

SOP: Z-score (=standard score) calculation

Definition:

A **z-score** (aka, a standard **score**) is the number of standard deviations by which the value of an observation or data point differs from the mean value of what is being observed or measured. Observed values above the mean have positive standard scores, while values below the mean have negative standard scores. A **z-score** can be calculated from the following formula. $z = (X - \mu) / \sigma$ where **z** is the **z-score**, **X** is the value of the element, μ is the population mean, and σ is the standard deviation. For normally distributed samples (see fig. 1) a z-score of $-/+1$ means the data point lies outside 84.1 % of all samples in the population, $-/+2$ outside of 97.2 % of data, and $-/+3$ outside 99.9 % of data, and higher z-score correspondingly means the point lies well outside the population.

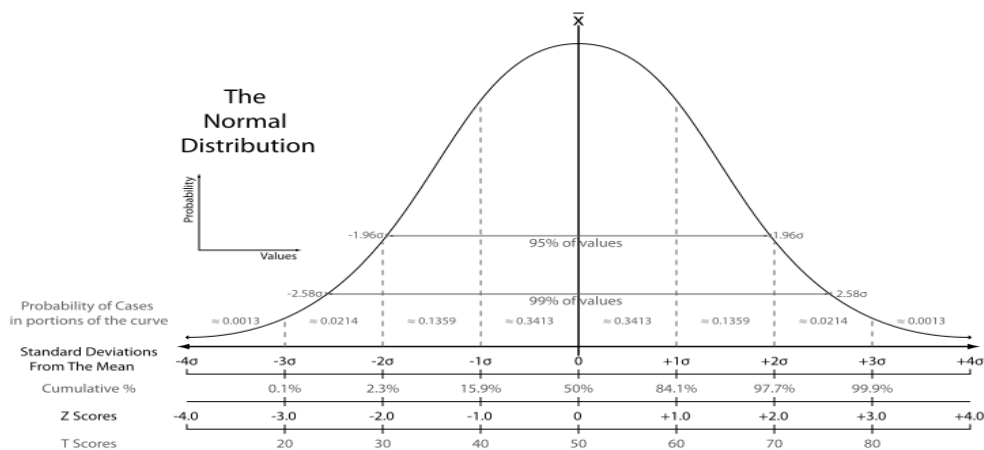


Fig. 1: Range of Z-scores along the normal distribution of data points. Z-scores correspond to the cumulative frequency of samples within the range covered by the z-score

Calculation in Excel:

Apply the following functions in Excel to your spreadsheet of data:

For the population mean: $f(\text{AVERAGE})$ or in German $f(\text{MITTELWERT})$

For the standard deviation: $f(\text{STDEV})$ or in German $f(\text{STABW.S})$

Then calculate the z-score for each data point thusly:

$$Z = (X - \text{AVERAGE}) / \text{STDEV}$$

Graphic representation in GraphPad Prism:

Chose New data table and graph → table format: XY, pick the options “X: Enter X values” and “Y: Enter and plot single Y value for each point” (or “Y: enter replicate values in side-by-side subcolumns” if there are repeat measurements)

Example of application:

-Longitudinal changes of surface marker expression (Fig. 2)

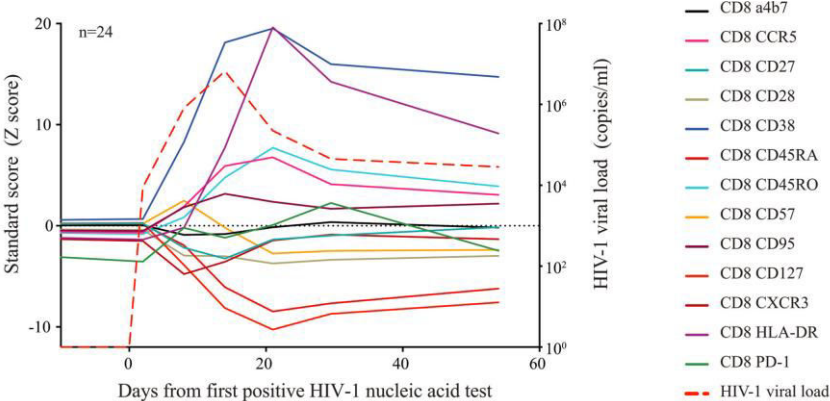


Fig. 2: Example of an application for z-score analysis. From Eller et al. (JVI 2016, 90:8), longitudinal changes of surface maker expression on CD8 T cells during acute HIV infection. Relative changes (Z-score) of surface markers on CD8 T cells during acute HIV infection. Average viral loads are depicted as red-segmented lines from the first RNA-positive HIV nucleic acid test.