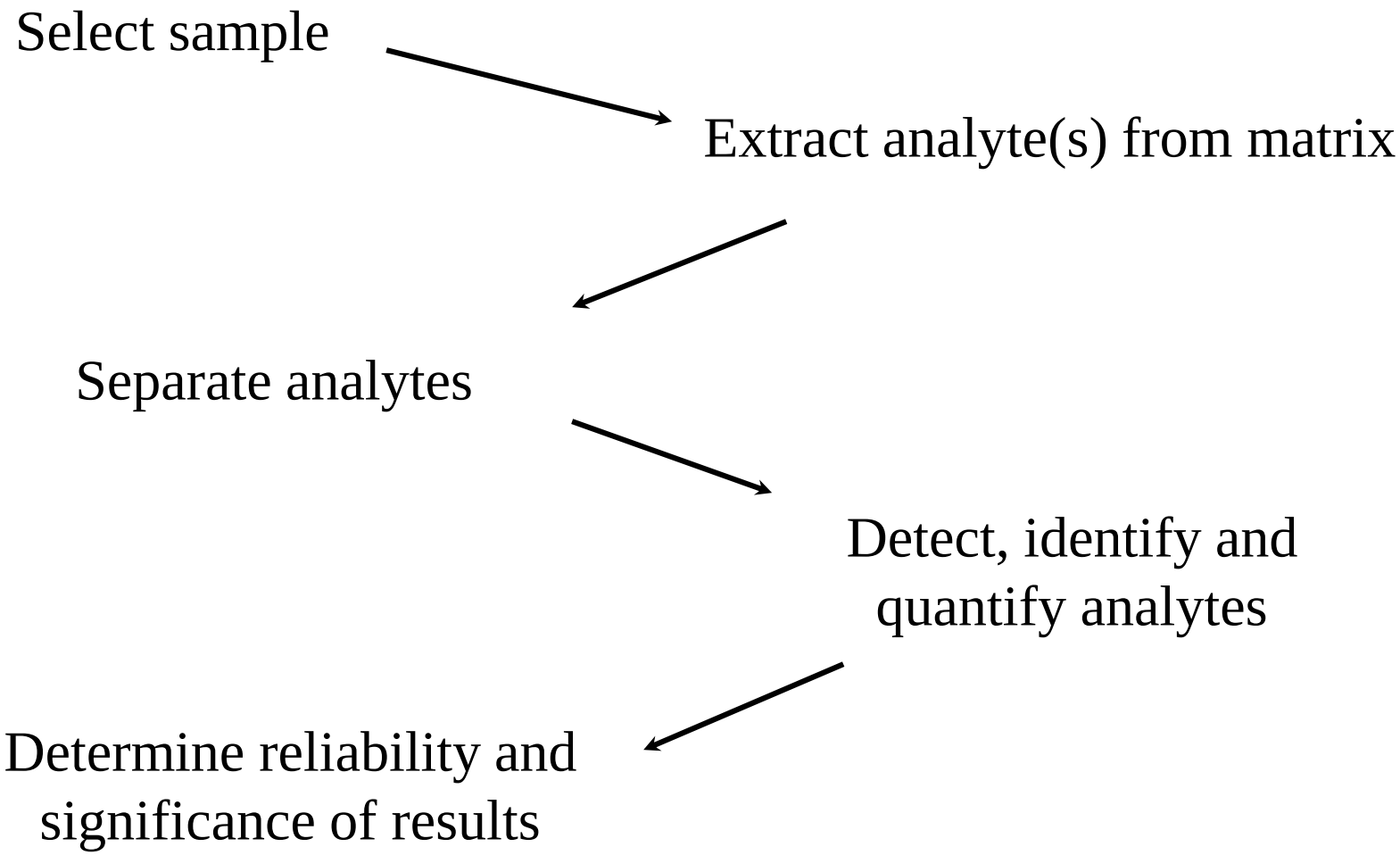
Lecture 4
Statistics for Analytical Chemistry

Dr. Rusul H. Hamza

The General Analytical Problem



Mean

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

also known as average: is a central value of a finite set of numbers.
the sum of the values divided by the number of values.

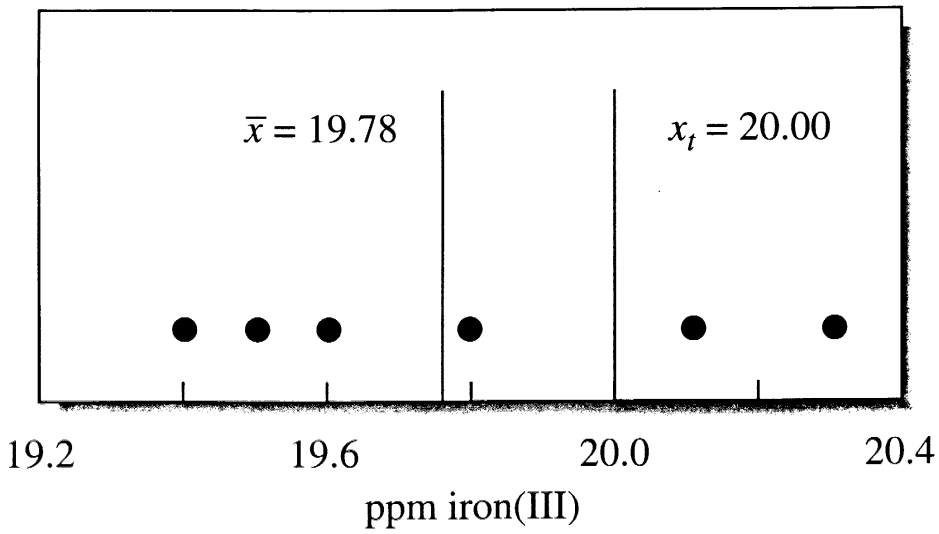
Where x_i = individual values of x
N = number of replicate measurements

Median

the **median** is the value separating the higher half from the lower half of a data sample, a population. For a data set, it may be thought of as "**the middle**" value.

Illustration of "Mean" and "Median"

Results of 6 determinations of the Fe(III) content of a solution, known to contain 20 ppm:



Note: The **mean** value is 19.78 ppm (i.e. **19.8ppm**) - the **median** value is **19.7 ppm**

Precision

Relates to *reproducibility* of results..
How similar are values obtained in exactly the same way?

Useful for measuring this:
Deviation from the mean:

$$d_i = \left| x_i - \bar{x} \right|$$

x_i = individual values
X= mean

Accuracy

Measurement of agreement between experimental mean and true value (which may not be known!).

Measures of accuracy:

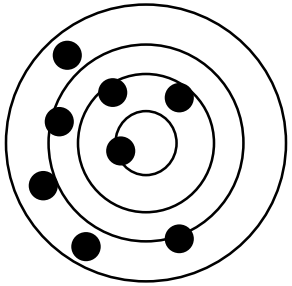
Absolute error: $E = x_i - x_t$

(x_t = true value, x_i =individual values)

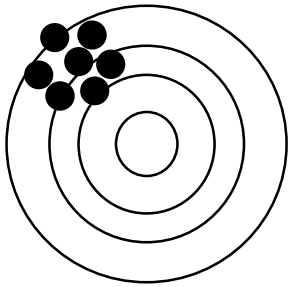
Relative error: $$E_r = \frac{x_i - x_t}{x_t} \times 100\%$$

(latter is more useful in practice)

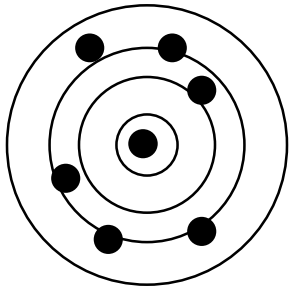
Illustrating the difference between "accuracy" and "precision"



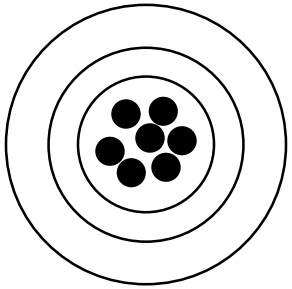
Low accuracy, low precision



Low accuracy, high precision



High accuracy, low precision



High accuracy, high precision

Types of Error in Experimental Data

Three types:

(1) **Random (indeterminate) Error**

Data scattered approx. symmetrically about a mean value. Affects **precision.**

(2) **Systematic (determinate) Error**

Several possible sources. Readings all too high or too low. Affects **accuracy.**

(3) **Gross Errors**

Usually obvious - give "*outlier*" readings.

Sample Standard Deviation, s

the **standard deviation** is a measure of the amount of variation or dispersion of a set of values.

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

Standard Error of a Mean

The standard deviation relates to the probable error in **a single measurement**.

The **standard error of the mean (Sm)**, is defined as follows:

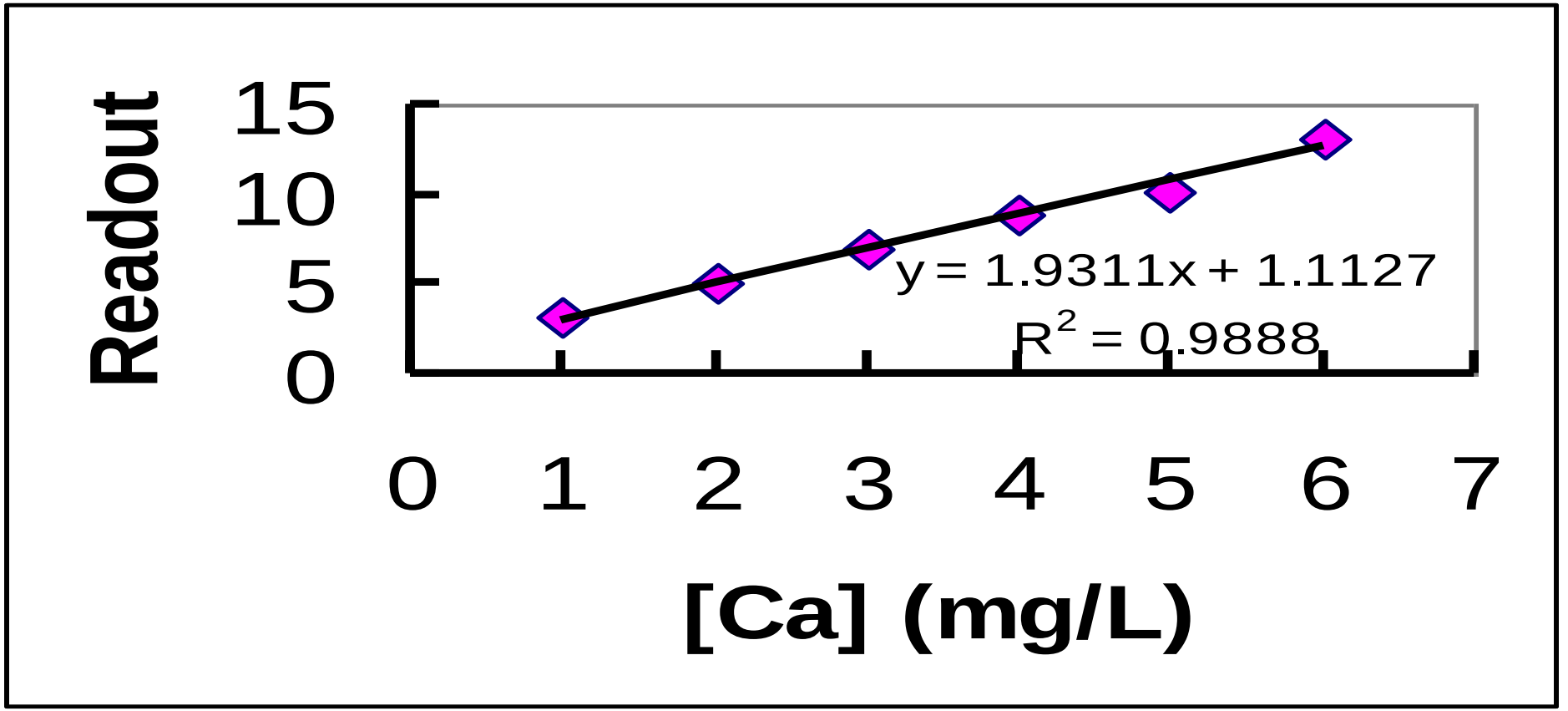
$$s_m = s / \sqrt{N}$$

N= measurements

Standard Curve

Not necessarily linear.

Linear is mathematically easier to deal with.



Textbooks

"Statistics for Analytical Chemistry" J.C. Miller and J.N. Miller, Second Edition, 1992, Ellis Horwood Limited

"Fundamentals of Analytical Chemistry"
Skoog, West and Holler, 7th Ed., 1996

Thank you