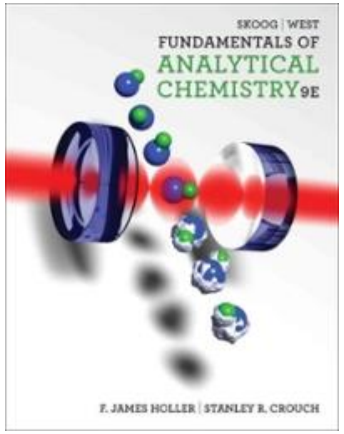
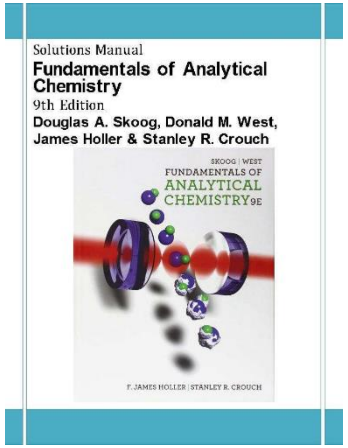
STATISTICS IN CHEMISTRY

1.



2.



1. Skoog DA, West DM, Holler FJ, Crouch SR. Fundamentals of Analytical Chemistry. Nelson Education; 2013.
2. Skoog DA, West DM, Holler FJ, Crouch SR. Solutions Manual of Fundamentals of Analytical Chemistry. Nelson Education; 2013.

Statistical Evaluation of Random Errors

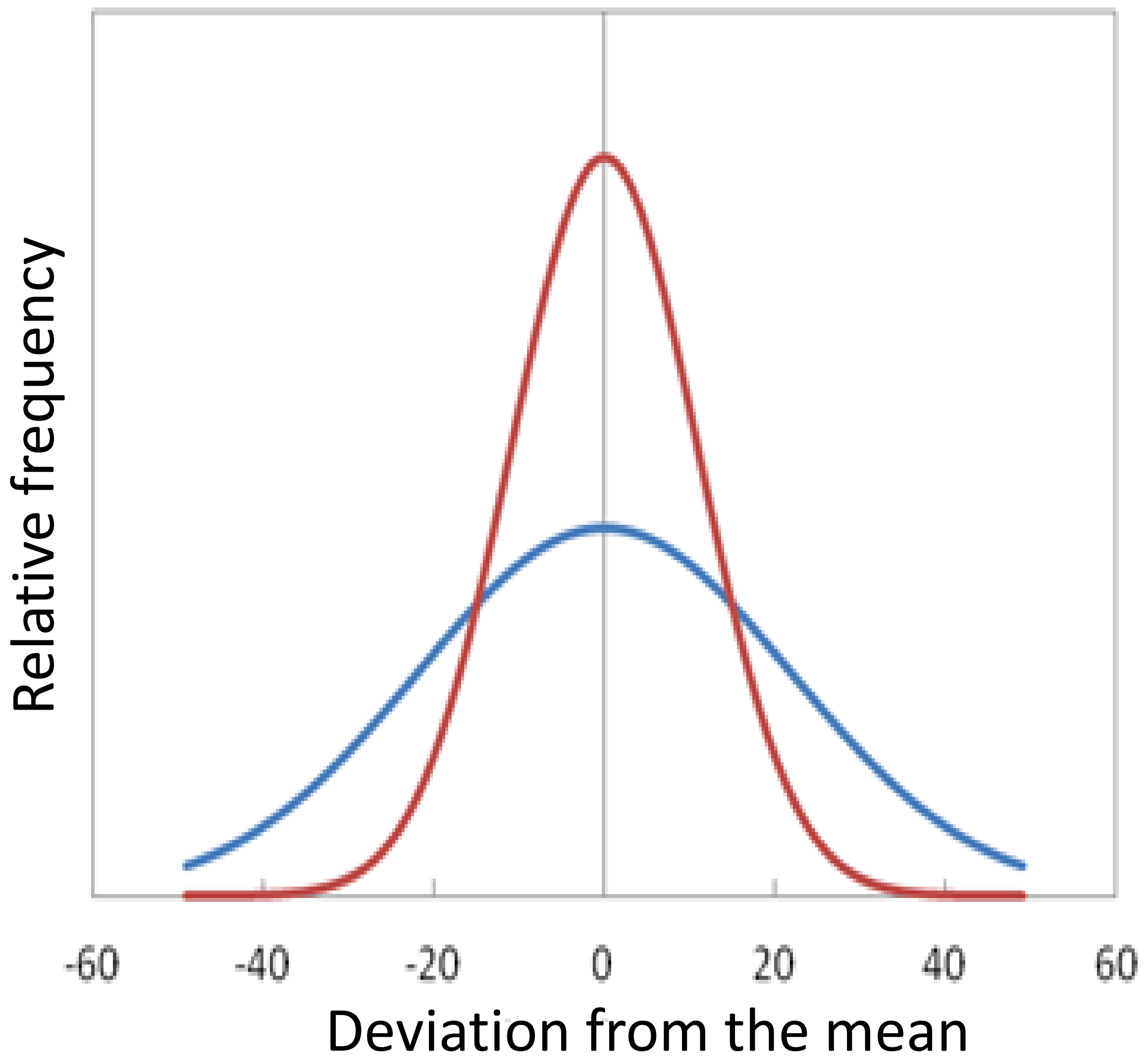
Two important concepts in the statistical evaluation of random errors are sample and population. A population is a collection of entities about which we want to gather information and is also called the universe. Examples of the population; a mining zone, a river, lakes, living tissues, cells, etc. The population can be finite in number and real or conceptual. Smaller subunits representing the population are called samples, and for chemists, the sample is the sample.

Statistical laws are derived for populations. Significant changes should be made when the statistical laws derived for the population are to be applied to small-scale samples.

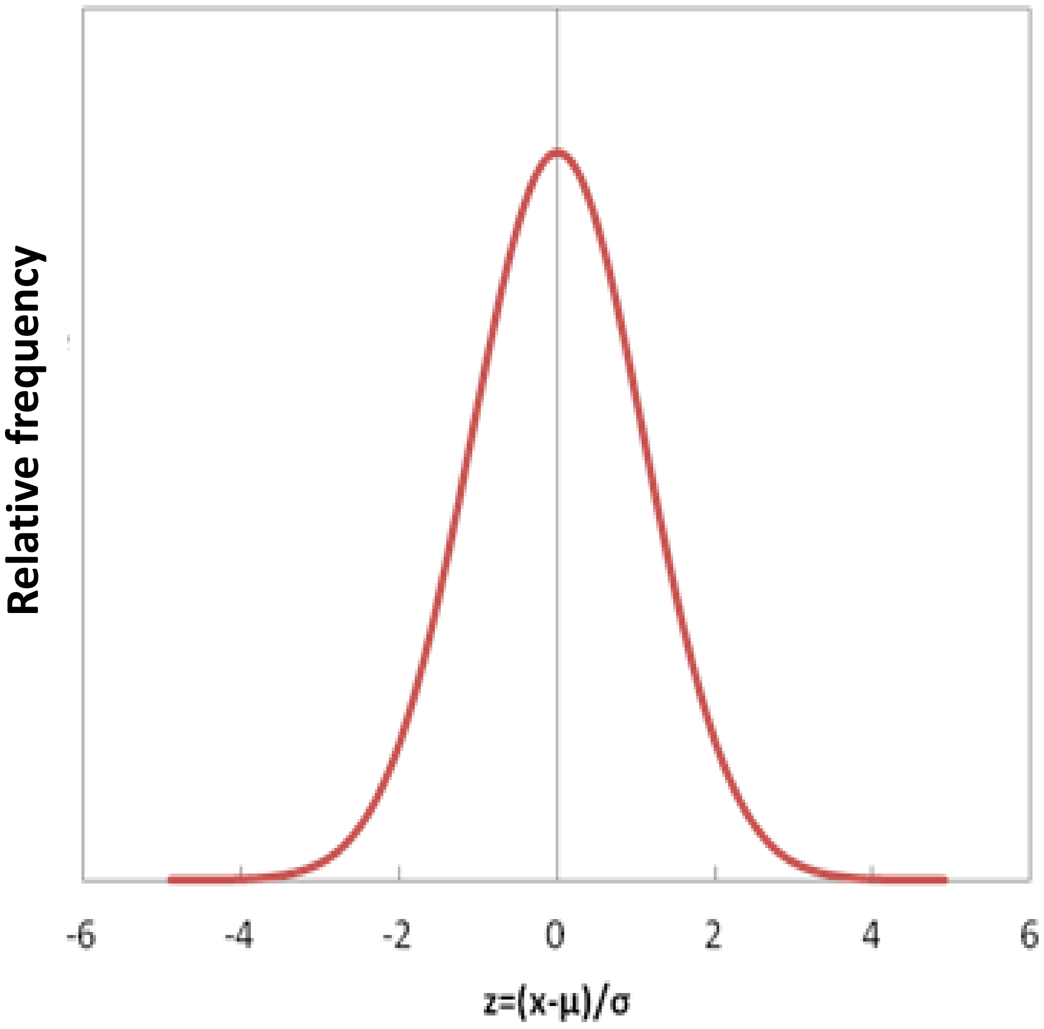
Properties of Gaussian Curve

We have stated that the Gaussian curve is the curve drawn by plotting the relative frequency of the error versus deviation from the mean. The equation representing the Gaussian curve is:

$$y = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$



In this equation (x-μ) is the deviation from the mean. μ represents the population average. In addition, σ in this equation represents population standard deviation. Since the deviation from the mean for each data set will be different, the Gaussian curve plotted against the deviation from the mean for each data set will be different.



However, the Gaussian curve plotted against the value of z, expressed as the deviation from the mean per unit standard deviation, is common to all datasets. The equation that gives the z value is as follows.

Accordingly, Gaussian curve equation can be written as follows:

$$y = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} = \frac{e^{-z^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$

Magnitudes such as μ and σ are called parameters. X is the variable that shows the measurement results. The estimated value of a parameter from a data sample is expressed by the term statistics.

Although the population mean (μ) and sample mean ($\bar{x}$) are calculated with similar equations, they have different meanings. The population mean (μ) refers to the arithmetic mean of all data in a population, while the sample mean ($\bar{x}$) is the arithmetic mean of a limited number of measurements selected from a data population.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \acute{x} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Population standard deviation (σ) is calculated as follows:

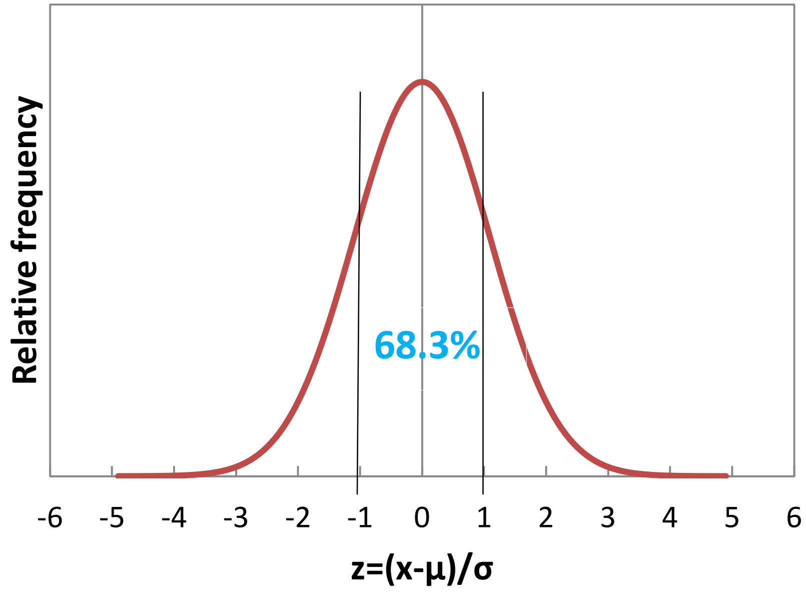
$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

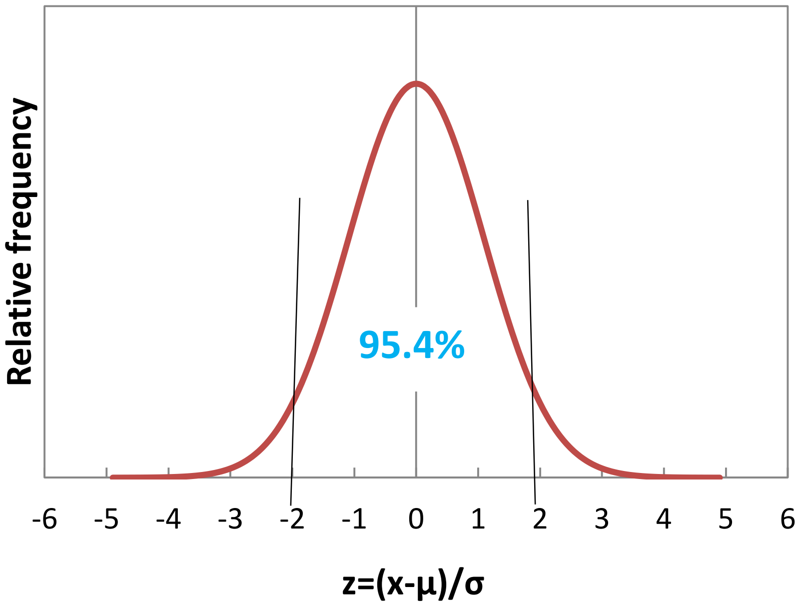
The sample standard deviation (x̅) is calculated as follows:

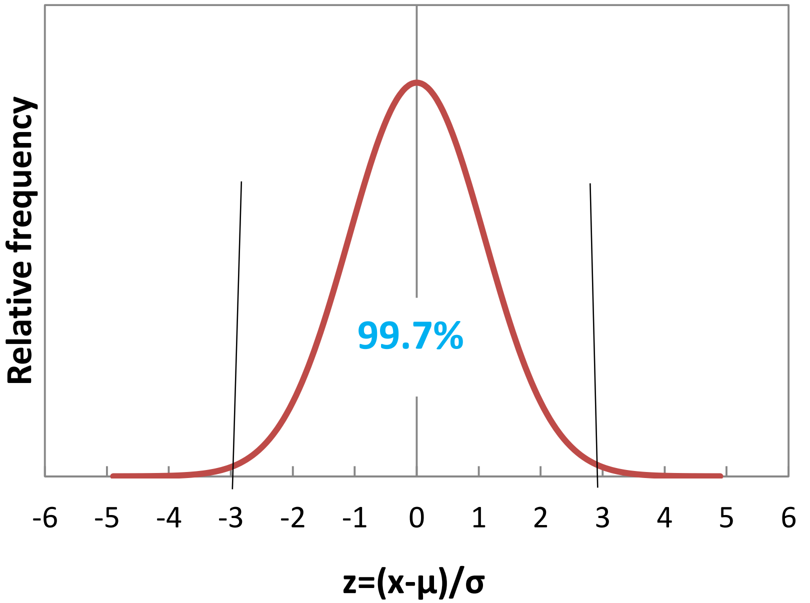
$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}}$$

The expression N-1 in this equation is called the degree of freedom.

Areas under the Gaussian Curve







In the Gaussian curve, the range of -1 to +1 of the z value is 68.3% of the total area. This means that 68.3% of all data is between -1 and +1.

In the Gaussian curve, the range of -2 to +2 of the z value is 95.4% of the total area. This means that 95.4% of all data is between -2 and +2.

In the Gaussian curve, the area of the z value from -3 to +3 is 99.7% of the total area. This means that 99.7% of all data is between -3 and +3.