

# Measures of Central Tendency and Variability

## Central tendency

Definition: the tendency of quantitative data to cluster around some central value. The closeness with which the values surround the central value is commonly quantified using the standard deviation. They are classified as summary statistics

### Measures of Central Tendency

1. Mean: The sum of all measurements divided by the number of observations. Can be used with discrete and continuous data. It is value that is most common
2. Median: The middle value that separates the higher half from the lower half. Mean and median can be compared with each other to determine if the population is of normal distribution or not. Numbers are arranged in either ascending or descending order. The middle number is then taken.
3. Mode: The most frequent value. It shows most popular option and is the highest bar in histogram. Example of use: to determine the most common blood group.
4. Geometric mean - the  $n$ th root of the product of the data values.
5. Harmonic mean - the reciprocal of the arithmetic mean of the reciprocals of the data values.
6. Weighted mean - an arithmetic mean that makes use of weighting to certain data elements.
7. Truncated mean - the arithmetic mean of data values that do not include the whole set of values, such as ignoring values after a certain number or discarding a fixed proportion of the highest and lower values.
8. Midrange - the arithmetic mean of the maximum and minimum values of a data set
9.  $\Delta$

## Variability (dispersion)

Definition: *dispersion is contrasted with location or central tendency, and together they are the most used properties of distributions. It is the variability or spread in a variable or a probability distribution i.e. They tell us how much observations in a data set vary.*<sup>[1]</sup> They allow us to summarise our data set with a single value hence giving a more accurate picture of our data set.

### Measures of variability

1. Variance: A measure of how far a set of numbers are spread out from each other. It describes how far the numbers lie from the mean (expected value). It is the square of standard deviation.
2. Standard deviation (SD): it is only used for data that are "normally distributed". *SD indicates how much a set of values is spread around the average.*<sup>[2]</sup> SD is determined by the variance (SD = the root of the variance).
3. Interquartile range (IQR): the interquartile range (IQR), is also known as the 'midspread' or 'middle fifty', *is a measure of statistical dispersion, being equal to the difference between the third and first quartiles*<sup>[3]</sup>.  $IQR = Q3 - Q1$ . Unlike (total) range, the interquartile range is a more commonly used statistic, since it excludes the lower 25% and upper 25%, therefore reflecting more accurately valid values and excluding the outliers.
4. Range: it is the length of the smallest interval which contains all the data and *is calculated by subtracting the smallest observation (sample minimum) from the greatest (sample maximum) and provides an indication of statistical dispersion*<sup>[4]</sup>. It bears the same units as the data used for calculating it. Because of its dependence on just two observations, it tends to be a poor and weak measure of dispersion, with the only exception being when the sample size is large.

## Links

### Related articles

### References

1. Wikipedia.org. *Statistical dispersion* [online]. The last revision 2011-11-08, [cit. 2011-11-28]. <[http://en.wikipedia.org/wiki/Statistical\\_dispersion](http://en.wikipedia.org/wiki/Statistical_dispersion)>.
2. HARRIS, Michael – TAYLOR, Gordon. *Medical and Health Science Statistics Made Easy* [online]. 2nd edition. Jones and Bartlett Publishers, 2009. Available from <Google books (<http://books.google.com/books?id=WqVbhD69WvMC&pg=PA16&ots=3KRluMt2GJ&dq=%22indicates%20how%20much%20a%20set%20of%20values%20is%20spread%20around%20the%20average.%22&hl=cs&pg=PP1#v=onepage&q=%22indicates%20how%20much%20a%20set%20of%20values%20is%20spread%20around%20the%20average.%22&f=false>)>. ISBN 9780763772659.
3. Wikipedia.org. *Interquartile range* [online]. The last revision 2011-10-31, [cit. 2011-11-28]. <[http://en.wikipedia.org/wiki/Interquartile\\_range](http://en.wikipedia.org/wiki/Interquartile_range)>.
4. Wikipedia.org. *Range (statistics)* [online]. The last revision 2011-11-14, [cit. 2011-11-

28]. <[http://en.wikipedia.org/wiki/Range\\_\(statistics\)](http://en.wikipedia.org/wiki/Range_(statistics))>.

## **Bibliography**

- BENCKO CHARLES UNIVERSITY, PRAGUE 2004, 270 P, V, et al. *Hygiene and epidemiology. Selected Chapters*. 2nd edition. Prague. 2008. ISBN 9788024607931.