

ERROR ANALYSIS

Physical Chemistry Laboratory

INTRODUCTION

Every measurement has a degree of uncertainty associated with it. The determination of the degree of uncertainty can be difficult and requires additional effort on the part of the measurer. Nevertheless, evaluation of the uncertainty cannot be neglected *because a measurement of totally unknown reliability is worthless.*

TYPES OF ERROR

1. *Systematic errors* have assignable causes and definite values. They are often unidirectional (cause results to be either too large or too small). Types of these errors include:
 - a) Instrumental uncertainties (Calibration procedures can usually eliminate these.)
 - b) Method uncertainties (nonideal behavior of substances, slowness of reactions, instability of species, etc.)
 - c) Personal uncertainties (personal judgments, bias, and mistakes).
2. *Random errors* or random fluctuations in results occur when replicate experimental data are collected. The specific causes are unknown because they have many sources, none large enough to be identified/detected.

ACCURACY

Individual results in a set of measurement are seldom identical, making it necessary to select a central "best" value. The mean or average of a set of measurements is often used for this purpose.

A measurement of accuracy requires the knowledge of the true value since it denotes the nearness of a measurement to its accepted value and is usually expressed as a comparison with a true or accepted value. Accuracy is usually described in terms of absolute error, the difference between the experimentally determined value and the accepted (sometimes theoretical or true) value:

$$E = \bar{x}_{\text{EXPERIMENT}} - \bar{x}_{\text{T}} \quad (1)$$

PRECISION/REPRODUCIBILITY

Precision is a measurement of the reproducibility of results. It is evaluated by performing replicate experiments under the same conditions. It is defined as an agreement between the numerical values between two or more measurements. Precision can be measured in several ways:

1) Range or spread is the numerical difference between the highest and lowest result:

$$(X_{\text{HIGH VALUE}} - X_{\text{LOW VALUE}}) \quad (2)$$

2) Deviation from the mean is a common method for describing precision and is simply the numerical difference between an experimental value, x_i and the average of the set:

$$|x_i - \bar{x}| \quad (3)$$

3) The standard deviation is a significant measure of precision. For a very large set of data, the standard deviation is given by equation 4:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}} \quad (4)$$

Where $(x_i - \mu)^2$ are the individual deviations from the mean and N represents the number of individual measurements.

A decrease in the number of individual measurements in a set of data decreases the reproducibility and creates a negative bias, or a tendency for the calculated standard deviation to be small as the number of measurements become smaller. Substituting the degrees of freedom (N-1) for N in equation 4 can largely eliminate this bias. Thus, the standard deviation for a small number of measurements is:

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}} \quad (5)$$

Relative Standard Deviation

The use of the relative standard deviation (RSD) can be a more informative way of reporting precision and can be expressed as a percentage or in parts per thousand as shown in equations 7 and 8 (respectively):

$$\text{RSD (\%)} = \text{SD} \times 100\% / \text{Mean} \quad (6)$$

$$\text{RSD (ppt)} = \text{SD} \times 1000 / \text{Mean} = \text{RSD (ppt)} \quad (7)$$

Regression Analysis

When an experimentally measured quantity (y) plotted against (x) approximates a straight line, data seldom fall exactly on that line. Regression analysis refers to the statistical methods used for the generation of such a line and for estimation of the uncertainties associated with its use. The line generated by a least-squares evaluation is the one that minimizes the squares of the individual displacements. Most scientific calculators will calculate the standard deviation of the slope and intercept of the best fit line by inputting the sets of measured x and y values.

REFERENCE

Skoog, Douglas A. and West, Donald M. *Analytical Chemistry: An Introduction*, 4th Ed., Saunders College Publishing, 1985.