitions

Definition 1.5. Absolute deviation
The absolute deviation, or **average absolute deviation**, of a set of values $(x_{i=1}^n)$ is the arithmetic mean of the absolute deviations from the mean[1]. That is :

$$\text{absolute deviation} = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$$

$$\text{absolute deviation} = \sum_{i=1}^{n} \alpha_i |x_i - \bar{x}| \qquad \text{with } \alpha_i \text{ the weight of } x_i$$

1.3 Standard deviation and variance

The standard deviation [2] is the most common value of the dispersion of a variable. It is usually referred as σ.

Definition 1.6. Standard deviation
The standard deviation of a set of values $(x_{i=1}^n)$ is the squared root of the average of the squares of the deviation from the mean. Thus :

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\sigma_x = \sqrt{\sum_{i=1}^{n}\alpha_i(x_i - \bar{x})^2} \qquad \text{with } \alpha_i \text{ the weight of } x_i$$

Definition 1.7. Coefficient of variation
The coefficient of variation or **relative standard deviation** of a set of values $(x_{i=1}^n)$, usually expressed in percentages, is the ratio of the standard deviation by the mean of the $(x_{i=1}^n)$. Thus

$$\text{Coefficient of variation} = \frac{\sigma}{\bar{x}}$$

Definition 1.8. Variance
The variance is the square of the standard deviation, noted as σ^2. Thus :

$$\text{variance} = \sigma^2$$

1.3.1 Properties

Translation The standard deviation of the $(x_{i=1}^n)$ is the same as the standard deviation of $(x_{i=1}^n + b)$

[1] For a continuous quantitative variable, the mean is usually noted μ
[2] in French : "écart-type"

Product The standard deviation of the $(a * x_{i=1}^n)$, with a a constant real number, equals to a times the standard deviation of the $(x_{i=1}^n)$.

Formally :

$$\sigma_{x+b} = \sigma_x$$
$$\sigma_{ax} = a * \sigma_x$$

2 Values of concentration

2.1 Medial

2.2 Gini coefficient and Lorenz curve

* * *