

# Chapter Seven

## *Statistics and Analytical Uncertainty*

### 7. General Remarks About Statistics in Analysis

As is common with all analytical methods, x-ray analysis is imperfectly reproducible, and any measurement must always be associated with some finite uncertainty or as it is commonly referred to, "error". In spite of the sophistication of instrument design and computing algorithms, it is unfortunate that many analysts report data with little or no sophistication or appreciation of analytical uncertainty. Just because the computer can print out a result to six digits does not mean that the digits are all statistically significant. It is the goal of this chapter to provide you with the background and encouragement to report your analyses honestly and with proper regard to uncertainties. Although we will focus on x-ray microanalysis, the fundamentals presented here are equally applicable to any series of measurements in science.

From a statistical point of view, the *reliability* of a series of measurements or a single analysis is a combination of *accuracy* and *precision*. Of these two parameters, accuracy is almost always impossible to quantify and largely remains unknown. In microbeam analysis, accuracy depends on factors such as 1) the absolute accuracy of the standards; 2) the stability and efficiency of the machine, 3) the homogeneity of the standards and specimen; and 4) the validity of all of the correction factors applied to the "raw" data. Reproducibility or precision, on the other hand, can usually be determined, and in some cases even controlled. Precision depends on factors such as 1) the counting statistics; 2) the reproducibility (precision) of the x-ray spectrometers and detectors; and 3) operator "stability" (e.g., for things like proper focus, etc.).

#### 7.1 Statistical Principles

Any act that involves a measurement is imperfectly reproducible; replicate measurements can only serve to define a mean value and some measure of the dispersion of the individual measurements about the mean value. For example, if we repeatedly measure the intensity of an x-ray analytical line, as represented by counts of photons and denoted by  $N$ , on the same sample with constant machine settings, we end up with  $N_1, N_2, N_3, \dots, N_n$  after  $n$  replicate measurements. From these replicate measurements we can calculate the mean intensity:

$$\bar{N} = \frac{\sum_{i=1}^n N_i}{n}$$

eq. 7-1

and the variance:

$$s^2 = \frac{\sum_{i=1}^n (N_i - \bar{N})^2}{(n - 1)}$$

eq. 7-2

The standard deviation<sup>1</sup> is defined as  $s$ . A small number of replicate measurements is only a sampling from the total parent population of all possible replicate measurements that could be made. As the number of individual measurements becomes large, i.e., as  $n > 8$ , we assume that the frequency distribution of individual measurements will begin to correspond more and more closely to a **Gaussian (Normal) Distribution**. This assumption is valid only if the causes of irreproducibility are truly random. This qualification rules out, for example, a normal distribution in the case of a detector with serious incomplete charge collection, replicate x-ray counts taken over a period of time in which the beam current was drifting systematically, gravimetric analyses taken over a period of time in which the bottled chemical reagents were reacting with air, and many more. If such nonrandom sources of error are eliminated, the frequency distribution of a large number of replicate measurements will follow the familiar Gaussian equation.

$$P_N = \left( \frac{1}{\sigma(2\pi)^{1/2}} \right) \exp \left[ \frac{-(N - \mu)^2}{2(\sigma)^2} \right]$$

eq. 7-3

where  $P_N$  is the probability of encountering a value of  $N$  out of the total population,  $\sigma^2$  is the variance of the population,  $\sigma$  is the standard deviation of the population, and  $\mu$  is the mean value of  $N$  in the population. A graph of a normal or Gaussian distribution is shown in Figure 7-1. To the extent that our population of replicate measurements actually follows the normal distribution, we can use the standard deviation to place statistical confidence limits around the mean value. This is done by integrating equation (7-3) between any two limits ( $\mu \pm N$ ). For example, as indicated in Figure 7-1 and the accompanying table, there is a 68.27% chance that any single measurement will give a value between  $(\mu - \sigma)$  and  $(\mu + \sigma)$  and a 99.73% probability that any single measurement will yield a result between  $(\mu - 3\sigma)$  and  $(\mu + 3\sigma)$ . Putting things another way, reporting a mean value and an uncertainty of  $\pm 1\sigma$  allows us to speak of that

<sup>1</sup> In these notes we will designate the mean and standard deviation of actual data as  $\bar{N}$  and  $s$  respectively. The equivalent terms for the **theoretical** mean and standard deviation will be designated as  $\mu$  and  $\sigma$  respectively.

measurement with only a 68.27% confidence limit whereas reporting a mean value and an uncertainty of  $\pm 2\sigma$  allows us to speak of that measurement with 95.45% confidence.

<b>N</b>	<b>P<sub>N</sub></b>	<b>P<sub>integral</sub> (<math>\mu \pm N</math>)</b>
$\mu$	0.3989	0.0000
$\mu \pm .2\sigma$	0.3910	0.1586
$\mu \pm .4\sigma$	0.3683	0.3108
$\mu \pm .6\sigma$	0.3332	0.4515
$\mu \pm .8\sigma$	0.2897	0.5763
$\mu \pm 1.0\sigma$	0.2420	0.6827
$\mu \pm 1.2\sigma$	0.1942	0.7699
$\mu \pm 1.4\sigma$	0.1497	0.8385
$\mu \pm 1.6\sigma$	0.1109	0.8904
$\mu \pm 1.8\sigma$	0.0790	0.9281
$\mu \pm 2.0\sigma$	0.0540	0.9545
$\mu \pm 2.2\sigma$	0.0355	0.9722
$\mu \pm 2.4\sigma$	0.0224	0.9836
$\mu \pm 2.6\sigma$	0.0136	0.9907
$\mu \pm 2.8\sigma$	0.0079	0.9949
$\mu \pm 3.0\sigma$	0.0044	0.9973

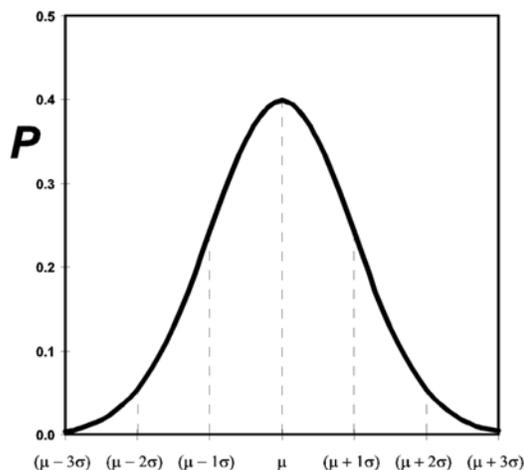


Figure 7-1 Gaussian distribution function.

Clearly, an analytical method that samples a population of measurements with a tight distribution (small  $\sigma$ ) is more precise than one that relies on sampling a population with a broader distribution (larger  $\sigma$ ). To the extent that a measured value can actually be associated with the "truth", sigma is also a measure of accuracy. This statement actually applies to microprobe analyses to a greater extent than some other non-comparative methods, but only to the degree that the composition of the comparison standard we use in probe analysis can be considered as the truth. For certain analytical techniques (e.g., "wet" chemical analysis), the preceding discussion is academic for the most part and of no special interest because replicate measurements are too time-consuming and there is no "truth" against which to standardize.

In the case of analytical methods that make use of x-ray photon counts, statistics have special significance for the following reasons.

- (1) Not only can replicate counting experiments be rapidly and cheaply performed, but it is also true that there is a special relationship between the mean value and the standard deviation. The special relationship is:

$$\mu = \sigma^2 \text{ (or } \bar{N} = \sigma^2 \text{ for a finite sample taken from the entire population)}$$

- (2) Owing to the fact that  $\bar{N} = \sigma^2$  (and  $\sigma = N^{1/2}$ ), it is possible to specify the entire normal distribution by simply controlling just the parameter  $\bar{N}$ . The analyst can, in effect, select the precision (and maybe even the accuracy) of his or her measurements by choosing the value of  $\bar{N}$ . For example, accumulating x-ray counts over a 1 second counting interval may result in an average of 1,000 counts and a precision of  $\pm (1000)^{1/2}$

or  $\pm 3.2\%$  at the .6827 probability confidence limit ( $1\sigma$ ). For the same sample and instrument conditions, counting for 10 seconds would result in an average of 10,000 counts and a precision of  $\pm(10000)^{1/2}$  or  $\pm 1.0\%$  at the .6827 probability confidence limit ( $1\sigma$ ).

Caveat: Strictly speaking, this special relationship holds true only for the variance in observed x-ray counts that is caused by the random production of photons. It does not apply to the variance caused by other random factors in the analytical technique. In other words, the total variance should be written as

$$\sigma^2_{\text{total}} = \sigma^2_{\text{x-ray production}} + \sigma^2_{\text{other random factors}}$$

and the relative precision (in percent) at the one standard deviation level is

$$\frac{(\sigma_{\text{total}})}{\bar{N}} \times 100 = \frac{(\bar{N} + \sigma^2_{\text{other random factors}})^{1/2}}{\bar{N}} \times 100$$

eq. 7-4

Looking at the formula above, one might be tempted to think that  $\bar{N}$  could be increased indefinitely in order to keep improving the precision. Unfortunately, the real world is full of real limitations, and the variance associated with other random factors can also increase as the total counting time increases. More importantly, systematic errors such as instrument instability or specimen damage are more likely to creep into the picture with increased counting time.

Fortunately, x-ray analyses are usually performed on instruments carefully designed and operated so that the "other random factors" are small. If a set of counting experiments are performed on a well-tuned instrument by a well-tuned operator, the distribution of replicate measurements will actually closely approximate a normal distribution with  $\sigma^2 = \bar{N}$  unless  $\bar{N}$  is made very small (<10 counts) or very large (>100000 counts). In any case, ***the variance due to random x-ray photon production is the irreducible minimum variance we can expect from an otherwise perfectly executed set of measurements made on a perfectly functioning machine.***

## 7.2 Explanation for the Special Relationship, $\sigma^2 = \bar{N}$ .

Given the importance we place on equation (7-4), it is prudent to justify the special relationship between the variance and the mean for data collected by microprobe analysis. The normal distribution is a direct result of the fact that the production of x-ray photons is randomly distributed in time. The random production of x-rays is required by quantum physics and is based on the notion that if the atoms in a specimen volume are exposed to a constant flux of electrons (i.e., constant beam current), a fixed fraction of the atoms will always be in an excited state of inner-shell ionization. While in an excited state, the production of x-ray photons can be considered to be a truly random process.

If we perform a replicate set of counting experiments under a perfectly constant electron beam for an interval of time  $\Delta t$ , we will observe that after a moderate number (e.g., 10) of replicate measurements, the mean of the counts observed ( $\bar{N}$ ) changes only insignificantly. If we divide the counting interval ( $\Delta t$ ) into a large number ( $\nu$ ) of much shorter time intervals of duration  $\Delta t/\nu$ , and if the average number of photons produced in  $\Delta t$  is  $\bar{N}$ , the probability of one photon occurring during any interval  $\Delta t(\nu)$  will be given by  $(\bar{N}/\nu)$ . If  $\nu \gg \bar{N}$ , it follows that the probability of more than one photon being produced in the same ( $\Delta t/\nu$ ) interval is very much less than  $(\bar{N}/\nu)$ , and we can ignore this probability.

The question we need to answer is: "*what is the probability of observing  $N$  counts during  $\Delta t$ ?*" In order to address this question we must consider not only the probability of producing 1 photon during any one short time interval ( $\Delta t/\nu$ ), but also the probability of **not** producing 1 photon in any one short time interval. The production probability is given by  $(\bar{N}/\nu)$ , and the non-production probability is  $(1 - \{\bar{N}/\nu\})$ . In other words, during any small time interval we may produce a photon (P) or not produce a photon (NP). The total number of photons produced over the long time interval  $\Delta t$  is the result of many possible combinations of P or NP in the large number of short time intervals ( $\Delta t/\nu$ ). The problem of finding the total number of possible sequences boils down to the familiar one of finding "*how many different ways can we place  $N$  'P-results' and  $(\nu - N)$  'NP-results' into  $\nu$  'time boxes'?*" The **combinatorial formula** can be used to determine the total # of distinct sequences that will yield  $N$  photons.

$$\text{Total number of sequences} = \frac{(\nu)!}{(N)! (\nu - N)!}$$

eq. 7-5

Since each sequence is equally probable, the overall probability of coming up with the result of  $N$  counts after counting for  $\Delta t$  seconds is obtained by multiplying the total number of distinct sequences by the probability of each individual event sequence (P or NP). The overall probability of the result " $N$  counts during  $\Delta t$ " is

$$P_N = \frac{(\nu)!}{(N)! (\nu - N)!} \left(\frac{\bar{N}}{\nu}\right)^N \left(1 - \frac{\bar{N}}{\nu}\right)^{(\nu - N)}$$

eq. 7-6

What we have just derived (Equation 7-6) is known as the **Binomial Probability Distribution Function**. As  $\nu$  becomes large (the case we have been considering) the binomial function can be accurately approximated by the analytically much simpler equation:

$$P_N = \frac{(\bar{N})^N e^{-\bar{N}}}{(N)!}$$

eq. 7-7

Equation (7-7) is known as the **Poisson Probability Distribution function**. The Poisson Distribution function has the special property that its variance is equal to its mean:

$$\sigma^2 = \bar{N}$$

eq. 7-8

Figure 7-2 illustrates a plot of the Poisson Distribution Function of  $P_N$  for the case where  $\bar{N} = 2$ .

$N$	$P_N$	$P_N \times (N - \bar{N})^2$
0	0.135	0.541
1	0.271	0.271
2	0.271	0.000
3	0.180	0.180
4	0.090	0.361
5	0.036	0.325
6	0.012	0.193
7	0.003	0.086
8	0.001	0.031
9	0.000	0.009
10	0.000	0.002
11	0.000	0.001
12	0.000	0.000
		2.0 = $\sigma^2 = \bar{N}$

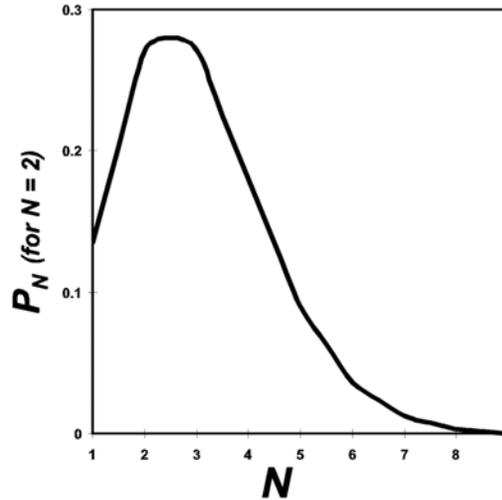


Figure 7-2 Poisson distribution function.

As  $\bar{N}$  becomes large (x-ray counts are always accumulated over a period of time sufficient to gather a large number of counts) the Poisson Distribution function approaches a Gaussian distribution with  $\sigma^2 = \bar{N}$ , i.e.,

$$\frac{(\bar{N})^N e^{-\bar{N}}}{(N)!} \approx \left( \frac{1}{(2\pi\bar{N})^{1/2}} \right) \exp \left[ \frac{-(N - \bar{N})^2}{2\bar{N}} \right]$$

eq. 7-9

Comparison of Figures 7-2 and 7-3 shows that the Poisson distribution certainly does not compare well with a Gaussian distribution for  $\bar{N}$  as small as 2. The two probabilities merge, however, quickly. The comparison is greatly improved at means as low as 10. At  $\bar{N} = 40$  the two are quite close (see Figure 7-3), and at  $\bar{N} > 100$  the two distributions are nearly identical. Clearly, given typical x-ray count rates, we are entirely justified in assuming that the precision of the measurement resulting from the random generation of x-rays in the specimen is governed by statistical factors that allow us to assume that under ideal conditions, the variance ( $\sigma^2$ ) and the standard deviation ( $\sigma$ ) of the analysis can be controlled by the total number of counts accumulated in a single measurement. Furthermore, this special relationship affords us with the opportunity to describe the precision of a single spot analysis or of a number of replicate analyses on a crystal within a given specimen.

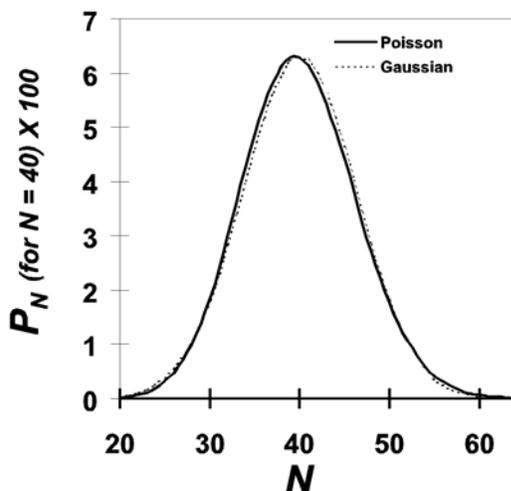


Figure 7-3 Comparison of the *Gaussian* (dashed) and *Poisson* distributions for a sample population characterized by a mean value of 40.

### 7.3 Practical Considerations

All of the statistical principles presented above are based on the fundamental assumption that **all** causes of variation in the number of x-ray counts we record (and ultimately use to calculate elemental or oxide compositions) are completely random in character. This means that if the analyst is to make completely meaningful statements about the precision of his or her analyses, it is necessary to address questions regarding the degree to which a given spot analysis or a set of replicate analyses are truly controlled by random processes. In practice, most of our work does not require us to fully analyze the statistics of our analyses (some special projects may require this, but not "routine" analyses). In general, we simply make every effort to eliminate any non-random contributions associated with the instrument or the operator. Indeed, a major justification for this course is to give you enough understanding of the instruments so you can "tune" them and operate them in a fashion that will reduce or hopefully or at least minimize, non-random errors.

To the extent that we can eliminate non-random errors, the microprobe analyst is in a special position of being able to control the precision of the analysis by selecting  $\bar{N}$ . The table below shows how the lower and upper limits and the relative deviation change with  $\bar{N}$  (all taken at the 0.683 probability or one standard deviation level).

$\bar{N}$	$\bar{N} - (\bar{N})^{1/2}$	$\bar{N} + (\bar{N})^{1/2}$	$\pm 100[(\bar{N})^{1/2} / \bar{N}]$
100	90	110	$\pm 10.0\%$
1,000	968	1,032	$\pm 3.2\%$
10,000	9,900	10,100	$\pm 1.0\%$
100,000	99,684	100,316	$\pm 0.3\%$

If we actually obtain the average from a reasonably large number of replicate counting experiments in which all sources of variance other than x-ray production statistics were negligible,  $\bar{N}$  can be regarded as a true value (or as close to it as we can get). In some applications, it is worth our while to obtain  $\bar{N}$  to see how close the variance is to the irreducible minimum,  $s^2 = \bar{N}$ . If  $s^2$  is significantly larger than  $\bar{N}$ , the machine is not functioning properly or there is some other source of variance associated with the sample itself. The main "other source" of variation is sample homogeneity, a factor that will be discussed in more detail later in this chapter.

In practice, we usually perform one of two types of analyses:

Case I - Multiple Spot Analyses:

Multiple<sup>2</sup> measurements are made, but the number of separate measurements is typically small (e.g., 10-15) relative to that needed to confidently define the entire population of a homogeneous sample. This is perhaps the most common form of microprobe analysis. For example, in one experiment, we may analyze 15 different spots on three different phases. In this case, we actually have two choices of how to estimate the precision of the analysis:

- (1) Calculate the arithmetic mean ( $\bar{N}$ ) from equation (7-1) and assume that the points analyzed are representative of a normal Gaussian population. With this assumption, the variance is then given by  $\sigma^2 = \bar{N}$ , and the standard deviation (at the one-sigma level) is simply given by  $\sigma = \bar{N}^{1/2}$ .
- (2) Calculate the arithmetic mean ( $\bar{N}$ ) from equation (7-1) and the variance from equation (7-2). In other words, we make no assumptions about the measured analyses, or of their relationship to an assumed (i.e., Gaussian) population.

If the specimen being analyzed is perfectly homogeneous, other random errors are minimal and no systematic errors are involved, the two choices should produce identical results. In the analysis of typical materials, if one compares the two alternative methods, the second

---

<sup>2</sup> "Multiple" measurements do not strictly mean "replicate" measurements. The two would be equivalent if the sample were perfectly homogeneous.

option (i.e., using equations 7-1 and 7-2) generally yields a higher variance. In nearly all cases, this can be shown to be due to sample inhomogeneity. Many natural minerals and synthetic materials are chemically zoned, and this zoning may manifest itself by increasing the variance of the mean. Unless there is a need to treat the measurements as individual analyses, we urge you to calculate the mean and standard deviation with equations 7-1 and 7-2. With this approach, the reported mean and standard deviation will constitute a valid representation of the precision and variability of the analyses. If you choose not to report the standard deviation for each oxide (or element) in the analysis, you must make certain that the numbers you do report do not overstate the precision of the analyses.

A general rule of thumb for microprobe analyses is to report weight percentages to three significant digits (including zero). Adoption of this general rule means that the following hypothetical values would be considered valid:  $\text{SiO}_2 = 60.4 \text{ wt.}\%$ ;  $\text{MnO} = 1.23 \text{ wt.}\%$ ;  $\text{ZnO} = 0.45 \text{ wt.}\%$ . In contrast, the following hypothetical values would be considered invalid:  $\text{SiO}_2 = 60.41 \text{ wt.}\%$ ;  $\text{MnO} = 1.234 \text{ wt.}\%$ ;  $\text{ZnO} = 0.451 \text{ wt.}\%$ . The bottom line is that you should either report the actual standard deviation (and state the confidence limit) **or** report only those digits that are statistically significant, given the actual variance.

### Case II - Single Spot Analyses

In many applications, it is necessary to report individual spot analyses. Examples might include a determination of a chemical profile across a zoned mineral grain, or the partitioning of Fe and Mg between the rim of an olivine grain and an adjacent (touching) pyroxene grain. In such cases, we are forced to assume that our single measurement comes from a parent population with a normal distribution ( $\sigma = \bar{N}^{1/2}$ ). It follows that a single measurement  $N$  will fall within the limits  $N = \bar{N} \pm (\bar{N}^{1/2})$  with 0.683 probability. Therefore, when a single measurement is assigned the error limits  $N \pm (\bar{N}^{1/2})$ , it signifies that there is a 0.683 probability that the true value,  $\bar{N}$  will lie within those limits. This assumption is valid as long as the sample is homogeneous.

## **7.4 Sample Homogeneity and Sigma Ratios**

As noted above, a fundamental assumption in both Gaussian and Poisson statistics is that the total population comes from a homogeneous sample. Indeed it is through application of such statistics that we define "homogeneous". In most routine mineralogical and petrological applications we are not greatly concerned with sample homogeneity. In industrial applications, however, sample homogeneity often plays a critical role. Even if single spot analyses are performed on unknown samples, it is customary to analyze multiple spots (Case I) on all standards to which the unknowns will be compared. In the perfect world, all standards would be completely homogeneous. Although homogeneity is a prime criterion by which we judge potential standards, we must admit that many of our standards are slightly inhomogeneous. In order for the x-ray counts obtained on the standards to be representative of the average composition of the material, it is necessary to analyze multiple spots on different grains within the standard mount.

How do we define *sample homogeneity*? Although we can never completely rule out other causes, disparate count rates (or total x-ray counts) on different spots of a single mineral or of different mineral grains, is appropriately attributed to inhomogeneity. A qualitative evaluation of sample homogeneity can be obtained from what is called the *sigma ratio*:

$$\frac{s}{\sigma} \quad \text{or} \quad \frac{s}{\sqrt{N}}$$

eq. 7-10

The sigma ratio compares the standard deviation calculated for a "real" population (cf., equation 7-2) with that of a Gaussian distribution having the same mean. A real population based on a random counting technique cannot produce a standard deviation less than that of the associated Gaussian distribution. Therefore, sigma ratios are always equal to or greater than 1. A sigma ratio of 1 would correspond to a perfectly homogeneous sample (or standard). Generally, a sigma ratio of less than 1.5 is considered to classify a sample as homogeneous. Sigma ratios greater than 3 are not, in our opinion, good for use as standards. The sigma ratio is included in the routine output from our CAMECA software. In terms of analyzing standards, if the sigma ratio is greater than 2, the analyst should be sure to analyze enough points on the standard to have a representative suite of spots to average for the standard's chemical composition. As a general rule of thumb, we like to analyze 8-12 spots on each standard.

## 7.5 Background Measurements and Minimum Detection Limits

So far we have only discussed the statistics associated with counting x-ray photons corresponding to the "peak" of an analytical line of interest. The concept of a "peak" is straightforward as long as the counts obtained on the peak are significantly higher than those corresponding to the background. The analysis of trace elements is complicated by low peak to background ratios, and any statements made concerning the abundance of trace elements must only be made with an understanding of the uncertainties in both the peak counts and the background counts. Examination of the literature containing tables of electron microprobe analyses reveals numerous examples in which trace elements are listed with concentrations of "0.00". If you have learned anything from the above discussion of statistics, it should be apparent that it is fundamentally impossible to state that an element is not present (0.00 wt.%) with complete (100%) certainty. Instead, you should realize that the counting statistics associated with background measurements tell us something about the lower limit of x-ray detection. Put in other words, how few counts can we have in a completely random process, and still say that an element is present (or absent) with a specified degree of confidence? When dealing with analytical chemistry, this concept is generally referred to as the *minimum detection limit*. Although it is customary to report analyses for major elements at the  $\pm 1\sigma$  level (0.683 probability), in order to make a statement about the presence or absence of an element, we ought to be more confident that just 68.3% -- indeed, we should be as confident as possible about such a statement. For these reasons, the accepted standard for defining the minimum detection limit is taken as the 3 sigma level of the background counts. In other words,

$$N_{\text{detected}} = N_{\text{background}} + (3\sigma)_{\text{background}}$$

Use of 3 sigma implies that statements made concerning minimum detection limits are made at the 99.73% confidence level.

As an example, say we want to determine the minimum detection limit for Si in plagioclase feldspar. Tuning an x-ray spectrometer to Si K $\alpha$  and counting for 20 seconds on plagioclase containing 50.0 wt.% SiO<sub>2</sub> produced 87,000 counts. Detuning the spectrometer to either side of the Si peak, produced 100 background counts during 20 seconds. From these data, we determine that  $\sigma = 10$  and  $3\sigma = 30$  counts. 30 counts corresponds to 0.02 wt.% SiO<sub>2</sub>, so the minimum detection limit for SiO<sub>2</sub> is 0.02 wt.%. This amount translates into approximately 0.01 wt.% Si, or about 100 ppm (actually 93 ppm).

As a rule of thumb, regardless of your confidence criteria, to lower the detection limit by a factor of two, you should quadruple the counting time (the relationship is relative to the inverse square).

### 7.5.1 Uncertainties For Off-Peak Corrected Intensities

The discussion above should convince you that the minimum, irreducible uncertainty of a microprobe analysis consists of more than just the counting statistics of the x-ray peak. It should be apparent that the total uncertainty must include the uncertainty of not only the peak counts on the unknown, but also the background measurements and the uncertainty associated with counting the peak for the standard. The total uncertainty must be obtained by *propagating* all of the relevant uncertainties in each component of the analysis.

The standard deviation,  $\sigma$ , is defined as the square root of the variance, which may be a combination of variances,

$$\sigma_{a+b+\dots} = \sqrt{\sigma_a^2 + \sigma_b^2 + \dots}$$

eq. 7-11

Uncertainties are propagated for simple addition or subtraction by adding *absolute* variances.

$$\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2$$

eq. 7-12

and for simple multiplication and division by adding *relative* variances.

$$\left(\frac{\sigma_{xy}}{xy}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$$

eq. 7-13

Note both sigmas above are absolute errors which have been normalized **relative** to x and y respectively. One needs to work with either absolute or relative errors exclusively, i.e., do not mix apples and oranges. The nature of the error propagation requires that both absolute and relative errors be used, depending on the operation. Take note of which type is being used and the conversions from one to the other.

According to Poisson statistics, the absolute variance for counting x-rays is simply the counts obtained over time,  $N$ , and the relative variance is the square of the normalized absolute error:

$$\sigma_{N(abs)}^2 = N, \quad \text{and} \quad \sigma_{N(rel)}^2 = \left(\frac{\sqrt{N}}{N}\right)^2 = \frac{N}{N^2}$$

eq. 7-14

To derive the variance associated with subtracting background counts, we must add the associated absolute variances for the peak and the background measurements,

$$\sigma_{(N-N_b)}^2 = \sigma_N^2 + \sigma_{N_b}^2 = N + N_b$$

eq. 7-15

Equation 7-15 is valid as long as the time spent counting on the peak is the same as that spent counting on the background. A more general form of equation 7-15 that allows for a different counting time for on-peak and off-peak measurements is:

$$\sigma_{(N-N_b)}^2 = N + \left(\frac{t_p}{t_b}\right)N_b$$

eq. 7-16

where the *t-ratio* is that of peak to background counting times. If background is determined by measuring at locations above and below the peak location, for which each background integration time is 1/2 the peak integration time, then as far as the above equation is concerned the ratio  $t_p/t_b$  is one. That is

$$\sigma_{NN_b}^2 = N + N_b I + N_b 2, \quad \text{where} \quad N_b = N_b I + N_b 2$$

eq. 7-17

We now have an absolute error associated with the counting statistics for measuring a single spot on our unknown. In order to relate this error to the calculated weight fraction, i.e., the error associated with the **k-ratio**, the ratio of x-ray counts from the unknown to that of the

standard, we need to incorporate the error associated with sampling the standard. Division is implied so we will need to work with the relative error associated with the unknown, i.e., relative to actual x-ray counts (less background),

$$\left( \frac{\sigma_{(N-N_b)}}{N-N_b} \right)^2 = \frac{N + \left( \frac{t_p}{t_b} \right) N_b}{(N-N_b)^2}$$

eq. 7-18

The error associated with the standard is different in the sense that several points are generally measured and then averaged (i.e., Case I above). Counting statistics should always be considered, but it is more meaningful if the error for the standard is associated with the real mean and variance calculated from equations 7-1 and 7-2 respectively. The standard deviation calculated with equation 7-2 is an absolute error, and we must normalize this standard deviation to N before incorporating it into the overall error calculation (i.e., it must be expressed as a relative error).

We now have all the equations we need to calculate a realistic error for the reported weight fractions of a single unknown analysis. For a given element (all variances being relative), its relative error is given by:

$$\frac{\sigma_{k\text{-ratio}}}{k\text{-ratio}} = \sqrt{\left( \frac{S\bar{N}}{N} \right)_{\text{standard}}^2 + \left( \frac{\sigma_{(N-N_b)}}{N-N_b} \right)_{\text{unknown}}^2}$$

eq. 7-19

You might question why we don't include the error associated with measuring the background on the standard. Two factors justify this omission: 1) Standards are usually selected because they have a relatively high concentration of the element of interest. Therefore the peak/background ratios are high and the uncertainty in measuring the background is not very significant. 2) We typically measure multiple points on the standard and average them. Standard inhomogeneity is usually greater than the variance associated with the background, so we are again justified in ignoring the small background contribution.

Note that the "error" value printed by most software is often an indicator of the minimum detection limit. Rather than actually measuring the minimum detection limit, the analyst can easily examine the relative error for the trace concentration. If, for example, the error associated with a trace element is 50%, then 2-sigma would imply that you are not 95% confident that the element was detected because the weight fraction = 0 is within the limits of uncertainty. In fact, limits of detection are generally stricter than 95% confident and are typically relative to 3-sigma, such that if an error is near 35% it should suggest that the element or oxide is at or below the minimum detection limit.

## 7.5.2 Anticipating your error and practical application of count times

Detection limits should be understood with consideration of the background. When x-ray intensities approach the noise then the amount of noise, or more exactly the error associated with measuring the noise, becomes the primary contribution to the error. As a rule of thumb, regardless of your confidence criteria, to lower the detection limit by a factor of two you should quadruple the counting time (the relationship is relative to the inverse square). Since the software will now take into account the background counting time, give it some consideration as well; i.e., for trace elements, always measure background on both sides of the peak for a significant amount of time — half the peak counting time.

The following formula will prove handy if an analyst has in mind a specific minimum detection level (i.e., a weight fraction which has an error of 35%), or if he/she has in mind a minimum error for a given weight fraction. To a first approximation, the time required for counting the x-rays can be calculated with data gathered from the standards. To a *first approximation* means we will make a few assumptions, the first being the before mentioned standard deviation from standardizing will be insignificant, and that background for the standard is similar to that of the unknown. The equation below also assumes the count time for background will be the same as the peak, which is the case when measuring background above and below the peak (the most accurate method when analyzing trace elements). The time,  $t$  (seconds), required for a given weight fraction,  $X$ , and desired relative error,  $\sigma$ , can be calculated as follows:

$$t = \frac{\frac{X(P - B)}{C} + 2B}{\left(\frac{X(P - B)}{C}\right)^2 \sigma^2}$$

eq. 7-20

where,

$C$  is the weight fraction of the element in the standard,

$P$  is the count rate on the peak uncorrected for background (read from standard's data),

$B$  is the count rate for the background (read from standard's data).

Equation 7-20 calculates the counting time,  $t$ , for a desired sensitivity,  $X$ , and desired error,  $\sigma$ , for most applications to EPMA x-ray counting and measurement sensitivity. We can also make this equation more useful, but now need to redefine a few of the variables. Let's first redefine

$$(P - B) \equiv P_s$$

where the subscript denotes the elemental concentration in the standard and apply the subscript to  $C$  as well. We need further define it as the beam current normalized x-ray intensity minus background. That is, its units are counts/sec/nA. So to make substitution accurate then we need to include the beam current,  $I_b$ , ... and we now have

$$t = \frac{\frac{XI_bP_s}{C_s} + 2B}{\left(\frac{XI_bP_s}{C_s}\right)^2 \sigma^2}$$

eq. 7-21

This equation should be more useful as you can now play *what if* with  $I_b$ . However, regarding the background counts variable,  $B$ , note that, depending on the software, it may not be beam current normalized, so one might need to scale it properly. That is, if you want to use  $I_b$  as 40nA but 'B' was acquired at 20nA, you would need to double  $B$ . Also regarding  $B$  ... to a first approximation you can use the value given for the standard, but it would be more accurate if you substituted a value obtained from your unknown ... especially if the composition of the unknown differed considerably from the standard.

Let's solve the above equation for  $X$  equal to the minimum detection limit,  $X_{mdl}$ . That is, when  $\sigma = 1/3$  and you cannot be 99% confident that zero isn't included in the  $3\sigma$  error range. Substituting ...

$$K = \frac{XI_bP_s}{C} \quad \text{and} \quad \sigma = \frac{1}{3}$$

and wanting to solve for  $K$  via a quadratic equation yields ...

$$K^2 \left(\frac{t}{9}\right) - K - 2B = 0$$

eq. 7-22

where  $a = (t/9)$ ,  $b = (-1)$  and  $c = (-2B)$ . Solving the quadratic for the only solution which makes sense yields ...

$$K = \frac{1 + \sqrt{1 + \frac{8}{9}tB}}{\frac{2}{9}t}$$

eq. 7-23

Finally substituting for  $K$  and solving for  $X_{mdl}$  yields ...

$$X_{mdl} = \left( \frac{C_s}{P_s I_b} \right) \frac{1 + \sqrt{1 + \frac{8}{9} t B}}{\frac{2}{9} t}$$

eq. 7-24

This equation may prove useful. You can play *what if* with it for various beam currents and count times. You can also incorporate it into any EPMA data spreadsheet and calculate minimum detection limit values for elements you measured at trace quantities ... that is, you shouldn't be reporting values measure at or below the MDL. The following analysis worksheet can help you organize your analytical procedure, and help you remember past parameters.

To calculate the counting time for measuring an element at a specific concentration,  $X$ , with a specific error,  $\sigma$ , use this equation:

$$t = \frac{\frac{X I_b P_s}{C_s} + 2B}{\left( \frac{X I_b P_s}{C_s} \right)^2 \sigma^2}$$

To calculate the *minimum detection limit* for a specific analysis via reference to its standardization, use this equation:

$$X_{mdl} = \left( \frac{C_s}{P_s I_b} \right) \frac{1 + \sqrt{1 + \frac{8}{9} t B}}{\frac{2}{9} t}$$

where,

- $P_s$  = peak counts less background counts for the standard (*counts/sec/nAmp*)
- $C_s$  = Concentration of element in standard (*Wt%*)
- $I_B$  = beam current used for analysis (*nAmp*)
- $B$  = background counts measured at beam current  $I_B$  (*counts/sec*).
- $t$  = counting time (*sec*)



## 7.6 Quantifying Heterogeneity: a case study of EPMA and statistics

Electron probe microanalysis (EPMA) is not only an excellent technique for measuring the chemical composition of a wide variety of materials, but also provides an effective means to quantitatively evaluate the chemical homogeneity of the material. In this contribution<sup>3</sup> we present an easy approach for determining the chemical homogeneity of specimens in a statistically meaningful way.

The degree of chemical heterogeneity may be determined by a variety of methods varying both in ease and rigor. One might simply apply a simple test for the sake of classifying the material as *homogeneous* or *heterogeneous*. One very useful method has long been to simply relate the observed standard deviation,  $s_n$ , for many microanalyses,  $n$ , to the statistical counting error, i.e., the square root of the average counts,  $(N_{ave})^{1/2}$ . This simple test is what is usually termed the *sigma ratio*...

$$\frac{s_n}{\sqrt{N_{ave}}}$$

eq. 7-25

A ratio of approximately unity would imply homogeneity, however, this approach does not really tell us anything about heterogeneity, or of the specific relationship of a given sigma ratio to a particular amount of heterogeneity (or degree of homogeneity). For example, what does a sigma ratio of 2 imply for heterogeneity?

Another “yes or no” type of evaluation may be made by comparing all measurements to the range defined by a 3 sigma counting error;

$$\left(N_{ave} - 3\sqrt{N_{ave}}\right) \leq N \leq \left(N_{ave} + 3\sqrt{N_{ave}}\right)$$

eq. 7-26

Utilizing this approach, if all measurements fall within the defined range, the material could be considered homogeneous. As analysts we often require a technique to quantitatively determine homogeneity so that meaningful comparisons can be made. Goldstein<sup>4</sup> makes reference to a *t-distribution* technique that would apparently calculate something more meaningful. That is, it would yield a number which would allow the analyst to say with confidence his sample was *at least* some weight percent heterogeneous. Such a value would be valuable and should be considered a measurement that could be shared amongst other similar

<sup>3</sup> Rice, J., Shaffer, M., Weill, D., course notes for *Electron Beam Analysis in Mineralogy and Petrology*, Department of Geological Sciences, University of Oregon.

<sup>4</sup> Goldstein, J., Newbury, D., Echlin, P., Joy, D., Fiori, C., Lifshin, E., *Scanning Electron Microscopy and X-ray Microanalysis*, Plenum Press, 1984. (p. 432)

measurements. However, what if the analyst wanted to describe his sample as *no more than* as well as *at least* some weight percent heterogeneous. The analyst would also require the technique to be based on statistical principles, to take into account the number of samples,  $n$ , the precision associated with  $N$ , and to be associated with a desired level of *confidence*. We might also ask the technique to separate the compositional variance from other variables, such as counting statistics.

Principle #1: Individual variances can contribute to the observed.

An observed variance can be considered the result of several specific variances summed in quadrature. Thus,

$$\sigma_{observed}^2 = \sigma_a^2 + \sigma_b^2 + \dots$$

eq. 7-27

The above equation relies on the universal acceptance of an *observed variance* being the sum of individual absolute variances ... that is, if not inter-dependent, variances can be individualized. Assuming we have justification for ignoring variances attributable directly to the instrument and/or analyst, we already have a quantitative handle on heterogeneity if we assign one variance to fundamental counting statistics and another to heterogeneity.

Principle #2: The chi-square distribution.

While it is convenient to discuss the ideal case of *a restricted number of "perfectly" executed replicate counting experiments*, in truth we never really know whether or not a set of analyses truly meets this ideal expectation. In practice all we can do is compare the frequency distribution of our sample analyses to the frequency distribution of its assumed parent population, and then try to decide whether the ideal assumption is indeed justified. This is what we do intuitively when we use the *sigma ratio* and state ... "when it is approximately equal to unity, our sample (characterized by  $N_{ave}$  and  $s_n$ ) comes from a parent population characterized by  $\sigma^2$  and  $\sigma$ ".

The *chi-square* function,  $\chi^2 = s_n^2/\sigma^2$ , is a statistical parameter which compares the observed variance of a sample to the variance of its parent population. Furthermore, there is a probability that can be calculated for drawing a sample with variance  $s^2$  from a population with variance  $\sigma^2$ , i.e., a probability of achieving a certain value,  $\chi^2$ . The probability of drawing a sample with  $s_n^2$  variance from a population of variance  $\sigma^2$  [denoted by  $P(\chi^2)$ ] depends on the value of  $\chi^2$  and on the degrees of freedom remaining after  $s_n^2$  has been calculated from the sample. For a case of a sample of  $n$  replicate counts, the degrees of freedom remaining is  $f = (n - 1)$ .

The usefulness of the *chi-square* distribution is its *probability function*. The analyst can ask this function to yield two numbers both associated with a desired confidence level. For the purpose of determining homogeneity, we can, at a given confidence level, evaluate a range defined by lower and upper limits. Functionally, these limits allow the analyst to characterize the specimen's heterogeneity with the terms *at least* or *no more than*, or both.

### Three counting experiments

It is instructive at this point to introduce three actual examples of replicate counting experiments, and introduce some terms:

1. Replicate counts for 100 different EPMA spot analyses on a optically homogeneous, synthetic glass, for which the counting period was relatively short and only ~900 counts were obtained:

$$N_{\text{ave}} = 901, \text{ and } s_n^2 = 993 \text{ and } s_n = 31.5$$

2. Replicate counts for 100 different EPMA spot analyses on a optically homogeneous synthetic, glass, for which the counting period was long enough to obtain ~9000 counts:

$$N_{\text{ave}} = 9005, \text{ and } s_n^2 = 10,609 \text{ and } s_n = 103$$

3. Replicate counts for 100 different EPMA spot analyses on an unknown glass specimen, for which the counting period was long enough to obtain ~9000 counts:

$$N_{\text{ave}} = 8974, \text{ and } s_n^2 = 247,596 \text{ and } s_n = 497$$

A calculation of the *sigma ratio* for each of the 3 cases yields 1.05, 1.08 and 5.25, respectively. Based on the sigma-ratio criterion, the first two glasses appear to be “homogeneous”.

#### 7.6.1 Putting quantitative limits on sample heterogeneity

The third counting experiment, however, suggests heterogeneity. There is practically no chance that our set of observations could have sampled a parent population with  $\sigma_{\text{observed}}^2 \approx N_{\text{ave}}$ . The assumption that it did would run counter to overwhelming odds and we must abandon it. Accordingly, we must describe the sample as being in all probability heterogeneous. We can be more quantitative, however, by utilizing the Principle #1 to describe the sample as having a parent population characterized by

$$\sigma_{\text{observed}}^2 = \sigma_{\text{heterogeneity}}^2 + \sigma_{\text{counting}}^2$$

eq. 7-28

where all units are counts for a given element (the final product being a conversion of *absolute* variances to *relative*, and applying the relative error to absolute weight percent values).

The first step in what we believe to be the correct approach to this problem is to exploit the relationship between the observed population and the parent as predicted by the *chi-square* distribution.

$$\chi^2 = \frac{S_n^2}{\sigma^2} = \left( \frac{S_n^2}{\sigma_{counts}^2 + \sigma_{heterogeneity}^2} \right) = \left( \frac{S_n^2}{N_{ave} + \sigma_{heterogeneity}^2} \right)$$

eq. 7-29

The next step is to decide at what probability (confidence) level we wish to specify the upper and lower limits of heterogeneity, thereby allowing us to state with a defined degree of confidence that “the heterogeneity exceeds  $p$  weight percent, but it does not exceed  $q$  weight percent”. Common confidences are 95% assured or absolutely assured at 99%.

Next, we utilize the integral  $\chi^2$  probability distribution as tabulated in most references on statistics (e.g., the table titled *percentage points, chi-square over degrees of freedom distribution* in the *CRC Handbook of Tables for Probability and Statistics*)<sup>5</sup>. For this example we choose the absolute confidence interval 99%, i.e., the tabulated values for  $P(\chi^2)$  for  $(1 - P) = 1.0$  and  $P = 99.0$ , and  $(100 - 1)$  as the degrees of freedom. Table 1 lists these as well as other comparative values.

$N$	$f$	$\chi^2 @ (1 - P) = 1\%$	$\chi^2 @ P = 99\%$	$\chi^2 @ (1 - P) = 5\%$	$\chi^2 @ P = 95\%$
5	4	.0742	3.3192	.1778	2.3720
10	9	.2320	2.4073	.3694	1.8799
30	29	.4916	1.7099	.6106	1.4675
50	49	.5906	1.5290	.6924	1.3539
100	99	.6993	1.3600	.7782	1.2447

Table 1. Values of  $\chi^2$  for useful examples of  $P$  and  $n$  replicate samplings

The table yields  $\chi^2_{.1} = 0.6993$  (this value sets the lower limit of  $\chi^2$  and therefore the upper limit of  $\sigma_{heterogeneity}^2$ ), and  $\chi^2_{.99} = 1.360$  (which sets the lower limit of  $\sigma_{heterogeneity}^2$ ).

Rearranging eq.7-25 implies

$$\sigma_{[heterogeneity(P)]} = \sqrt{\frac{S_n^2}{\chi^2_{(P)}} - N_{ave}} \quad \text{and} \quad \sigma_{[heterogeneity(1-P)]} = \sqrt{\frac{S_n^2}{\chi^2_{(1-P)}} - N_{ave}}$$

eq. 7-30

Using the data collected our obviously heterogeneous specimen and eq.7-26, we solve for  $\sigma_{heterogeneity}$  where  $\chi^2 = 0.6993$  and  $1.360$ , which yields the absolute counting values of 587 and 416, respectively. If we desire relative values we normalize by the average counts,  $N_{ave}$ , which yields with 99% confidence our sample’s heterogeneity varies at least 4.6%, but no more than 6.5%. These values can be converted into absolute weight percents by applying them to the concentration levels as measured with the electron microprobe.

<sup>5</sup> See also Bevington, P.R., *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, 1969.

## 7.6.2 Summarizing with comparisons

As previously described (cf. eq.7-24), the absolute variance for counting x-rays is simply the counts obtained over time,  $N$ , or the average counts,  $N_{ave}$ , obtained for many points,  $n$ .  $\sigma^2_{observed}$  can be replaced with the square of the observed standard deviation from averaging,  $s_n$ , thus yielding an estimate for the contribution heterogeneity makes to the observed error.

$$\sigma_{heterogeneity} = \sqrt{s_n^2 - N_{ave}}$$

eq. 7-31

For our heterogeneous specimen, a quick calculation yields 488 or 5.44%. Considering its simplicity, this is a reasonable value. It should be noted, however, that this calculation corresponds to confidence level of 68%. If the same sigma is calculated for a confidence level comparable to our *chi-square* evaluation, a value near 15% would result. This discrepancy is related to the fact that this equation does not take into account the large number of replicate samples.

The number of replicate samples ( $n$ ) has a large affect on the *chi-square* calculation. For example if  $n = 5$  ( $f = 4$ ), the values for *chi-square* for the same confidence level are 0.0742 and 3.3192, and the upper and lower values for heterogeneity are 20.3% and 2.9%. This heterogeneity range is clearly of little use compared to the more limited range obtained from a larger number of replicate analyses, and because  $N$  and  $s$  remain similar in spite of  $n$  being large or small. Recognizing that 100 random spot analyses on a single specimen may be considered excessive and not an efficient use of instrument time, consider  $n = 30$ ; the heterogeneity range for our sample (3) based on this number of replicate analyses is 7.8% and 4.1%. Compared with 6.5% and 4.6% (for  $n = 100$ ), and for less than a third of the instrument/operator time, 30 replicate samplings seem like a good compromise.

The counting experiments presented above for optically homogeneous glasses (1 and 2) appear similar in that application of the simplistic *sigma ratio* implies both materials to be “homogeneous”. More rigorous analysis using the technique described above shows that the statistical precision is different and analysis (1) either represents a case for a minor element not being analyzed, or a case where the analyst chose too short a counting time. Using case (1) and a 99% confidence level, this *chi-square* technique suggests the range of heterogeneity to be between 2.5% and “zero”, whereas for case (2) the calculated range is 0.87% and “zero”. That is, for both cases homogeneity is a possibility, but the statistical range is significantly larger for sample (1) than for sample (2). Even though the *sigma ratio* is better for sample (1), statistically it is less homogeneous than sample (2), as a result of the worse precision due to smaller  $N$ . The bottom line is that application of the *chi-square* technique, and other robust statistical techniques will be limited, or misleading for both small  $n$  and small  $N$ . For studies where quantitative information on the homogeneity of samples is required, the analyst must take the time to collect sufficient data for a reasonable result.

Such evaluation of homogeneity usually suffices for major elemental constituents. However, for minor and trace constituents, I need mention the more general case for the

counting error (see eq.7-24), which would include the variance associated with the background measurement, such that

$$\sigma_{counting}^2 = \sigma_{peak}^2 + \sigma_{background}^2$$

eq. 7-32

That is, we had been able to ignore the contribution to the counting error as long as  $N_p \gg N_b$ , but this is no longer true for minor elements. This contribution would need be considered in the equations which follow eq.7-24, but it may take one of several forms because the analyst might use one of several methods of measuring the background more accurately and more appropriately than for every  $N$ , thus minimizing its variance and optimizing the contribution to the accuracy of  $N_{ave}$ . For example, you could reduce the counting error for the background if you measured it very carefully and gathered 10000 counts (1% relative error). Of course, this would be time consuming for every point, but it is easy to justify the background being identical everywhere unless the sample were very inhomogeneous.

## 7.7 Anticipating analytical sensitivity with a spreadsheet

An equation similar to eq. 24 can be formulated especially for a spreadsheet. Since a spreadsheet can be used to graph a relationship, it would be useful to plot *sensitivity*,  $X$ , as a function of time  $\times$  beam current,  $ti$ , where the units are seconds  $\bullet$  nanoAmps. Recall the relationship of x-ray intensity with respect to both time and beam current is linear; therefore we would expect the same x-ray measurement if the count time were 10 seconds for a beam current of 10nA, or 20sec and 5nA ... both of which represent a  $ti$  value of 100. A resulting plot of sensitivity would allow the analyst to make a decision of how sensitivity may approach diminishing returns with respect to specimen damage.

### 7.7.1 Useful variables

*Sensitivity* should always include a connotation of confidence or *error*. A useful variable would be *counting error* (one sigma), and lets further define it as *relative error*. Recall our equation for *relative variance*, eq. 24, here simplified.

$$\sigma_{relative}^2 = \frac{(ti)(N + N_b)}{(ti)^2 (N - N_b)^2}$$

eq. 7-33

where we have made counts for peak ( $N$ ) and background ( $N_b$ ) independent of count time and beam current, and have instead brought  $ti$  aside. The units for  $N$  are therefore *counts per second per nanoamp* (make sure you transcribe, convert to, the proper values when you create your spreadsheet).

If the above x-ray counts represent the measurement of our unknown, let's replace them with the measurement data available from the standard. Recall the *first approximation*

$$\frac{X}{C} \cong \frac{(N - N_b)}{(N - N_b)_0} \quad \text{or} \quad (N - N_b) \cong \frac{X}{C}(N - N_b)_0$$

where  $X$  and  $C$  represent weight fractions (or percent ... or ppm) of the element in the analyte and standard, respectively, and therefore  $X$  will become our variable for representing *sensitivity*. Now, for a bit of a leap, but not too far from our first approximation. Let's assume  $N_b$  is the same for the standard and analyte, which will allow us to substitute the right side of the above equation into eq. 7-33 for  $(N - N_b)$ , and with minor rearranging, we realize

$$\sigma^2 = \frac{\left[ \frac{X}{C}(N - N_b) + 2N_b \right]}{(ti) \left( \frac{X}{C} \right)^2 (N - N_b)^2}$$

Finally, with quite a bit of rearranging and with the help of applying the *quadratic equation* for solving for the only sensible relationship,  $X = f(ti)$ , we finally end up with

$$X = \left( \frac{C}{N - N_b} \right) \frac{\left[ 1 + \sqrt{1 + 8\sigma^2(ti)N_b} \right]}{2\sigma^2(ti)}$$

eq. 7-34

You now have the relationship required to allow a spreadsheet program to either create for you a table or graphically present the relationship of *weight percent sensitivity* as a function of beam and counting conditions. **But wait ... there is more!** The relationship is also a function of *confidence*. That is, what if you need to measure nickel in olivine at 1000ppm, but 1000ppm is not a useful number, unless it has an error associated which is less than 100ppm, or 10%? If you plug 10% (0.10) in for  $\sigma$ , then the relationship becomes  $X \pm 10\%$  as a function of beam and counting conditions. On the other hand, if you need to predict with absolute confidence the detection limit, then plug into  $\sigma$  a value for which predicts the 99% confidence limit,  $1/2.85 = 0.35$ .

Modern spreadsheets will allow you to create a template once you have done this once. What you need is the peak and background values from the microprobe standard you want to use, and don't forget to normalize these values to beam current and time. You will also need to obtain the weight fraction of the element in the standard. If you set your spreadsheet up correctly, and calculate  $X$  for values of  $ti$  (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000), the point of diminishing returns should be made apparent, and changing the relationship as a function of *confidence* should be as easy as changing the  $\sigma$  value.

Another very interesting relationship, formulated just as easily, might be *confidence*, or error, ( $\sigma$ ) as a function of  $X$ , for a given  $ti$ . The graph would be useful if you inserted horizontal lines for  $\sigma = 10\%$  &  $35\%$ , therefore indicating important sensitivities where the relationship intercepts these values for  $\sigma$ .