CHEM 321: Analytical Chemistry Chapter 3 Notes

I. The Nature of Random Errors - also called "indeterminate" and follow a predictable pattern. Error is the deviation from the "true value" and random error results in values that are higher or lower than the "true value".

A. Mathematics of Random Error - When relative frequency of the error is plotted as a function of the error magnitude, a "bell-shaped or normal" curve is obtained. The plot is also called a Gaussian Curve (named after the mathematician who first described the formula for this curve)

B. Distribution of Experimental Data: has a mean (arithmetic average) and a standard deviation. The median is the middle value when all the values are ranked, while the mode is the most frequent value. Mean = median = mode in a true normal distribution.

II. The Statistical Treatment of Random Error

A. The Sample and the Population - population is an infinite number of observations (all the possible results in the universe!). The sample is a finite number of observations that are, hopefully, representative of the population.

B. Properties of a Gaussian Curve - has a population mean, μ, and a population standard deviation.

1. Population mean. In the absence of systematic error, μ, is the true value for the measurement. The sample mean, *x*, approaches μ and the number of observations approach infinity.

2. Population standard deviation. One population standard deviation contains ±34.15% of the most frequent values, while between 1 and 2 standard deviations contains ±13.6% of the next most frequent values. Thus, 95.5% of all the values are found within ±2 standard deviations in a Gaussian distribution.

We can also say that there are 95.5 chances out of a hundred that the measurement will lie within ±2 standard deviations of the mean value or "true value". The sample standard deviation, s, approaches population standard deviation as the number of data points, n, approaches infinity. The formula is:

3. Standard deviation

$$\text{Standard deviation} = s = \sqrt{\frac{(x_i - \bar{x})^2}{N-1}}$$

where (N - 1) is called the degrees of freedom and i is each individual value from the first value through the last value.

Always carry all digits during a statistical calculation until the end of the calculation, then round. The algorithm (see eq. 3-4 in your text) used by your calculator will give you erroneous results when you have five or more significant digits.

3. Standard error of the mean, s_m. As more measurements are made on the same sample, the mean will vary less from the true value. The standard deviation of this variation or s_m is inversely proportional to $(n)^{1/2}$. *This is only true for measurements made on the same population!*

C. The Reliability of s as a Measure of Precision - the more measurements that are made, the more reliable the value obtained for s. Usually 20 - 30 measurements are necessary.

1. Pooling data to improve the reliability of s. We can "pool" s values obtained for similar samples to obtain a more representative value of s (provided the individual s values are similar!). For replicate measurements made on 3 different samples:

$s_{pooled} = \{[\text{summation } (x_i - x_1)^2 + \text{summation } (x_j - x_2)^2 + \text{summation } (x_k - x_3)^2] \div [n_1 + n_2 + n_3 - 3]\}^{1/2}$, where x_i, x_j, x_k, and n_1, n_2, n_3 are the means and numbers for samples 1, 2 and 3, respectively.

$$s_{pooled} = \sqrt{\frac{\sum_{i=1}^{N_1}(x_i - \bar{x}_1)^2 + \sum_{j=1}^{N_2}(x_j - \bar{x}_2)^2 + \sum_{k=1}^{N_3}(x_k - \bar{x}_3)^2 + \ldots}{N_1 + N_2 + N_3 + \ldots - N_t}}$$

D. Alternative Terms for Expressing the Precision of Samples of Data

1. Variance = s^2. Variances are additive, values of s are not!

2. Relative standard deviation (RSD) & coefficient of variation (CV)

RSD = (s/x), ppt RSD = (s/x) 1000, %RSD = CV = (s/x) 100

3. Spread or range (w) = highest value - lowest value.

III. The Standard Deviation of Computed Results - See Table 3-4

A. The absolute standard deviation of sums and differences:

$s_y = (s_a^2 + s_b^2 + s_c^2)^{½}$, where y = the mean value for the final calculated result

B. The standard deviation of products and quotients:

$s_y/y = [(s_a/x_a)^2 + (s_b/x_b)^2 + (s_c/x_c)^2)]^{1/2}$, where s_y/y is the RSD for the result. This value must be multiplied by y to obtain s_y

C. The standard deviation of exponential calculations:

for $y = a^x$, the RSD is $s_y/y = x(s_a/a)$

D. The standard deviation of logarithms and antilogarithms

for y = log a, $s_y = 0.434(s_a/a)$

for y = antilog a, $s_a/y = 2.303\ s_a$

Examples 3-4 & 3-5

IV. Methods for Reporting Computed Data

A. Rules:

1. Disregard all initial zeros

2. Disregard all final zeros unless they follow a decimal point

3. All remaining digits including zeros between nonzero integers are significant

B. Sums and Differences - the smallest number of digits to the right of the decimal sets the significance

C. Products and Quotients - the smallest number of significant digits determines significance

D. Logarithms - for logs, keep as many digits to the right of the decimal as there are significant figures in the original number

Antilogarithms - keep as many digits as there are digits to the right of the decimal point in the original number

E. Rounding Data:

1. Round up for digits > 5, round down for digits < 5 and round to an even digit when remainder = 5 exactly

2. Use common sense when rounding. Remember that even though 3 significant figures may be permissible for a s value, s is ±term so that:

2.10 ± .0111 becomes 2.10 ± .01

3. Remember not to round off calculations until the final result is obtained!

Examples of significant figures: 3-8 & 3-9