5

Statistics of Counting

5.1 INTRODUCTION

In this chapter, I will examine the statistical nature of radioactivity counting. Statistics is unavoidably mathematical in nature and many equations will emerge from the discussion. However, only as much general statistical mathematics will be introduced as is necessary to understand the relevant matters. I will go on to discuss the statistical aspects of peak area measurement, background subtraction, choosing optimum counting parameters and the often superficially understood critical limits and minimum detectable activity. I end with an examination of some special counting situations.

At its simplest, radioactivity counting involves a source, a suitable detector for the radiation emitted by the source, a means of counting those decay events that are detected and a timer. If we measure the rate of detection of events, we can directly relate this to the number of radioactive atoms present in the source. The basic premise is that the decay rate of the source (R) is proportional to the number of atoms of radioactive nuclide present (N), the proportionality constant being the **decay constant**, λ. Thus:

$$R = \frac{dN}{dt} = \lambda N \tag{5.1}$$

R is, of course, what would normally be referred to as the **activity** of the sample. In principle, therefore, if we count the number of events, C, detected by the detector in a fixed period of time, Δt, we can estimate the decay rate as follows:

$$R = \frac{C}{\varepsilon \Delta t} \tag{5.2}$$

where ε, in Equation (5.2), is the effective efficiency of counting, taking into account the source–detector geometry, the intrinsic detection efficiency for the particular radiation and the probability of emission of the detected radiation.

While it is true to say that all scientific measurements are estimates of some unattainable true measurement, this is particularly true of radioactivity measurements because of the statistical nature of radioactive decay. Consider a collection of unstable atoms. We can be certain that all will eventually decay. We can expect that at any point in time the rate of decay will be that given by Equation (5.1). However, if we take any particular atom we can never know exactly when it will decay. It follows that we can never know exactly how many atoms will decay within our measurement period. Our measurement can, therefore, only be an *estimate* of the expected decay rate. If we were to make further measurements, these would provide more, slightly different, estimates. This fundamental uncertainty in the quantity we wish to measure, the decay rate, underlies all radioactivity measurements and is in addition to the usual uncertainties (random and systematic) imposed by the measurement process itself.

5.1.1 Statistical statements

At this point, it is appropriate to introduce a number of statistical relationships with which I can describe the distribution of a number of measurements. This section must necessarily be somewhat mathematical. However, textbooks on statistics will cover the theoretical basis of these parameters in much detail, and here I will content myself with a number of simple definitive statements. Later, these will become relevant to an understanding of counting statistics.

Let us assume we have m measurements, x_1, x_2, x_3, ... x_m, each of which is an estimate of some parameter. The nature of the parameter is not important: it might be a voltage, a length or, more relevantly, a number of events within a particular count period. The actual form, that is

Practical Gamma-ray Spectrometry – 2nd Edition Gordon R. Gilmore
 ISBN: 978-0-470-86196-7

shape of the distribution of the measurements, need not concern us at the moment. The distribution will have a value, $E(x)$, which we can expect our measurements to have. Thus:

$$\text{Expected value} = E(x) \tag{5.3}$$

The difference between any particular value, x_i, and the expected value gives some idea of how good an estimate that particular measurement was. Taking the differences for all of the measurements into account would give an idea of the overall uncertainty of the measurements. However, some measurements will be below the expected value and others above; taking a simple sum of the differences is likely to give a result of precisely zero. To get around this, the sum of the square of the differences is used. The resulting factor is called the **variance**, so that:

$$\begin{aligned}\text{var}(x) &= \text{expected value of } [x - E(x)]^2 \\ &\approx E\{[x - E(x)]^2\}\end{aligned} \tag{5.4}$$

Note that the variance is not a function of x but a parameter of the distribution of x. A more convenient factor, which indicates the spread of the values about the $E(x)$, is the **standard deviation**, σ_x. This is simply the square root of the variance:

$$\sigma_x = \sqrt{\text{var}(x)} \tag{5.5}$$

Standard deviation is more meaningful in the sense that it has an obvious relationship to the expected value and the spread of the distribution. Variance will play a large part in this discussion. Variance is additive, standard deviations are not. Calculating the standard deviation relative to the expected value gives the **relative standard deviation**, r_x, sometimes referred to as the **coefficient of variation**, and often expressed as a percentage:

$$r_x = 100\sigma_x / E(x) \tag{5.6}$$

If we have the results of two measurements that we wish to combine, say x and y, then it is a straightforward matter to show that the following relationships hold:

$$E(x+y) = E(x) + E(y) \tag{5.7}$$

$$E(xy) = E(x)E(y) + \text{cov}(x, y) \tag{5.8}$$

The term $\text{cov}(x, y)$ is the **covariance** of x and y and is analogous to the variance:

$$\text{cov}(x, y) = E\{[x - E(x)][y - E(y)]\} \tag{5.9}$$

Covariance is a measure of the interrelation, or correlation, between x and y. When there is no correlation, as is likely to be in all the cases discussed here, then $\text{cov}(x, y) = 0$.

$$\text{var}(x+y) = \text{var}(x) + \text{var}(y) \tag{5.10}$$

$$\text{var}(x-y) = \text{var}(x) + \text{var}(y) \tag{5.11}$$

$$\begin{aligned}\text{var}(xy) \approx{}& E(y)^2\text{var}(x) + E(x)^2\text{var}(y) \\ &+ 2E(x)E(y)\text{cov}(x, y)\end{aligned} \tag{5.12}$$

It can also be shown that, by making the covariance term negative, this relationship also holds for $\text{var}(x/y)$. More usefully, if as we expect $\text{cov}(x, y) = 0$, and using relative standard deviations, we can rearrange Equation (5.12) to:

$$r_{xy}^2 = r_{x/y}^2 = r_x^2 + r_y^2 \tag{5.13}$$

Finally, if k is a constant then:

$$\text{var}(k) = 0 \text{ and } \text{cov}(k, x) = 0 \tag{5.14}$$

$$E(kx) = kE(x) \text{ and } \text{var}(kx) = k^2\text{var}(x) \tag{5.15}$$

$$E(k+x) = k + E(x) \text{ and } \text{var}(k+x) = \text{var}(x) \tag{5.16}$$

These relationships are valid whatever the distribution of our measured values. When we make a radioactive count, our ultimate intention is to estimate the sample activity and a degree of confidence in that estimate of activity. Statistically we can achieve the former aim by identifying the measured count, C, as the expected number of decays, $E(n)$, and relating the confidence limit to the variance $\text{var}(n)$. Thus, in principle:

$$C \Rightarrow n \Rightarrow E(n) \Rightarrow \text{var}(n)$$

Both the expected value and the variance depend upon the form of the relevant statistical distribution and we can now move on to consider the particular case of the distribution of radioactive counts.

5.2 COUNTING DISTRIBUTIONS

5.2.1 The binomial distribution

In principle, the statistics of radioactive decay are binomial in nature. If we were to toss a handful of coins onto a table and then examine the arrangement, we would find coins in one of two dispositions – heads up or tails up. Similarly, if we could prepare a radioactive source and, during a particular period of time, monitor each individual

atom we would see that each has only one of two possible fates – to decay or not decay.

Let us suppose that we could determine exactly which of the atoms, and how many, decayed during the count period. If we were able to repeat the experiment, we would find that different atoms and a different number of atoms decayed in the same period of time. We can regard each such measurement, each count, as a sample in the statistical sense, an attempt to estimate the true decay rate. We would expect the distribution of these counts to fit a **binomial distribution** (sometimes called a **Bernoulli distribution**). This distribution applies because:

- There are two possible states for each atom.
- The probability of an atom decaying during the count period is independent of how often we look.
- The decay of one particular atom does not affect the probability of other atoms decaying.

If we consider each atom in our source there is a certain probability, p, that the atom will decay during the period we choose to make our measurement. This probability is related to the decay constant of the atom and it is straightforward to demonstrate that:

$$p = (1 - e^{-\lambda \Delta t}) \tag{5.17}$$

where Δt is the count period and λ the decay constant. Since there are only two possible outcomes for each atom the probability that the atom will not decay must be $1 - p$. The binomial distribution predicts that, in any particular sample of N atoms the probability of n atoms decaying in a given time, $P(n)$, is:

$$P(n) = \frac{N!}{(N-n)!n!} p^n (1-p)^{N-n} \tag{5.18}$$

So if we have, say, 20 atoms and the probability of decay during the count is 0.1, Equation (5.18) predicts that on 9 occasions out of 100 we would find that 4 atoms decayed. This means that if our detection system were 100 % efficient in detecting decays then we would collect 4 counts on 9 out of 100 occasions. Figure 5.1 shows this probability distribution when the probability, p, is 0.1, 0.5 and 0.9. Unless the probability is close to 0.5, the probability distribution is skewed.

Regardless of the shape of the distribution, the most likely number of decays is given by Equation (5.19):

$$E(n) = pN \tag{5.19}$$

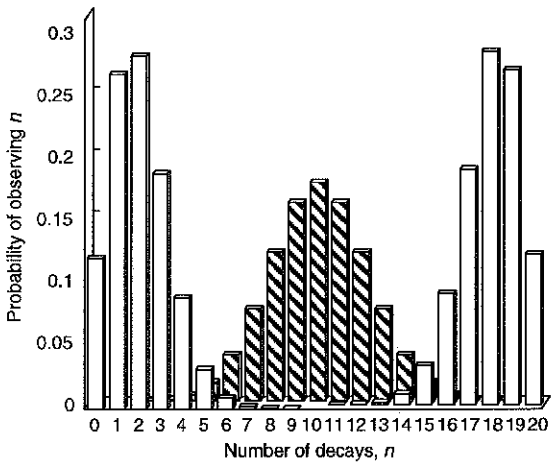

Figure 5.1 Binomial probability distributions for $p = 0.1$ (left), 0.5 (centre) and 0.9 (right)

In the specific cases plotted in Figure 5.1 the most likely counts would be 2, 10 and 18 for the three chosen probabilities. The variance of the distribution is:

$$\text{var}(n) = (1-p)E(n) = (1-p)pN \tag{5.20}$$

Taking the square root of the variance, we can calculate the standard deviation and, for the three specific cases, this would be 1.34, 2.24 and 1.34 decays (or counts, assuming 100 % efficiency). Equation (5.19) is interesting in that it predicts that as the probability becomes very small or very near to 1, the width of the distribution or, we might say, the uncertainty on the number of decays, tends to zero. This is not unreasonable. If $p = 1$, we can expect all atoms to decay and if $p = 0$ none to decay. In either case there is no uncertainty about the number of decays which would be observed.

To relate this to practice, suppose we have counted a sample on a detector with known efficiency, ε, and measured C counts in time Δt s. If the decay constant of the nuclide is known to be λ, then using Equation (5.17), p can be calculated. The overall probability of detection, as opposed to decay, is $p\varepsilon$ and the expected count could be:

$$E(C) = p\varepsilon N \tag{5.21}$$

If we take the measured count C as an estimate of the expected count then Equation (5.1) allows us to calculate the rate of decay, R, as:

$$R = \lambda N = \frac{\lambda C}{(1 - e^{-\lambda \Delta t})\varepsilon} \tag{5.22}$$

In most practical situations, the number of radioactive atoms present is exceedingly high and the probability of detection very small. This means that the number of decays detected (n decays or C counts) is very much smaller than the number of radioactive atoms present (N). (Exceptions to this general situation, when the efficiency of detection and probability of particle emission are very high and when the count period is comparable to the half-life of the nuclide, are discussed in Section 5.7.) In fact, if we assume the detection efficiency to be subsumed into p, it makes no difference to the statistics whether we consider number of decays or number of counts detected and from now on we can take n and C as equivalent. Under these circumstances, various mathematical approximations can be made to Equation (5.18) which lead to a new form for the probability distribution.

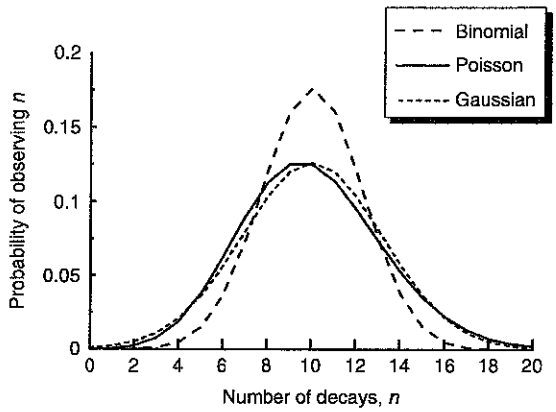

Figure 5.2 Comparison of the binomial, Poisson and Gaussian distributions for $E(n) = 10$

5.2.2 The Poisson and Gaussian distributions

The **Poisson distribution** is used in statistics whenever the total number of possible events, in our situation N, is unknown. The distribution is described by the equation:

$$P(n) = \frac{[E(n)]^n}{n!} e^{-E(n)} \tag{5.23}$$

As before, $P(n)$ is the probability that a count of n will be observed given that the expected count is $E(n)$. This distribution has, as might be expected, some similar properties to the binomial distribution. For example, Equation (5.19) is still valid; however, because $p << 1$, Equation (5.20) approximates to:

$$\text{var}(n) = E(n) \tag{5.24}$$

Curiously, a strict consideration of the mathematics produces the conclusion that if we observe this count n then the expected value, $E(n)$ is:

$$E(n) = n + 1 \tag{5.25}$$

This, at first, surprising statistical fact reminds us that if we were to detect no counts at all, the expected count need not be zero. In most situations, either n is large or is to be corrected for background and it is common practice to ignore this particular statistical fact and take n as a direct estimate of $E(n)$.

Figure 5.2 compares the binomial distribution and the Poisson distribution when both have $E(n) = 10$. The binomial case repeats the data in Figure 5.1 and represents 20 atoms and a probability of decay of 0.5. In the Poisson case, the number of atoms is unknown but large and p is very small. At such a low expected value, there are clear differences. The third distribution shown in Figure 5.2 is the **Gaussian** or **Normal distribution** for the specific case where the variance is equal to the expected value, again 10 counts. This is the distribution one would expect if the differences between the observed and expected counts were solely due to chance. The similarity between the Poisson and Normal distributions is not surprising. When the expected number of counts is greater than 100, then further mathematical approximations can be made to Equation (5.23) which yield the formula for a Gaussian distribution:

$$P(n) = \frac{1}{\sqrt{2\pi E(n)}} \exp\left\{\frac{-[n - E(n)]^2}{2E(n)}\right\} \tag{5.26}$$

To summarize, counting statistics are fundamentally binomial in nature. Under most counting circumstances, we can assume a Poisson distribution of counts. The exceptions to this general rule are:

- when the counting period is long compared to the half-life and the detection efficiency is high;
- when the total number of counts is very small.

These special situations will be discussed in Section 5.7.

5.3 SAMPLING STATISTICS

If we take a large number of measurements of the same parameter, we would find differences in the actual measured value from measurement to measurement. In effect, each measurement is a sample from the infinite number of possible measurements we could make. These

measurements will have a distribution and, of course, an expected value and a variance. If the difference between each measured value and the expected value is due purely to chance, then there is considerable evidence to suggest that the distribution will be Gaussian, often referred to as a Normal distribution in this context. In this case, the equation will have a form similar to the special case in Equation (5.27):

$$P(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-\bar{x})^2}{2\sigma^2}\right] \tag{5.27}$$

that is, the probability of measuring a value x given a particular expected value, $\bar{x}$, and a distribution with a standard deviation of σ (see Figure 5.3 below). Suppose, as I suggested earlier, we have m measurements, x_1, x_2, $x_3, \ldots x_m$. We can define the expected value, or **mean**, $\bar{x}$, of these measurements as:

$$\bar{x} = \frac{\sum x_i}{m} \tag{5.28}$$

where the summation is understood to include all of the measurements x_1 to x_m. The mean is also referred to as the **average**. We can show that as m becomes larger, then $\bar{x}$ becomes a more precise estimate of the expected value. If the true, but unknown, value of the parameter is X, then:

in the limit as m increases: $\bar{x}\ E(x) = X$

It is worth emphasizing this point. The mean, $\bar{x}$, is not the true value of the parameter, only a better, more reliable estimate. The width of the distribution of measured values gives an idea of the overall uncertainty of the measurements. The factor quantifying the width of a distribution is the variance, which is calculated as:

$$\text{var}(x) = \frac{\sum(x_i - \bar{x})^2}{m-1} = s^2 \tag{5.29}$$

where s is an estimated standard deviation, not to be confused with the true standard deviation of the distribution, σ, from which we have taken our sample. The denominator of Equation (5.29), $m-1$, is referred to as the number of **degrees of freedom**.

As with the mean, the more items taken together, the more precise the estimate of the standard deviation:

$$s^2\ \sigma^2, \text{ as } m \text{ increases} \tag{5.30}$$

It is becoming common to refer to **standard uncertainty**, that being the uncertainty on a value at the level of one standard deviation. This may seem an unnecessary addition to the vocabulary but the term does have the advantage of emphasizing that we are dealing with uncertain measurements. It is, perhaps, worth noting that in statistical texts it is more usual to discuss 'standard error' rather than 'standard uncertainty'. I shall keep to the latter usage as being descriptive of the actual situation, reserving the term 'error' for mistakes and the use of incorrect values (see also Section 5.8.1 relating to use of the terms 'accuracy' and 'precision').

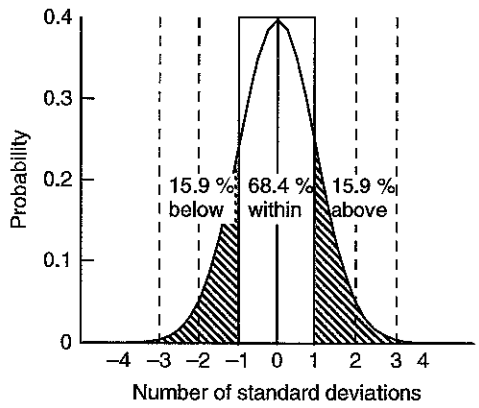

Figure 5.3 A Gaussian distribution with unit standard deviation

5.3.1 Confidence limits

When we quote the result of an experimental measurement, whatever the technique used, it is essential that it is accompanied by a realistic estimate of the uncertainty of the measurement. If we refer again to the Normal distribution of all possible results of a particular measurement, then the uncertainty of the measurement must be related to the width of the distribution. Suppose then that we were to quote our result as, say, $a \pm s$, where s represents one standard uncertainty (a not uncommon procedure). This statement says that the true result (which we can never know) is most likely to be close to a and is less likely to lie below $a-s$ or above $a+s$. We can see from Figure 5.3, where the Normal distribution is plotted with the abscissa scaled in units of one standard deviation, that there is a great deal of scope for the true value to lie outside of these limits and still be 'within' the distribution of results.

To be more certain that our quoted limits encompassed the true value, perhaps we should quote two or three times the standard uncertainty. Whatever limits we choose, we still need to quantify the likelihood of the true value being

outside them. Or, to put it the other way about, we must quote our limits in such a way that we have a stated degree of confidence that the true value lies somewhere within them – hence the term '**confidence limits**'. This degree of confidence is related to the area of the Normal distribution lying within the limits and can be calculated precisely from the parameters of the Normal distribution. The number of times the standard uncertainty we decide to quote to achieve our desired degree of confidence is called the **coverage factor**. Table 5.1 lists the degree of confidence associated with various coverage factors. So, for example, if we wish to be 90 % confident that the true result lies between the quoted limits we might quote the result as follows, with a coverage factor of 1.645:

$a \pm 1.645s$ (90 % confidence)

Table 5.1 Coverage factors and the associated degree of confidence[a]

Coverage factor	Area within confidence limits (%)
1.0	68.3
1.645	90.0
1.96	95.0
2.0	95.5
2.326	98.0
2.576	99.0
3.0	99.9

[a] Confidence limit = coverage factor × s.

The confidence limit quoted in this manner may be referred to as the **expanded uncertainty**. This particular result has confidence limits that are symmetrical about the mean because we have assumed that the distribution of the measurements is Normal. If the distribution were skewed in any way, or perhaps if we were aware that the measurement was possibly, for some reason, biased high (or low), then the lower and upper confidence limits would not be identical.

It is a common practice to quote confidence limits as a percentage of the value rather than as standard deviation. For example, Table 5.2 demonstrates a calculation of weighted mean (which will be explained in due course). If we take the first weighted mean result, we might quote it as:

10.33 (8.03 %, 1σ)

The advantage is that, expressed in this manner, the uncertainty of the result is immediately obvious whereas

Table 5.2 Illustration of weighted mean[a]

Set A

Count	Time (s)	cps	%RSD
102	10	10.2	9.90
53	5	10.6	13.74

Simple mean: 10.40 ± 0.28(2.72 %)
Weighted mean: 10.33 ± 0.83(8.03 %) pooled
or ± 0.19(1.81 %) weighted

Set B

Count	Time (s)	cps	%RSD
1020	100	10.2	3.13
560	50	11.2	4.23

Simple mean: 10.70 ± 0.71(6.61 %)
Weighted mean: 10.51 ± 0.26(2.52 %) pooled
or ± 0.46(4.41 %) weighted

[a] Figures are to be regarded as intermediate, un-rounded, values.

the alternative, 10.33 ± 0.83, needs a degree of mental arithmetic to appreciate whether the result is of good or poor quality.

While discussing the quoting of results it is, perhaps, appropriate to comment on the oft-abused matter of *rounding*. We might well have quoted the result above as $10.333 + 0.829$. Taken at face value, that implies that we are able to determine the confidence limit to one part in 829. That is, not 0.828 nor 0.830, but 0.829. In fact, we were not able to determine the actual value to better than 829 in 10 333 (about one part in 12). In the light of this, is it reasonable to suggest such a high degree of precision for the estimate of the confidence limit? This is a prime example of *spurious accuracy*.

A 'rule-of-thumb' suggested in a code of practice published by the National Physical Laboratory (1973) is as follows:

- Take the confidence limit and round it *always upwards* so as to leave only one significant figure, e.g. 0.829 becomes 0.9.
- Round the result itself *up or down* according to the normal rule to the same degree of precision as the confidence limit, e.g. 10.333 becomes 10.3.

The result above rounded according to these rules would become 10.3 ± 0.9, a more honest statement of what was achieved by the measurement. The recommendation by UKAS, the United Kingdom Accreditation Service (1997), which is consistent with the broader advice in the NPL code, is to quote a 95 % confidence limit (coverage factor 1.96 – often rounded to 2) and to round to two significant figures. Thus, the result above would become 10.3 ± 1.7.

5.3.2 Combining the results from different measurements

Suppose that we have made two measurements of the same parameter and have calculated the uncertainty associated with them. For example, we might have taken measurements on two separate sub-samples of the same radioactive sample and calculated the activity, a_1 and a_2, in Becquerels per gram with confidence limits of s_1 and s_2, which for simplicity we will take as the 68.3 % confidence limit (one standard uncertainty). We will assume that these confidence limits include all sources of uncertainty, not only those due to counting uncertainty:

e.g. $a_1 \pm s_1$ and $a_2 \pm s_2$

Unless the variances of the two results are equal, it is not statistically valid to take a simple mean. This is not unreasonable. A simple mean accords equal importance to each result. A result with a larger variance is less precise and should not be taken as much notice of. The correct procedure is to calculate a weighted mean, $\overline{a}$:

$$\overline{a} = \frac{\sum a_i w_i}{\sum w_i} \tag{5.31}$$

where w_i are weighting factors for each individual result and are simply the reciprocal of the variance of each result. (As usual, the summation is taken to mean the sum over all items.) For example, in the case suggested above: $w_1 = 1/{s_1}^2$ and $w_2 = 1/{s_2}^2$. The standard uncertainty of the combined result is calculated from the **pooled variance**:

$$\mathrm{var}(a)_{\mathrm{internal}} = \frac{1}{\sum w_i} \tag{5.32}$$

Because this calculation takes into account only the individual sample uncertainties, implicitly assuming that the distribution about the mean is satisfactory, this is also known as the **internal variance**. Table 5.2 gives a couple of numerical examples to illustrate the difference between simple and weighted means. In Set A, the simple and weighted means are similar but the simple standard uncertainly does not reflect the fact that both measurements are of poor precision. The weighted mean and pooled standard uncertainty give a much more realistic assessment of the data.

What, however, if the quoted uncertainties do not take into account all sources of uncertainty? In Set B, count times are taken ten times longer. The data is such that the precision of each result is better but the actual results are further apart. In this case, the pooled precision, 2.52 %, is consistent with the precision of the individual results (as it must be!) but does appear to be optimistic taking into account that the difference between the results is nearly 10 %.

Of course, such a large difference could happen by chance, by the statistical roll of the dice, but it is more likely that there are other sources of uncertainty in addition to that due to counting and not accounted for in the uncertainty quoted. We could, of course, simply ignore the uncertainties on the individual values and calculate a simple mean. That, however, would not take into account the relative degrees of reliability of the individual values. In such cases, a standard deviation derived from the **weighted variance** might be quoted, calculated as follows:

$$\mathrm{var}(a)_{\mathrm{external}} = \frac{\sum (a_i - \overline{a})^2 w_i}{\sum w_i (m-1)} \tag{5.33}$$

Because this takes into account the spread of the results about the mean, it is also known as the **external variance**. This is quoted in Table 5.2 as the **weighted uncertainty**. For Set B, the weighted uncertainty of 4.41 % is a more satisfactory estimate of the actual uncertainty than the pooled estimate. In practice, particularly if the work is done by computer, it would make sense to calculate both estimates and quote as the best result the weighted mean together with the larger of the two uncertainty estimates. There is no merit in underestimating uncertainties.

Calculating both has in any case diagnostic value. If experience of a particular measurement scheme shows that the pooled variance is always a significant underestimate of the actual variance, then the measurement process should be looked at in detail to track down the hidden sources of that extra uncertainty.

It should not be lost on us that a single radioactive count has an inherent uncertainty and this should be borne in mind when combining simple count data. A weighted mean should always be used. In fact, because the variance of a count is numerically equal to the count itself, simply combining the count data together will do just that as long as there are no significant sources of uncertainty other than counting uncertainty. As an example, take Set A data from Table 5.2. Simply adding together the counts $(102 + 53 = 155)$ and dividing by the sum of the count times $(10 + 5 = 15)$ provides the weighted mean result of 10.33 cps with an uncertainty (1σ) of 0.83 cps (i.e. $\sqrt{155}/15$), precisely the result shown in the table. Note, though, that applying the same procedure to the data in Set B would give an unsatisfactory result because of the extra, unknown uncertainties.

When calculating weighted means, it is important that the variances used only include those items of uncertainty that are different from measurement to measurement.

Common uncertainties should not be included; otherwise, correlations within the data are introduced.

5.3.3 Propagation of uncertainty

The previous section discussed combining the results of different measurements to obtain a better overall result. We noted that data Set B in Table 5.2 must have undisclosed sources of uncertainty. Let us suppose that it becomes apparent that the preparation of the sources had introduced an extra uncertainty of 6.5 % in the case of the first source and 5.3 % for the second. How can we include the information? The calculation of the uncertainty for each data item, using the example of Set B, is as follows:

- for the count of 1020: $\sqrt{(3.13^2+6.5^2)} = 7.21$ %;
- for the count of 560: $\sqrt{(4.23^2+5.3^2)} = 6.78$ %.

This would provide us with a weighted mean of 10.68 with a pooled uncertainty of 4.94 %, consistent with the actual spread of the data suggested by an external uncertainty of 4.68 %. This is an example of **propagation of uncertainty**. Because the source preparation factor is multiplicative, Equation (5.13) from Section 5.1.1 can be used to combine the uncertainties. The uncertainties are said to have been combined **in quadrature**. (We will meet this again later when discussing the factors that combine to create the width of gamma-ray peaks.)

In our example here, if the source preparation uncertainty were a fixed amount for the method it would be an item common to both sources. It should not, therefore, be included when the uncertainties on the individual results are calculated. It should be taken into account by adding in quadrature to the weighted mean result. If, in our example, the sample preparation uncertainty were 6.5 % for both samples, then the overall uncertainty of the weighted mean for Set B would be $\sqrt{(4.41^2+6.5^2)} = 7.85$ %. The weighted mean value would be unchanged.

In a radioactivity measurement, we may have several sources of uncertainty, all of which must be taken into account in our final uncertainty. For example, we might have:

$$r_T = \sqrt{r_A^2 + r_P^2 + r_S^2 + r_E^2} \tag{5.34}$$

where the various factors are the relative standard deviations of, in order, the total, peak area measurement, source preparation, standard calibration and the efficiency estimate (which would, in turn, include uncertainties due to gamma-ray emission probability and half-life).

Equation (5.13) can only be used in this way when the various factors are multiplied together. If the factors contributing to the overall result are additive, then Equations (5.10) and (5.11) are relevant. For example, assume the result is calculated by an equation which includes additive and multiplicative factors, for example:

$$R = (C - B) \times Y/E$$

The process of combining uncertainties will have to be done in separate stages. In this example, the overall uncertainty of $(C - B)$ must be calculated by using Equation (5.11). This uncertainty, expressed in relative terms, can then be combined in quadrature with the relative uncertainties of Y and E. Combination of uncertainties will be discussed further in Section 5.8, where uncertainty budgets are discussed.

5.4 PEAK AREA MEASUREMENT

In Chapter 3, I explained that a gamma-ray spectrum consists of a large number of 'channels' in each of which are accumulated all of those counts which fall within a small energy range. We might have, for example, a 4096 channel spectrum covering an energy range of 2048 keV, the content of each channel representing the number of counts received within a 0.5 keV energy window. Successive channels represent increasing energy. Within such a spectrum, a gamma-ray appears as a distribution of counts, approximately Gaussian, about a central point which we can take to represent the gamma-ray energy (Figure 5.4). In principle, the actual distribution of counts in a peak is irrelevant; measurement of the peak area should require no more than a simple summation of the number of counts in each of those channels that we consider to be part of the peak and subtraction of an allowance for the background beneath the peak.

The background beneath gamma-ray spectrum peaks can arise from many sources. In most cases, the background will represent the Compton continuum from other gamma-ray interactions within the detector, within the sample itself and from general background radiation interaction with the shielding and the detector. Both background radionuclides and other radionuclides in the sample will contribute to this peak background. Unlike simple counting where total counts are accumulated, the measurement of natural background is of little use in estimating the continuum background beneath a peak. In some cases, those where the radionuclide we wish to measure can be detected in the natural background (^{60}Co, for example), allowance will have to be made for this additional peaked-background over and above the continuum background. This will be considered later, and for the time being we will make the assumption that the continuum beneath the peak is linear.

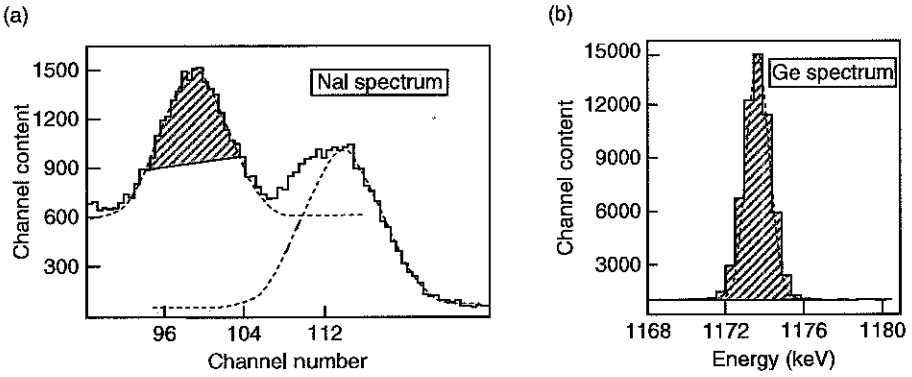

Figure 5.4 Portions of the spectrum of ^{60}Co gamma-rays measured on (a) NaI(Tl) scintillation, and (b) Ge semiconductor detectors (the dotted lines represent the underlying Gaussian distribution of counts)

Over the years, a number of simple algorithms for peak area calculation have been used. The *Covell method* was used in the early days of digital gamma-ray spectrometry for measuring peak areas in sodium iodide scintillation spectra. The procedure was to locate the highest channel in the peak and then to mark the peak limits an equal number of channels away from the centroid channel. When using low-resolution scintillation detectors, peak interference was frequent and it was often necessary to restrict the portion of the peak measured to minimize the effect of neighbouring, possibly overlapping, peaks (see Figure 5.4). The fact that not all of the peak area was taken into account was compensated for by ensuring that the same fraction of the total peak (i.e. the same measurement width) was used for all samples and standards.

With the advent of high-resolution detectors, peak interference became the exception rather than the rule and the peak limits were extended down the sides of the peak to the background continuum level. This, the **total peak area method**, is now the standard method for peak area estimation for single un-interfered peaks. Other methods for estimating peak background, such as the *Wasson* and *Quittner methods*, found limited favour but, except for a few special situations, these offered no overall advantages. As an example, the Quittner method, which involved fitting a polynomial function to the background channels either side of the peak, is more accurate when the peak sits on an obvious nonlinear background, such as the top of a Compton edge.

5.4.1 Simple peak integration

In both the Covell and total peak area methods, the background level is estimated by using the channel contents at the upper and lower edges of the peak region (Figure 5.5).

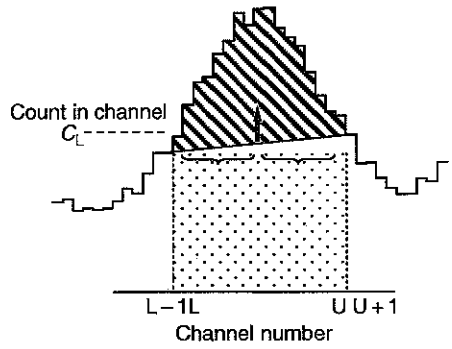

Figure 5.5 Calculation of peak area using the Covell method

If we take the first channel on each side of the peak beyond what we consider as being the peak region as representative of the background, then the gross (or integral) area of the peak is:

$$G = \sum_{i=L}^{U} C_i \tag{5.35}$$

where C_i are the counts in the ith channel (see Figure 5.5). The background beneath the peak is estimated as:

$$B = n(C_{L-1} + C_{U+1})/2 \tag{5.36}$$

where n is the number of channels within the peak region and C_{L-1} and C_{U+1} are the counts in the channels immediately beyond the lower and upper edge channels L and U. This background is, mathematically, the area of the background trapezium beneath the peak. It is more useful to think of this as the mean background count per channel

beneath the peak, multiplied by the number of channels within the peak region.

The net peak area, A, is then:

$$A = G - B = \sum_{i=L}^{U} C_i - n(C_{L-1} + C_{U+1})/2 \qquad (5.37)$$

It is important to appreciate that while we can calculate precisely the number of counts within the peak region (G), we can only ever estimate the number of background counts beneath the peak. We can never know which counts within the peak region are due to background and which are the peak counts. In most spectra, the peak background continuum derives from the sample itself. Unlike simple total activity counting, such as Geiger–Müller counting, we cannot take away the sample to determine a precise background count. In certain circumstances, in particular, when small peaks lie on large backgrounds, the uncertainty on the background estimate can dominate the total uncertainty of the peak area measurement.

Background estimates can be made more precise (i.e. less uncertain) by using more channels to estimate the mean count per channel under the peak. Figure 5.6 shows the general principle. Instead of a single channel, m channels beyond each side of the peak region are used to estimate the background beneath the peak.

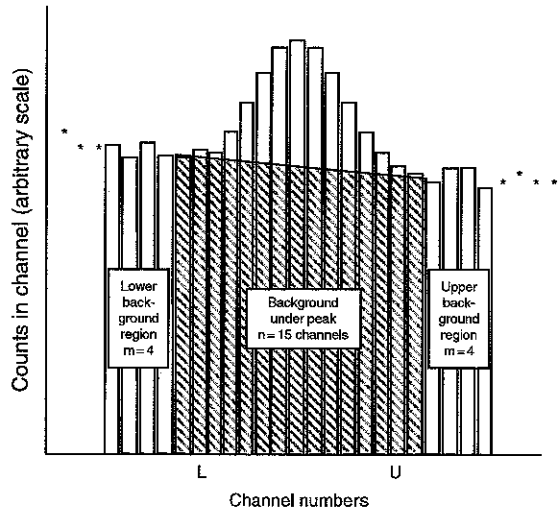

Figure 5.6 Calculation of peak area using extended background regions

Extending Equations (5.35) to (5.37) we find:

$$A = \sum_{i=L}^{U} C_i - n\left(\sum_{i=L-m}^{L-1} C_i + \sum_{i=U+1}^{U+m} C_i\right)/2m \qquad (5.38)$$

Again, the background term is the mean background count per channel, but now estimated using upper and lower background regions, m channels wide, multiplied by the number of channels within the peak region. There is little point in estimating a peak area unless the statistical uncertainty of that peak area is also calculated. If, as we have stated above, $A = G - B$, then, according to Equation (5.11), the variance of the net peak area is given by the sum of the variances of these two terms, giving:

$$\text{var}(A) = \text{var}(G) + \text{var}(B) \qquad (5.39)$$

Substituting for the individual variances and using Equation (5.15):

$$\text{var}(A) = \sum_{i=L}^{U} C_i + n^2\left(\sum_{i=L-m}^{L-1} C_i + \sum_{i=U+1}^{U+m} C_i\right)/4m^2 \qquad (5.40)$$

From this, we can calculate the standard deviation, σ_A.

This simple method described here assumes that the background is linear from the bottom to the top edge of the peak. In fact, examination of well-defined peaks shows that the background appears to have a 'step' beneath the peak (see Figure 8.6). Nevertheless, for most everyday purposes the method provides satisfactory results. The simple method cannot, of course, provide satisfactory results in cases where peaks are overlapped.

It is still possible to find incorrect expressions for the calculation of peak area uncertainty in the literature. The confusion arises because of a failure to appreciate that unlike single background counts where the variance of the count is numerically equal to the count itself, the variance of a peak background depends upon the number of background channels used. The offending expressions are variation of the form:

$$\sigma_A = \sqrt{(A + 2B)}, \text{ or } \sigma_A = \sqrt{(G + B)} \qquad (5.41)$$

These expressions are certainly correct for a single count plus background count, for example, from a simple beta counter. They are not valid for peak area calculations where Equation (5.40) must be used, resulting in the correct expression:

$$\sigma_A = \sqrt{[A + B(1 + n/2m)]} \qquad (5.42)$$

In the simple case of a single count plus background count, var(B), according to Poisson statistics is indeed equal to B and the expressions (5.39) and (5.40) are equivalent. In the peak area case, while var(G) is numerically equal to G, a sum of counts, the variance of the background estimate, var(B), depends upon the number of channels used to estimate (as opposed to measure!) the background as we saw earlier. Equation (5.41) does not take this into account and must, therefore, be generally incorrect. It is only true for the single case when $n = 2m$.

5.4.2 Peaked-background correction

So far, we have discussed only the situation when the background to the peak is a continuum. Measurements of radionuclides that are detectable in natural background must take that additional background component into account. In these cases, the backgrounds to the peaks will be peaks themselves and will be unavoidably included within the overall calculated peak areas. A background spectrum must be measured and the appropriate peak areas determined and subtracted from the sample peak areas. **Peaked-background** is most likely when dealing with environmental samples where, one hopes, the sample activity is near to background levels.

In some analysis programs, peaked-background correction is made after peak areas and the background contribution have been separately converted to nuclide activities. Since the calculation of activity necessarily introduces extra uncertainties, it makes sense to make the background correction at the earliest possible stage of the analysis process. Ideally, analysis programs should allow the correction to be made in terms of peak count rate in counts per second.

For example, if the peak area is A counts accumulated over Δt seconds of live time, the net peak area will be:

$$A_{\mathrm{NET}} = A - B_{\mathrm{PBC}} \times \Delta t \tag{5.43}$$

where B_{PBC} is the background peak count rate in counts per second and, as in any background correction, the variance (from which the standard uncertainty and %RSD can be calculated) will be:

$$\mathrm{var}(A_{\mathrm{NET}}) = \mathrm{var}(A) + \mathrm{var}(B_{\mathrm{PBC}} \times \Delta t)$$

$$\mathrm{var}(A_{\mathrm{NET}}) = A + (B_{\mathrm{PBC}} \times r_{\mathrm{PBC}} \times \Delta t)^2 \tag{5.44}$$

where r_{PBC} is the uncertainty on B_{PBC}, expressed as a relative standard uncertainty (*not* as a percentage). Although commercial spectrum analysis programs will consider peaked-background correction, at least one, GammaVision™, takes no account of the uncertainty on the peaked-background. The effect of that is to increase the number of false positive results when there is little or no nuclide present over and above natural background.

Apart from natural background, peaked-backgrounds can be experienced if the detector is used in an area where there is an enhanced neutron flux. Gamma spectrometry close to nuclear reactors and accelerators can be a problem in this respect. Although the neutron fluxes may not be significant from a safety aspect, activation of the materials of the detector system and prompt gamma-rays from neutron capture can sometimes be a problem. Appendix C lists the prompt gamma-ray from the activation of ^{114}Cd within a graded shield.

Leaving such special cases aside, it cannot be assumed that background is constant. An obvious example is the common background nuclide, ^{60}Co, which decays with a half-life of 5.27 years by about 1 % per month. Many peaks within the natural background spectrum originate in the uranium and thorium decay series. The degree of ventilation in a counting room might alter the amount of radon within the room and the amount of daughter nuclides in equilibrium with it. Even the external cosmic-ray background can change over a period of time. Background spectra should therefore be measured regularly. Because of this variation, it is advisable to collate the analysis of several backgrounds to establish a true uncertainty over time, rather than depend upon the measurement uncertainty of a single measurement.

There is further discussion of the sources of background in Chapter 13.

5.5 OPTIMIZING COUNTING CONDITIONS

5.5.1 Optimum background width

Equation (5.40) implies that the uncertainty of the estimate of the background must depend upon the number of background channels used. Since the more channels that are used the better the background estimate, it would appear that the more channels the better.

However, as one uses more channels there are decreasing returns and one must not overlook the possibility of neighbouring peaks causing a wide background region to be nonlinear. What is the optimum number of channels to use? This depends upon the circumstances.

Figure 5.7 summarizes the results of an assessment of the measurement of a particular ill-defined peak taken from an actual gamma-ray spectrum as a function of the width of the background region. It is apparent that the uncertainty on the peak area estimate (expressed as percentage relative standard deviation in Figure 5.7) decreases as the number of channels used to estimate the background increases. It is obvious that two channels is a

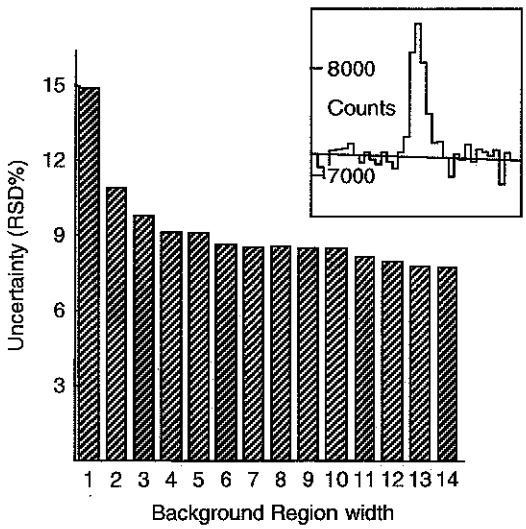

Figure 5.7 Variation of peak area uncertainty with background region width (the inset figure shows the actual peak measured)

considerable improvement on one, and three rather better than two are, but the reduction in uncertainty with each extra channel used gets smaller and smaller. There would be little extra value in using more than nine or ten channels and, in practice, the presence of neighbouring peaks may automatically limit the width of the background region.

If the peak is well defined and has a large area, then there may be little to be gained by using more than three or four channels. In such cases, the background uncertainty will have a much smaller effect on the uncertainty of the net peak area estimate. Note that the number of channels used for background estimate does not have any statistically significant effect on the net area, only on the uncertainty with which it is measured.

In an automatic spectrum analysis system, a compromise is usually made. Most commercial MCA and spectrum analysis programs use 3, 4 or 5 channels, depending upon the manufacturer and the situation. Note that there is no fundamental reason why the width of the background region should be the same above and below the peak region. If there were a potentially interfering neighbour above the peak, it would be sensible to use, say, three channels above and, perhaps, ten below. In such a case, the term $2m$ in Equations (5.38) and (5.42) would be replaced by $(m_L + m_U)$, where m_L and m_U are the lower and upper background region widths. When its 'Automatic' peak background width option is selected, GammaVision™ chooses 5, 3 or 1 channel widths on each side of the peak independently, depending upon whether the channels are deemed to represent a flat portion of the background continuum.

5.5.2 Optimum spectrum size

How does the peak area uncertainty alter with the number of channels in the spectrum? Conventional advice is often to use as many channels as possible. If you have an 8192 channel MCA system, use 8192 channels, if 16384 use 16384. The argument is that as detector resolution increases with advances in detector manufacture, the number of channels within each peak becomes smaller at a constant energy range. From the point of view of the spectrum analysis program, it may be advantageous to have more, rather than fewer, channels in each peak. However, what is not always taken into account is that as the spectrum is spread over more channels, for a constant counting time, the numbers of counts within the channels decrease. In order to compensate we must increase the number of channels in the peak region and ought to increase the number of channels in the background regions. Unfortunately, while the former will be done automatically by the spectrum analysis software there may be no option of altering the background region width. Figure 5.8(a) demonstrates how the uncertainty of a peak area estimation deteriorates as the number of channels in the spectrum is increased without a corresponding increase in background width.

The curves were calculated for a peak of energy 1332 keV, measured with a resolution of 1.8 keV and for peak areas of 500, 1000 and 10000 counts on a background of 1000 counts per keV with an overall range of 2048 keV. Figure 5.8(b) shows the appearance of the same 500 count peak in spectra ranging from 4096 to 32 768 channels. Although the overall peak area is unchanged, the uncertainty (i.e. the scatter from channel to channel) of the background is much greater because the counts are spread over more channels. The consequence is poorer precision for the area measurement. Even if the background region width is adjusted to suit the change in spectrum size by doubling the width for a doubling of spectrum size, there is no advantage, from the point of view of peak area precision, of using a larger spectrum size.

We should not forget though that peak width varies with energy. (The effect this might have on conversion gain was discussed in Chapter 4, Section 4.11.4.) If mainly low energy, and narrower, peaks are to be measured, then more channels per keV might be arranged either by increasing amplifier gain or doubling the spectrum size. If only high energy, and wider, peaks are of interest, then a smaller amplifier gain or spectrum size might be preferable.

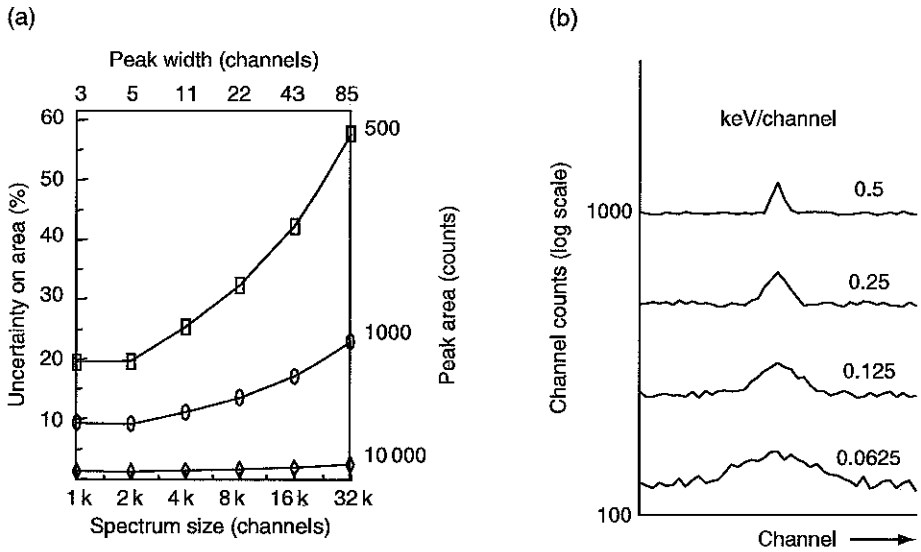

Figure 5.8 (a) Deterioration of peak area precision with increase in spectrum size. (b) Broadening of peak and increase in background 'scatter' with increase in spectrum size

Taking into account peak area measurement uncertainty, and the need for a reasonable number of channels within a peak (to facilitate peak searches and fitting), it would seem that spectrum sizes of 4096 and 8192 channels would be optimum. With current detectors and spectrum analysis software, there seems little point in seeking larger spectrum sizes.

5.5.3 Optimum counting time

In many laboratories, samples will not be submitted one by one for individual attention but in batches all to be counted within as little time as possible. Efficient use of counting equipment in terms of the time devoted to counting each sample can pay dividends when time and equipment are limited.

The first matter to be decided is the precision required of the final result. Let us suppose that, as an example, the reason for the count is to assess whether the ^{137}Cs in a sample of lamb is above or below some action limit. It might be that a poor precision result from a count of only five or six hundred seconds might answer the question for the majority of samples where the amount of ^{137}Cs was much lower, or indeed much greater, than the action limit. This would leave much more time available to achieve more precise results for those samples that are near to the action limit.

For a single sample with unlimited counting time available, the optimum count period is that which will provide sufficient counts in the spectrum to allow a peak area estimate with the predetermined satisfactory precision or to achieve a stated upper limit on activity. All multichannel counting systems will allow count periods to be automatically terminated after a preset time – usually, live and real time presets are options. Some systems will allow preset maximum count or total count within a channel or spectrum region. While such options can be useful, their value is limited by the fact that the precision of small peak areas may depend largely on the background continuum level. If this varies greatly from sample to sample, then preset-maximum count or gross region count are unlikely to be of any real use. In this respect, the most useful systems are those which monitor the actual peak uncertainty (as %RSD) continuously during the count and allow acquisition to be terminated when the desired precision is achieved.

More thought is needed when there are a number of samples to be counted with differing activities. Unless the facility for continuous monitoring of precision is available, there seems to be few options. Either one can select a count period which is expected to be satisfactory for the majority of samples and accept that some will be 'over-counted' and some 'under-counted', or one might split the samples into groups of roughly equal activity and count each group under optimum conditions for that group.

A special case arises when a background spectrum is necessary. An example might be the measurement of low levels of ^{60}Co where a peaked-background correction must be made. Let us assume that we have a batch of

samples to count, plus background, within a fixed overall time period. Is there an optimum way to split the available counting time between samples and background to achieve the best precision for the net count rates?

Taking simple single channel counts as an example, if we measure C counts in time Δt_C and measure a background count B in Δt_B, then the net count rate (R) is:

$$R = C/\Delta t_C - B/\Delta t_B \tag{5.45}$$

and the variance of this net count, V, is, according to Equations (5.11) and (5.15):

$$V = C/\Delta t_C{}^2 + B/\Delta t_B{}^2 \tag{5.46}$$

Now, if we have a fixed total count time, $\Delta t_C + \Delta t_B$, then the optimum sharing of the time will be found when the variance is at a minimum, i.e. when $dV/dR = 0$. If the mathematics is followed through, we find that this condition is obtained when:

$$\Delta t_C/\Delta t_B = \sqrt{(C/B)} \tag{5.47}$$

Now because C, the total count, can never be less than B, then Δt_B, the time devoted to background counting, should never be greater than Δt_B, or otherwise the precision of sample measurement will suffer. For a sample of four times background, we would achieve the best precision if we counted the sample for two thirds of the available time and the background for one third.

This is counter to the instinct to devote more counting time to the background – on the basis that because the background correction is applied to all sample counts, it should therefore be of high precision. If the sample count is near to background, then both are equally important in terms of precision. If the sample count rate is higher than background, then the background is proportionately less important and can be counted for a shorter time. Ultimately, of course, as the sample activity becomes very large the background becomes insignificant and we might choose not to measure it at all. If there is more than one sample, the conclusions are still valid. If the activity of the samples is unknown, divide the counting time so as to give the same counting time for each sample and background. If the samples are known to be greater than background, then reduce the background time appropriately and share the saved time equally between the samples.

5.6 COUNTING DECISION LIMITS

There is a great deal of confusion about the meaning of such terms as 'limit of detection', 'minimum detectable activity' and 'critical limit'. The terms are often treated as if interchangeable and there appears to be a considerable degree of freedom of choice in the manner in which they are calculated. Of these, minimum detectable activity (MDA) appears to be the most variable and I shall discuss this later. For now, I shall define a number of statistically determined levels that answer the following questions:

- **Critical limit** (L_C) – a decision level: '*Is the net count significant?*'
- **Upper limit** (L_U) – '*Given that this count is not statistically significant, what is the maximum statistically reasonable count?*'
- **Detection limit** (L_D) – '*What is the minimum number of counts I can be confident of detecting?*'
- **Determination limit** (L_Q) – '*How many counts would I have to have to achieve a particular statistical uncertainty?*'
- **Minimum detectable activity** (*MDA*) – '*What is the least amount of activity I can be confident of detecting?*'

These are considered in some detail by Currie (1968) and from a different perspective by Sumerling and Darby (1981). Note that, with exception of the MDA, the limits are calculated as a count rather than as an activity or other derived quantity. Note also that critical limit and upper limit relate to a measurement just made, whereas detection limit (and the associated MDA) and determination limit pose hypothetical 'what if' questions.

5.6.1 Critical limit (L_C)

'*Is the net count significant?*' After a peak area has been measured, it is important to establish its statistical significance. Since a peak becomes non-significant only by being 'lost' in the background, this cannot be done by reference to the peak area alone but must take into account the uncertainties of the background.

Let us suppose that a sample with no radioactivity at all in it was measured a large number of times. A series of counts – effectively background counts – would be obtained for which the mean net count above background was zero but distributed in a Gaussian fashion above and below zero (Figure 5.9). The spread, or standard deviation, of this distribution we will call σ_0.

How can we decide whether any particular measurement near to zero is truly zero or represents a true positive count? There must be some level, which we can call the **critical limit**, above which we can be confident, to a degree, that a net count is valid. We might decide that if the count, A, were above a certain number of standard

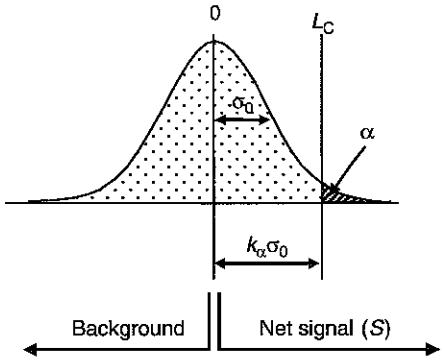

Figure 5.9 Definition of critical limit (the vertical axis represents the frequency of observing a particular count)

uncertainties of the distribution of counts we would be confident that the count existed, namely:

- if $A > k_\alpha \times \sigma_0$, the count is statistically significant;
- if $A \leq k_\alpha \times \sigma_0$, the count is not significant.

The factor k_α would be selected to provide a predetermined degree of confidence in the conclusion. For example, we may consider that it would be acceptable that if a count happened to be at the critical limit there would only be a 1 in 20, or 5 %, chance that we would judge the count to be present when in reality it was not. This is the same as saying that at the critical limit we would be 95 % certain that the count was not statistically significant. In this case, in statistical probability terms $\alpha = 0.05$ and, from one-tailed probability tables (Table 5.3), we find that k_α would be 1.645. (We use the one-tailed tables because we are only interested in the level being exceeded on one side, the higher, of the distribution.)

Table 5.3 k_α factors for particular probability intervals and the associated degrees of confidence

Probability interval, α	1-tailed confidence	2-tailed confidence	kα factor
0.1587	84.13	68.27	1.0
0.1	90.00	80.00	1.282
0.05	95.00	90.00	1.645
0.025	97.50	95.00	1.96
0.022 75	97.73	95.45	2.00
0.01	99.00	98.00	2.326
0.006 21	99.38	98.75	2.5
0.005	99.50	99.00	2.576
0.001 35	99.87	99.73	3.0

$$L_C = 1.645\sigma_0 \quad \text{(95 \% confidence limit)} \tag{5.48}$$

Operationally, this is applied as follows. If the net count is above L_C, we can say that the activity has been detected and can legitimately quote a value together with an associated uncertainty (confidence limit). Otherwise, we must judge the count not significant and quote an upper limit (see below). There is, of course, nothing sacrosanct about the 95 % confidence level. A higher or lower level might be chosen with an appropriate change to the value of k_α. Whatever the value chosen, it should be a positive decision in the context of the overall measurement system rather than a selection by default.

In practice, we do not know σ_0, the standard deviation (or uncertainty) of the net background count distribution. All we do have are the sample and background estimates. Taking Equation (5.39) again and remembering that var(count) = count, we can deduce that:

$$\text{var(net count)} = \text{net count} + \text{background} + \text{var(background)} \tag{5.49}$$

This is true whether we are dealing with single counts or peak areas. In the case where single counts are measured, the total count is C, and the background the single count B. If the net count $C - B$ is N then:

$$\text{var}(N) = N + B + \text{var}(B) \tag{5.50}$$

Now σ_0^2 is the variance of N when $N = 0$ and, for a *single count*, $\text{var}(B) = B$. Therefore:

$$\text{var}(N = 0) = \sigma_0^2 = B + \text{var}(B) = 2B \tag{5.51}$$

and it follows from Equation (5.48) that the critical limit, for 95 % confidence, is given by:

$$L_C = 1.645\sqrt{(2B)} = 2.33\sqrt{B} \tag{5.52}$$

For peak area calculations, the situation is complicated by the fact that the uncertainty of the background estimate depends upon the numbers of channels in peak and background regions. The principles leading to Equation (5.50) are still valid but now B is not a single background count but an *estimate* of a background that has an uncertainty that is not numerically equal to B itself. If we return and consider Equation (5.39) again, we remember that the second term is in fact the variance of the background estimate. Taking the second term of Equation (5.38) as B then:

$$\text{var}(B) = nB/2m \tag{5.53}$$

For a peak area, the expression equivalent to Equation (5.50) is therefore:

$$\text{var}(A) = A + B + nB/2m \tag{5.54}$$

This is, in fact a restatement of Equation (5.42).

Taking the net peak area A as zero, and rearranging:

$$\text{var}(A = 0) = \sigma_0^2 = B(1 + n/2m) \tag{5.55}$$

and:

$$L_C = 1.645\sqrt{[B(1 + n/2m)]} \quad \text{for } \alpha = 0.05 \tag{5.56}$$

Note that when the total number of channels used for the background estimation equals the peak width ($n = 2m$), Equations (5.52) and (5.56) become identical. Equating the peak background and its variance is a common misconception that appears in some current analysis programs. The effect is to underestimate the critical limit. For example, in the case of a peak 21 channels wide, when background regions of three channels are used, the factor used to calculate L_C should be 3.49 (Equation (5.56)) rather than 2.33 (Equation (5.52)) – an underestimate of 33 % which could lead to false positive identifications.

In most cases, the background to the peak will simply be the Compton continuum. However, if there is a peaked-background in addition, that must also be taken into account.

5.6.2 Upper limit (L_U)

'Given that this count is not statistically significant, what is the maximum statistically reasonable count?' The critical limit is used to assess the statistical validity of a calculated net count. If the net count, N, is below or equal to L_C, then the activity must be declared 'not detected' and an upper limit or 'less-than' level quoted. So we wish to define a level which we can be confident (to an appropriate degree) exceeds the actual peak area, if any. We can relate this to the notional distribution of counts we might obtain if we were to count the particular sample a large number of times (distribution (b) in Figure 5.10)

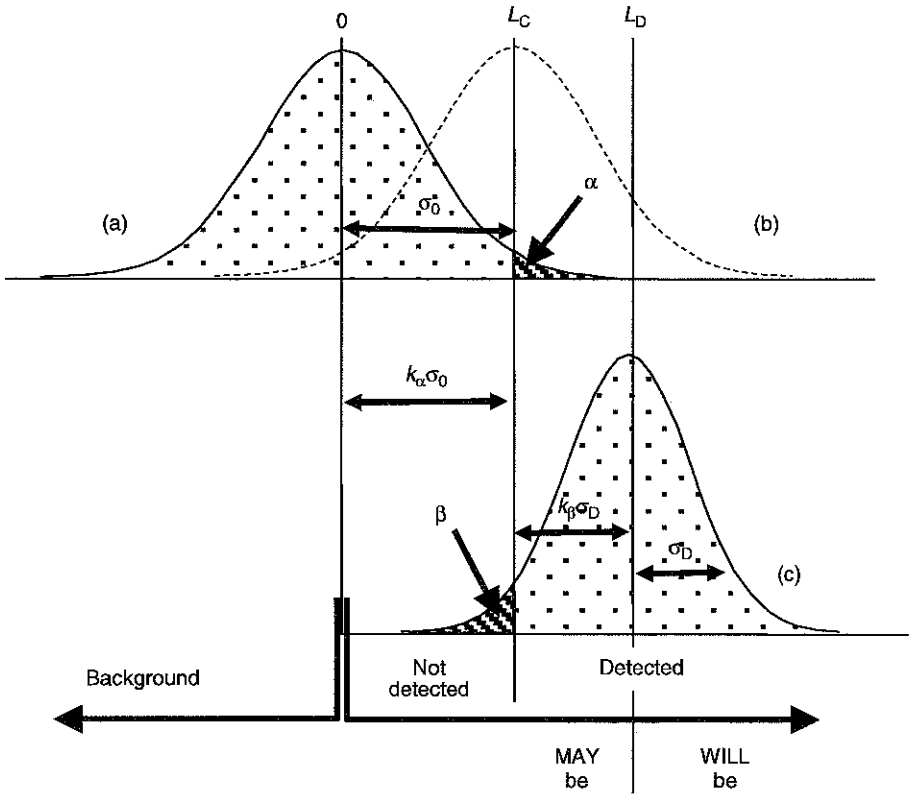

Figure 5.10 Definition of detection limit (the vertical axis represents the frequency of observing a particular count)

and define L_u accordingly. In fact for *any* distribution of counts, above or below the critical limit, we can say that:

$$L_U = N + k_\alpha \sigma_N \tag{5.57}$$

where σ_N is the uncertainty of the actual measured value; k_α is again the one-sided confidence interval and if we take again as our confidence level 95 % ($\alpha = 0.05$), then we can be sure that there is only a 1 in 20 chance that the true activity is greater than L_U. If N is less than zero then, although statistically reasonable, it does not represent a true situation and N should not be included in the calculation of L_U. There is little point in underestimating the upper limit. For 95 % confidence then:

For a simple count:

$$L_U = N + 1.645\sqrt{(N + 2B)} \tag{5.58}$$

For a peak area:

$$L_U = A + 1.645\sqrt{[A + B(1 + n/2m)]} \tag{5.59}$$

In both cases, the square root term is the standard deviation of the count, or of the estimated peak area, calculated in the normal manner, rather than of the background. You may notice that if N happens to be precisely zero, then Equations (5.58) and (5.59) reduce to the critical limit expressions (Equations (5.52) and (5.56)). Quite so. If the net count were zero, we would be 95 % certain that the true count were less than L_C – which is the definition of the critical limit. In spite of this, the upper and critical limits should not be used interchangeably.

5.6.3 Confidence limits

If a count, N, is found to be valid (i.e. greater than L_C), then the result may be quoted as a value with an appropriate confidence limit represented by k standard deviations of N, as explained in Section 5.3.1:

$$N \pm k_\alpha \sigma_N$$

The intention here is to state that the count or peak area we have measured lies, within a defined degree of confidence, between the two limits, $N - k_\alpha \sigma_N$ and $N + k_\alpha \sigma_N$. In this case, the factor for the two-tailed probability distribution should be used (see Table 5.3) and for 95 % confidence we might chose to present the result as:

$$N \pm 1.96\,\sigma_N$$

In the case of a single count, σ_N is $\sqrt{(N + 2B)}$ and for a peak area $\sqrt{[A + B(1 + n/2m)]}$.

5.6.4 Detection limit (L_D)

'*What is the minimum number of counts I can be confident of detecting?*' It is important to appreciate that the critical limit and upper limit are both *a posteriori* estimates based upon actual measured counts. They are statements of *what has been achieved* in the measurement. The detection limit answers the *a priori* question '*If you were to measure a sample, what would the count have to be for, say, 95 % certainty of detection?*' it is, therefore, a statement of *what might be* achieved. Detection limit is often confused with the critical limit. However, if the sample activity did happen to be exactly L_C (distribution (b) in Figure 5.10), statistically we would only be able to be sure (or 95 % sure!) of detection in 50 % of cases because the counts would be distributed symmetrically about L_C. It is clear that L_D must be some way above L_C (see distribution (c) in Figure 5.10).

Imagine that we have a sample with an activity that will provide a count precisely at our limit of detection. The distribution of counts, were we to measure the sample a large number of times, would have a standard deviation of σ_D. We wish to be certain, to a degree determined by k_β, that the chance of not detecting the activity when it is really there is only β, namely:

$$L_D = L_C + k_\beta \sigma_D = k_\alpha \sigma_0 + k_\beta \sigma_D \tag{5.60}$$

If α and β are both taken to be 0.05 (although there is no reason, other than convenience, why they should be so), then $k_\beta = k_\alpha = 1.645$.

Taking the single count situation where the net count is equal to the detection limit (i.e. $N = C - B$ and $N = L_D$), we could make the following statements:

- the variance of the distribution of counts = $\sigma_D{}^2 = C + B$;
- at the detection limit, C must be $L_D + B$;
- for a single count, $\sigma_0{}^2 = 2B$;
- combining these, $\sigma_D{}^2 = L_D + \sigma_0{}^2$
- hence from Equation (5.60), $L_D = k_\alpha \sigma_0 + k_\alpha (L_D + \sigma_0{}^2)^{1/2}$.

Rearranging this equation produces the simple relationship:

$$L_D = k_\alpha{}^2 + 2k_\alpha \sigma_0 \tag{5.61}$$

Putting $k_\alpha = 1.645$ and remembering from above that $\sigma_0{}^2 = 2B$ gives:

$$L_D = 2.71 + 4.65\sqrt{B} \tag{5.62}$$

Although for the peak area case the expression for σ_0 is more complicated, the mathematics is identical except that the final expression becomes:

$$L_D = 2.71 + 3.29\sqrt{[B(1+n/2m)]} \quad (5.63)$$

In practice, the calculation of L_D would be made once a background or spectrum, one which represented the particular situation for which detection limit is needed, had been measured. Again, although there are circumstances for which $L_D = L_C$, it is important to distinguish between these limits. Note that it is not essential to make $\alpha = \beta$. If they are not equal, the principle remains but the final expressions, derived from Equation (5.60) rather than Equation (5.61), will be more complicated.

From Figure 5.10, we can see that if the expected count due to a sample was below the critical limit we would almost certainly not detect the activity. If the expected count was above the critical limit but below the detection limit then we might detect the activity. If it was above the detection limit, then it is more likely than not that we would detect the activity. It is important to realize that it *is* possible to detect a count *below* the L_D – the detection limit. At first, this seems perverse. Consider a gamma spectrum. Is it not reasonable to suppose that, if a peak was detectable in 95 % of cases, it would be visible within the spectrum? Indeed, it is quite easy to show, by mathematically creating a continuum plus peaks, as in Figure 5.11, that if a peak contains a number of counts equivalent to L_D that it is indeed, in most cases, visible by eye – and, one would hope, by the spectrum analysis software. Practical experience of visual examination of gamma-ray spectra over many years leads me to suggest that if a peak cannot be identified visually it is not there – regardless of what the spectrum analysis software decides.

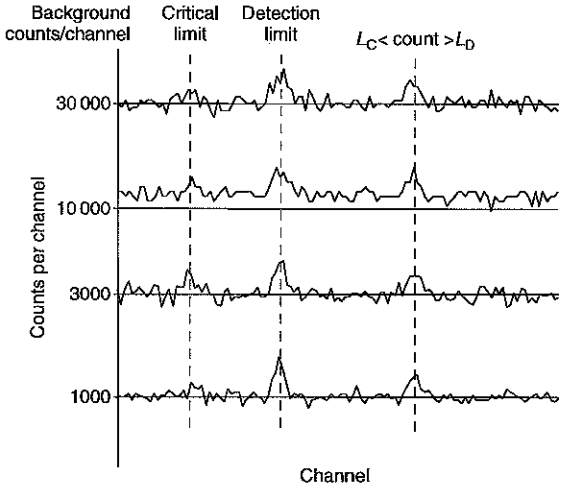

Figure 5.11 Peaks containing numbers of counts equivalent to L_C, L_D and midway between, on different levels of background continuum

Figure 5.11 shows that even peaks below L_D but above L_C can be seen. It follows that if a peak is not visible, and not detected, that the actual number of counts present, if any, must be less than L_D. However, the detection limit relates to a particular *confidence* of detection – in the equations derived above, 95 % confidence. Below L_D, detection will be less certain but will often still be possible.

5.6.5 Determination limit (L_Q)

'*How many counts would I have to have to achieve a particular statistical uncertainty?*' This limit is similar in concept to the detection limit and is also an *a priori* calculation but answers the question '*How many counts must there be to provide a result with, say, 10 % uncertainty?*' This implies that for a count, or peak area, equal to L_Q, the standard deviation σ_Q would be 10 % of L_Q or:

$$L_Q = k_Q \sigma_Q \quad (5.64)$$

where k_Q is the inverse of the required relative standard deviation. Following the logic of the mathematics summarized above for calculating L_D, it can be shown that:

$$L_Q = k_Q(L_Q + \sigma_0^2) \quad (5.65)$$

and the solution of this quadratic equation gives:

$$L_Q = k_Q^2[1 + (1 + 4\sigma_0^2/k_Q^2)^{1/2}]/2 \quad (5.66)$$

For example, if the required precision is 10 %, then $k_Q = 10$ and:

$$L_Q = 50[1 + (1 + B/12.5)^{1/2}] \quad (5.67)$$

for the simple count case; appropriate adjustment for σ_0 in the peak area case gives:

$$L_Q = 50\{1 + [1 + B(1 + n/2m)/25]^{1/2}\} \quad (5.68)$$

5.6.6 Other calculation options

Note that all these expressions are in terms of counts – the basic unit of uncertainty in radioactivity measurement – and assume equal count and background measurement

times. If count times are not equal, then adjustment is needed to account for this by using count rates, bearing in mind that, if the count rate is $C/\Delta t$ then the variance of the single count rate is $C/\Delta t^2$. The adjustment is easily made by altering every occurrence of B in Equations (5.52), (5.58), (5.62) and (5.67) to $B\Delta t_C^2/\Delta t_B^2$, where t_C and t_B are the sample and background count periods, respectively. In the case of peak area calculation, the sample and background are derived from the same spectrum and the question does not arise.

In the mathematics above, k_α and k_β were set equal. There is no reason, other than convenience, why this should be so and, in general, these confidence levels can be set independently.

5.6.7 Minimum detectable activity (MDA)

'*What is the least amount of activity I can be confident of measuring?*' This is a term often used loosely without qualification or rigorous definition and different interpretations can be made. An acceptable general definition would be that, given the circumstances of the particular spectrum measurement, the MDA is the minimum amount of radioactive nuclide that we can be ***confident*** that we can detect. First, this limit is, then, an activity rather than a count limit. It is often equated to the activity equivalent of the detection limit, L_D. However, there is a problem. As we defined it above, the detection limit is that count which we can be 95 % certain of detecting in the particular spectrum. However, as we saw, the detection limit is some way above the critical limit. We could, therefore, have the situation where a peak area measurement gave a net area which was significant (i.e. above the critical limit) but below the detection limit. Our activity result would then be below the minimum detectable activity. In fact, the minimum detectable activity is *not* the minimum activity detectable! In Figure 5.11, there is visible proof that we can detect peaks that would give an activity below the MDA. How should we interpret that?

The problem stems from the fact that there is a general misunderstanding of the meaning of the limit of detection from which MDA is derived. L_D is that number of counts that we can expect to detect in 95 % of cases (assuming α is 0.05). From this, we can calculate the MDA, which then becomes the activity that we can expect to detect in 95 % of cases. It answers the *a priori* question '*How good is your method?*' This is what should be quoted on tenders or in documentation describing methods. *The MDA should not be quoted as an estimate of upper limit* when a peak is *not* detected. It is unfortunate that not all of the commercial spectrum analysis programs give the option of quoting anything but MDA when a peak is not detected.

If a peak is not detected, the client, the recipient of the information, will wish to know an upper limit on the activity in the particular sample, not what amount of activity the analyst would be 95 % confident of measuring in an arbitrary similar sample in future.

I recommend the following strategy for normal gamma spectrometry, which I should say, is counter to common practice:

- Examine the region in the spectrum where the peak is expected to be. Calculate the net peak area, its uncertainty and the uncertainty of the peak-background correction.
- Calculate the critical limit, L_C, and compare with the net peak area.
- If the peak area is greater than L_C, quote a result with an appropriate confidence limit.
- If the peak area is not significant ($A < L_C$), then calculate the upper count limit, L_U, and pass that value through the calculation to produce an eventual upper activity limit.

Note that no account is taken of whether the peak has been explicitly detected or not. If the net peak area is significant, we can imply that it would have been detected if sought. This procedure has the advantage that exactly the same calculation is performed, whether the peak is present or not. If the peak is not present, one can have some confidence that the upper limit is realistic.

If the question arises '*What is the performance of your method?*' – perhaps to satisfy the requirements of a tender to provide gamma spectrometry services – the procedure should be:

- Find a spectrum that adequately represents the analyses to be performed. For example, if tendering for measurements on soil, find a typical soil spectrum measured under the conditions, sample size, count period, etc., which are intended to be used. (One could also make a case for selecting a worst-case or a best-case sample spectrum.)
- Using data from the spectrum, calculate the L_D at the appropriate region of the spectrum and convert to activity. That is then is the MDA for 95 % confidence of detection.

If all those making tenders are using the same procedure, then comparing tenders will be realistic and fair. Unfortunately, that may not be the case. A major problem is that there is a variety of equations that are used, rightly or wrongly, to calculate the MDA. Those quoted in one tender might not be comparable to those in another. Even worse

is the fact that in some software, even when the correct principle is applied, the equations properly relevant to the single count case are used for peak area measurement. Any spectrum analysis algorithms that equate the standard deviation of the background with the square root of the background are in error and will, in most cases, underestimate the limit. Further comments about the manner in which the MDA is calculated in spectrum analysis programs can be found in Chapter 9, Section 9.13.3.

In 1999, the UK Gamma Spectrometry Users Forum (now combined with the Alpha Spectrometry Users Forum as the Nuclear Spectrometry Users Forum) set itself the task of considering which were the 'correct' equations to use for MDA calculations. The intention was to make recommendations to users and software manufacturers alike. In the event, the scope was widened to encompass the correct use of the MDA and L_U. Recommendations were drawn up but, regrettably, have not to date been published. The reasons, apparently, are concerned with difficulties associated with recommendations about the uncertainties in the parameters used to calculate the MDA. To me, this seems to be an unnecessary delay, as I shall explain below.

5.6.8 Uncertainty of the (L_U) and MDA

The upper limit, in counts, is converted into an activity limit, A_U, by an equation of the form:

$$A_U = \frac{L_U}{m \times LT \times \varepsilon \times P_\gamma} \tag{5.69}$$

where m is the sample mass, LT the count period, ε the detector efficiency and P_γ the gamma emission probability. In some circumstances, there may be other factors involved.

L_U is estimated from the uncertainty on the peak-background continuum. If the sample were measured several times, the value of L_U would change statistically. However, *when determined from one particular spectrum*, it has no uncertainty – it is the number of counts below which we are 95 % confident that the true number of counts lies *in that spectrum*. On the other hand, all of the terms of the denominator have some degree of uncertainty, which could be propagated to an uncertainty on the value of A_U.

The unresolved, almost philosophical, question is whether the calculated A_U should be quoted 'as-is' or increased to take into account the uncertainties on the denominator. Suppose A_U were 100 Bq/g and the uncertainties on the denominator amount to, say, 5 %. Should we then quote $100 + 1.645 \times (A_U \times 5\%) = 108$ as the activity upper limit?

Replacing L_U in Equation (5.69) with L_D would allow calculation of the MDA and the same arguments apply. Again, there is an uncertainty on the MDA. Bearing in mind that the MDA is an *a priori* parameter, one could suggest that it should be determined from a number of measurements of an actual sample to estimate the true variability. In fact, the uncertainties on m and LT are very small and, for every measurement of a particular gamma-ray, the values used for ε and P_γ would be exactly the same; the variability of the MDA would be entirely due to the counting uncertainty on the background continuum. Experience shows, not surprisingly, that there is a much greater variability in the MDA between different samples than between different measurements on the same counting sample. Deriving a justifiable MDA is much more dependent on selecting an appropriate spectrum than on the details of the calculation.

Considering the fact that the uncertainties on m and LT are very small and, hopefully, those on ε and P_γ are also small, for the present, it seems reasonable to quote the MDA as calculated with no extra allowance for uncertainty. If the basis of reporting the MDA is to be altered, it is more important to achieve consistency between spectrum analysis programs than to worry about uncertainties on the MDA.

5.6.9 An example by way of summary

Statistics do not make light reading and the whole matter of decisions limits is undoubtedly confusing until one 'sees the light'. The following example might clarify matters. Consider the portion of a spectrum shown in Figure 5.12. The following have been calculated:

Gross counts in the peak region, G:	30 374	($n = 11$ channels)
Sum of background region counts, S:	12 040	($2m = 6$ channels)
Background correction, $B = nS/2m$:	30 100	(Equation (5.38))
Net peak area (counts), $A = G - B$:	274	
Critical limit (counts), L_C:	534	(Equation (5.56))

Because A is less than L_C, we can conclude that the peak was not detected. We must therefore calculate an upper limit:

Upper limit (counts), L_U:	809	(Equation (5.59))

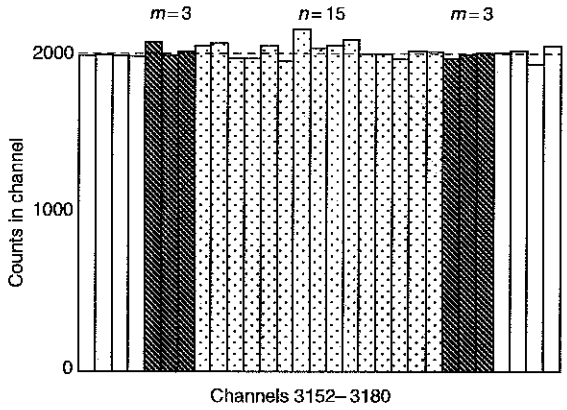

Figure 5.12 The peak used for the example calculations of critical, upper and detection limits (see text for further details)

The interpretation of this upper limit is that we are 95% certain that the actual number of counts in the peak is less than 809. (In fact, because this is a test peak, we know that the actual number of counts in the peak is 250.) We may wish to determine the detection limit:

Detection limit (counts), L_D: 904 (Equation (5.65))

This means that, given the general level of background counts in this spectrum, if we were to measure another spectrum we could be 95% certain of detecting a peak that had 904 counts in it. The difference between the upper and detection limits arises because we are asking different questions.

5.7 SPECIAL COUNTING SITUATIONS

In general, we assume that the statistics of counting can be adequately described by the Poisson distribution. When we calculated the various decision limits, we effectively assumed, for simplicity, the Normal distribution for the counts. We know, however, that Poisson statistics are only applicable when the probability of detection of the decay of any particular radioactive atom within the count period is small and when the statistical sample size is large. There are a number of circumstances when these conditions may not be met and we should consider whether the statistical treatment above is still valid.

5.7.1 Non-Poisson counting

If a sample is counted for a long time compared to the half life, the probability of decay within the count period is high. This condition is seldom met in most routine gamma spectrometry situations where half-lives are long, but may be met frequently when dealing with short-lived radionuclides. In activation analysis, for example, the measurement of radionuclides with half-lives as short as a few seconds is commonplace and a typical count period could be a number of half-lives. Even in such circumstances, the day is saved because the efficiency of detection of gamma radiation is usually very small, partly due to the intrinsic detector efficiency and partly due to geometry factors, reducing the probability of detection.

However, if the count period were long and the detection efficiency high (perhaps a low-energy gamma-ray emitted with high probability and measured close to the detector or inside a well detector), then the assumptions underpinning our use of the Poisson distribution are no longer valid. It is then necessary to return to the binomial distribution. There is no place here for the mathematics involved but it can be shown that if we observe a count of n, then the expected true count is:

$$E(n) = n + 1 - p\varepsilon \tag{5.70}$$

where p is the probability of decay, calculated using Equation (5.17), and ε is the known effective detector efficiency, taking into account the emission probability, the intrinsic detector efficiency and geometry factors. Similarly, the variance of n can be shown to be:

$$\text{var}(n) = (n+1)(1-p\varepsilon) \tag{5.71}$$

If $p\varepsilon$ is very small, then these relationships approximate to Equations (5.25) and (5.24) discussed above. If $p\varepsilon$ is large, then var(n) tends towards zero. This is not unreasonable; if we could detect every disintegration, then there would be no uncertainty associated with the number of counts detected.

Note, however, that even if the counts due to the measured species cannot be assumed to be Poisson distributed, it is most likely that the background count, be it a single count or a peak, will be. If this is so, then the gross count will also be Poisson distributed.

5.7.2 Low numbers of counts

If the number of counts accumulated is small, then, even though the count distribution will be Poisson, the approximation to a Normal distribution will not be valid. This means that the relationships to calculate the decision limits given above will not be valid. For number of counts of less than 25, we must resort to the Poisson distribution itself.

For example, if we wish to calculate the critical limit, L_C, we must consider the distribution of counts when the

sample has zero activity – background count, in effect. Suppose that we have a background count B in time Δt, then the probability of accumulating n counts is given by Equation (5.23), thus:

$$P(n) = \frac{(B)^n}{n!}e^{-B} \tag{5.72}$$

The probability that a blank sample could have a count greater than L_C, for a particular degree of confidence defined by α, is given by:

$$\sum_{i=n}^{\infty} \frac{[B]^n}{n!}e^{-B} \leq \alpha \tag{5.73}$$

So, L_C is the minimum value of n for which this condition is satisfied. As an example, Table 5.4 lists factors taken from published tables of the Poisson distribution. If our background count were, say, 2 counts per counting period, then the critical level for 95 % confidence would be 6 counts. Any count below that would have to be interpreted as 'not detected'. Similar considerations apply to the calculation of the limit of detection, which, it turns out, depends only on the critical limit as long the background is well known. Again, the appropriate limit can be taken from tables. In the case above, where the critical limit was 6 counts, the detection limit would be 10.5 counts. If the background is not well known, as might be the case for a gamma-ray peak with very small numbers of counts, then it becomes more appropriate to consider the problem as a binomial one.

Table 5.4 Decision limits in low count situations (for 95 % confidence in each case)

Background (counts)[a]	L_C (counts)	L_D (counts)
0.1	2	3.0
0.2	2	3.0
0.4	3	6.3
1.0	4	7.8
2.0	6	10.5
4.0	9	14.5
10	16	23.0
20	27	36.0

[a] The average background count within the sampling time.

For small numbers of counts, the Poisson distribution is not symmetrical. This implies that confidence limits associated with a small count will not be symmetrical either. Again, appropriate limits can be tabulated and examples, taken from Sumerling and Darby (1981), are shown in Table 5.5. Taking as an example a gross count of 10 counts, and using Table 5.5, we could only say that there was a 95 % probability of the true count being somewhere between 4.8 and 18.39 counts.

Table 5.5 95 % confidence limits for low numbers of gross counts

Gross count	Confidence limit (counts)	
	Lower	Upper
0.0	0.0	0.025
1.0	0.025	5.57
2.0	0.242	7.22
4.0	1.09	10.24
10	4.80	18.39
20	12.22	30.89

A number of authors have commented on the failure of the Currie equations to cater adequately for the low count situation and proposed alternative equations for critical and detection limits for the single count situation. Strom and MacLellan (2001) compare eight equations for calculating the critical limit, looking at their tendency to allow false positive identifications. None of the rules appears to be completely satisfactory but the authors advocate the use of a 'Stapleton rule' rather than the usual Currie equation in situations where background counts are very low and α, as in Section 5.6.1, is 0.05 or less.

In gamma spectrometry, we can be comforted by the fact that when the number of background counts measured beneath our peaks is 100 or more, the Currie equations are valid. In practice, there will be few occasions when this is not so.

5.7.3 Non-Poisson statistics due to pile-up rejection and loss-free counting

It is not generally recognized that the uncertainty of counts in a spectrum may not be adequately described by Poisson statistics when there are non-random losses of counts, especially when counting rates are high. A good summary of the situation is given by Pomme *et al.* (2000).

Measurements under a fixed real count time may suffer from an increased uncertainty when count losses exceed 20 %. The increase could be a significant extra contribution to an uncertainty budget. When using a loss-free counting system, artificial counts are added to the spectrum to compensate for lost counts in a manner that is not random. That distorts the statistics to a small extent.

Finally, although Poisson statistics can be relied upon for almost all normal counting with an extending live time, when pile-up rejection is used at high count rate the counting uncertainty will be underestimated. Under these conditions, a significant proportion of the total counts might be rejected and the counts lost during the rejection dead time periods will not be truly random.

All spectrum analysis programs will make the assumption that count uncertainties are described by Poisson statistics. If the actual count situation is one of those described above, an extra uncertainty allowance would have to be made externally from the program.

5.8 UNCERTAINTY BUDGETS

5.8.1 Introduction

When we present the results of our gamma spectrometry, we have a duty to take care that the results are as accurate as possible and that the uncertainty that we quote is realistic. It must take into account all of the known sources of uncertainty within the measurement process. Identifying and quantifying those uncertainties provides us with an ***uncertainty budget***. Laboratories that are seeking accreditation from bodies such as UKAS will have no choice but to create a satisfactory uncertainty budget.

The most obvious source of uncertainty is, of course, that due to the statistical nature of radioactivity counting. In many cases, it will be the major component of the total uncertainty and it is not uncommon for that alone to be the basis of the quoted confidence limits. However, that alone must underestimate the true uncertainty, especially when the counting uncertainties themselves are small and other sources of uncertainty become dominant. In setting up the budget, the uncertainty of all parameters which contribute to the final result must be assessed and quantified. Some may be justifiably found to be negligible. That conclusion in itself is part of the budget and should be documented.

Setting up an uncertainty budget is a valuable exercise, irrespective of the primary need to quote realistic uncertainties. While assessing in detail all of the relevant sources of uncertainty, it may be possible to remove some of them completely. For example, we might consider that positioning the sample on the detector would not make a significant difference to the result, even though there is potential for placing the sample in slightly different positions. To fulfil the requirements of our budget, we ought to check that out practically. However, if we provide positive sample location, preventing variable positioning, we remove that source of uncertainty completely. Similarly, preparing reference sources by mass, rather than by volume, will reduce the uncertainty on the result considerably.

It can be pointed out that matters such as errors caused by equipment malfunction or operator error do not form part of an uncertainty budget. Nevertheless, consideration of ways in which such problems can be avoided and, equally important, detected, should they happen, is a useful exercise that can only improve the robustness of the overall analytical procedure.

A good point at which to start setting up an uncertainty budget is to look at the way in which the result is calculated. The equation converting net peak counts, C, to activity per unit mass, A, might be:

$$A = \frac{(C - B_p) \times e^{-\lambda t} \times R \times S}{m \times LT \times \varepsilon \times P_\gamma} \tag{5.74}$$

where the parameters have the same meaning as in Equation (5.69), with the addition of $e^{-\lambda t}$, the decay correction, B_p, the peaked-background correction, R, a possible random summing correction and S, a possible self-absorption correction. Within those parameters, there may be others hidden from view. For example, in addition to the normal counting uncertainty there may be other factors affecting the count rate, such as sample positioning or differences in sample height. In the case of the decay correction, we must take into account the uncertainty on the correction due to the uncertainty of the half-life. In principle, we ought to include the uncertainty of the calculated decay period, t, but, under normal circumstances, this would be known very accurately.

When we consider the efficiency, ε, we find we have a separate uncertainty budget to consider taking into account all of the factors involved in preparing the efficiency calibration reference source and the measurement of the calibration data. This would take into account the uncertainty of the certified source from which the calibration source was prepared and the uncertainty of interpolation of the calibration curve.

Combining all of these different sources of uncertainty into a single value can be confusing. There is no other advice to be offered other than to work carefully through the factors one by one and quantify the uncertainty of each before combining them. One may be asked by a client or regulatory body for a single overall uncertainty value summarizing the budget. Frankly, in radioactivity measurements, this is not possible. The overall uncertainty on a result depends critically on the counting uncertainty, that, in turn, depending upon the magnitude of the background to the peak. Measurement of different samples with the same activity of a particular nuclide might give results with considerably different uncertainties if they have different amounts of other nuclides in them. It is

possible, however, to give a reasonable uncertainty budget with the counting uncertainty listed as a separate, variable item.

5.8.2 Accuracy and precision

The performance of measurement systems has been traditionally defined in terms of *accuracy* and *precision*. **Accuracy** can be defined as a measure of how close a result is to the actual value and ***precision*** is thought of as the uncertainty of the result, which we could identify with the standard uncertainty. Modern usage in the context of quality of analytical results tends to avoid these terms. This is because there has been a more fundamental appreciation of the actual measurement process. For example, accuracy or, perhaps we should say, inaccuracy, involves bias within a measurement process as well as statistically determined factors that cause the result to be different from the true result. What, at one time, we would have blithely termed precision is now discussed as **repeatability**, the variability of a method when applied to measurements on a single sample within a laboratory, and **reproducibility**, which applies to measurements of that sample when applied by different laboratories using different instruments operated by different operators.

The IUPAC recommendations on this matter have been published by Currie (1995). On a day-to-day basis, there is little harm in applying the traditional usage. However, when producing formal documentation I would recommend that the IUPAC usage should be adopted.

5.8.3 Types of uncertainty

It was, at one time, conventional to identify uncertainties as 'random' or 'systematic'. Experience showed that it was not always possible to assign any particular uncertainty to one category or the other. For example, there may be sources of uncertainty with a Normal distribution – and therefore ought to be categorized as 'random' – which would be more understandable as 'systematic' in origin. Modern usage is to treat each source of uncertainty separately and calculate the standard uncertainty, taking into account the type of distribution involved, before combining with other uncertainties.

However, a new distinction has arisen – Type A and Type B uncertainties. Type A uncertainties are defined as those that have been determined by repeated measurements to assess the magnitude and distribution of the parameter. Type B uncertainties are those whose magnitude has been derived in any other manner. For example, the uncertainty on gamma-ray emission probability is Type B because the data will have come from literature sources, as will the uncertainty on a certified source activity that will have been taken from the calibration certificate.

Counting uncertainties are a special case. Unlike all other measurements, the nature of radioactive decay, and a considerable body of theory and practice, means that we can establish the uncertainty of a count rate from a single measurement. Because of that, we regard counting uncertainty as Type A because we measure it, although not by repeated measurement. The designation Type A or Type B has no bearing at all on how the uncertainty is incorporated into the budget.

5.8.4 Types of distribution

So far, we have taken it as read that the parameters we are dealing with have a Normal distribution. In many cases, this will be so, but there may be exceptions. For example, suppose samples are placed manually on the cap of a detector for counting. The variation in sample position is likely to alter the sample count rate. What is the shape of the distribution of count rates when a large number are measured? One could argue that the most likely position would be the centre of the detector cap and that the count rates would, therefore, be distributed Normally. On the other hand, one might make an equally attractive case for saying that, taking into account the fact that different operators, with differing perceptions of where the centre was, would be placing samples on the detector with equal probability over a limited area of the detector face. In that case, there might be a rectangular distribution of count rates. Other situations might generate triangular or U-shaped distributions. How do we handle these different distributions? Fortunately, the answer is simple. It should be remembered that all of the statistical relationships and the equations for combining uncertainties are valid, whatever the shape of the distribution, as long as we are consistent in the use of standard uncertainty. It is only necessary to work out the standard uncertainty of the assumed distribution. This might be done by repeated practical measurements. However, if the extreme limits of the measurement are known, the appropriate factor in Table 5.6 may be used to convert the range between those extremes into a standard uncertainty.

5.8.5 Uncertainty on sample preparation

Ideally, samples presented for gamma spectrometry would be homogeneous. Unfortunately, in the real world samples are often far from homogeneous, much less representative. A 100 g sample of a decommissioning waste submitted

Table 5.6 Calculation of standard uncertainty[a] for different distributions

Distribution	Parameter	Divisor
Normal	68 % Confidence limit (1σ)	1
Normal	95 % Confidence limit (2σ)	2
Rectangular	Half-range	$\sqrt{3}$
Triangular	Half-range	$\sqrt{6}$
U-shaped	Half-range	$\sqrt{2}$

[a] Standard uncertainty = parameter/divisor

for gamma spectrometry can hardly be said to be representative of the tons of material to be disposed of, which might be crushed brick, concrete, soil or even floor sweepings. My personal feeling is that, because the gamma spectrometry laboratory has no control over the sampling procedure, the unrepresentative nature of the sample cannot be included as part of the uncertainty budget.

Once the sample has been received, however, it is a different matter. In principle, it is the duty of the analyst to provide an analysis that is representative of the sample provided. However, how far should the analyst go in achieving that? If the sample is clearly not homogeneous, steps need to be taken to make it so, especially if it must be sub-sampled. That may involve crushing or grinding to remove large lumps of material and perhaps segregation into clearly different portions of the sample. From a gamma spectrometry point of view, it may only be necessary to make the sample macroscopically, rather than microscopically, homogeneous – i.e. very small grain size is not essential. Although such procedures can be expected to reduce the uncertainty on the composition of the counting sample, whatever procedure is adopted, the laboratory should have some idea of the final uncertainty. How that is achieved is another matter. Homogeneity can only be properly assessed by a number of measurements on sub-samples. That may be acceptable in a research environment where any amount of time and effort can be devoted to a final high-quality measurement. However in a commercial environment the client is unlikely to wish to pay for anything other than a notional attempt at homogenization and a single measurement. One could, perhaps suggest that representative samples of a 'typical' matrix were homogenized and measured and several measurements made to assess the uncertainty on the composition of a sub-sample. However, practical experience suggests that in a commercial laboratory with a wide range of received samples there is little which can be described as 'typical'. I can only suggest that in such cases, prior to any work being done, an agreement should be reached with the client on the procedure to be carried out on the sample to achieve assumed homogeneity. Having done that, one could reasonably exclude homogeneity from the uncertainty budget, unless one does indeed have a reasonable idea of the magnitude.

Having achieved a (notionally) homogeneous material, it must be weighed into a counting container. Two sources of uncertainty remain; the mass and height (shape) of the sample. The uncertainty on the mass is small and can easily be quantified in relation to the significance of the least significant digits on the balance display. Uncertainty of the count rate with height of the sample must be assessed experimentally. When samples are measured close to the detector, as is usually the case with low-activity samples, this uncertainty can be significant. Ideally, variation in sample height should be eliminated by using a plunger to lightly compress the sample to a standard height. Chapter 7, Section 7.6.6 describes an empirical correction to count rate for sample height. Determining the factors needed to make that correction would in itself give an idea of the variability of count rate with sample height.

5.8.6 Counting uncertainties

The very act of placing the sample in its counting position is uncertain unless there is a positive sample location. I would recommend that sample locators are always used.

Statistical counting uncertainties are always present, of course, but are always taken into account within the spectrum analysis program. Because these uncertainties vary from sample to sample, from nuclide to nuclide within the sample and even from peak to peak of each nuclide, it makes no sense to include counting uncertainty as part of the uncertainty budget, except to point out that it is variable and taken into account. If peaked-background corrections are involved, you should be aware that the spectrum analysis program may not take into account the uncertainty of the background correction itself.

In routine gamma spectrometry, uncertainty on the timing of the count can be ignored. Only if the count rate varies considerably during the count, for example, if measuring very short-lived nuclides, is there likely to be any live timing problem.

If counting losses due to random summing and/or self-absorption are corrected for, then these corrections will themselves have an uncertainty that must be accounted for. If these corrections are made by the spectrum analysis program, you should make sure, by reading the manual and by validation measurements, that the uncertainties assigned by the program are reasonable.

5.8.7 Calibration uncertainties

Nuclear data uncertainty

The value of the analysis result ultimately depends upon the value of the gamma-ray emission probability. Reputable nuclear data tables will provide you with an uncertainty on these values.

There is, of course, also an uncertainty on the half-life and, if a decay correction is made, the uncertainty on that should also be included. The commercial spectrum analysis program libraries may only allow a single nuclide uncertainty factor to be accounted. If so, it will be necessary to devise a single factor, taking into account the likely magnitude of any decay correction and the various emission probability uncertainties for each nuclide.

Should uncertainty on the nuclear data represent part of one's uncertainty budget? In an ideal world, every laboratory would use the same, well evaluated, set of nuclear data. For the nuclides within the DDEP database (see Chapter 15, Section 15.2 and Appendix B), the standard uncertainties are small; in general, less than 1 %. However, for other nuclides they may be much larger; for the 68.28 keV gamma-ray of ^{234}Th, the emission probability is quoted in the LARA database as 0.048 ± 0.006, a relative uncertainty of 12.2 %. This means that, if taken into account, no laboratory, however careful and skilful, can provide ^{234}Th results with an uncertainty of better than 12 %. Within an intercomparison exercise, where the intention of the measurements is to compare methods or laboratories, if everyone were using the same nuclear data the inclusion of the uncertainty on gamma emission probability would obscure underlying differences due to methodology. However, under normal circumstances (and from the point of view of the recipient) the nuclide data uncertainty does represent part of the overall uncertainty of the result and should be included.

Uncertainty on efficiency calibration standards

When purchased, the reference material from which the calibration sources are prepared will be accompanied by a calibration certificate. This will list, for each nuclide, the activity per unit mass and the overall uncertainty on that activity. These uncertainties should then be taken into account when the efficiency calibration curve is created (see Chapter 7, Sections 7.6 and 9.9). Ideally, they would be used to weight the corresponding points within the fitting process.

It is unlikely that the calibration points will lie exactly on the fitted calibration line. The degree of scatter of the calibration points around the line can be said to represent both the 'goodness of fit' of the calibration data and the uncertainty of estimating the efficiency obtained by calculation from the calibration equation. (We are ignoring here the effect of true coincidence summing, which would make the scatter worse.) This 'interpolation uncertainty' is the figure that one would wish to include in the uncertainty budget.

There will, of course, also be uncertainties introduced when preparing the calibration source. However, they will be a constant amount on each calibration point and it would not be useful to include them in the weighting process. In fact, little more than weighing will be involved in most cases and the extra uncertainty is likely to be small. Nevertheless, it should be accounted for by combining with the interpolation uncertainty. Note that the individual uncertainties on the amount of each nuclide in the reference material do not appear directly in the budget. These will contribute to the scatter on the calibration curve. On the other hand, if individual efficiencies, for particular gamma-rays of particular nuclides are used, the uncertainty on the amount of nuclide in the calibration source should be taken into account.

5.8.8 An example of an uncertainty budget

Table 5.7 shows a notional uncertainty budget based on Equation (5.74). The data in the table are quoted by way of examples; they should not be taken too seriously. Indeed, even the choice of items may not be relevant to other detector and analysis systems. The procedure one should follow is as follows:

- List all identified sources of uncertainty. It may help to group them into categories such as 'Source preparation', 'Calibration', 'Counting' (Column 1).
- For each assess, or measure, the magnitude of the uncertainty (Column 2).
- Decide what that magnitude means. Is it a standard uncertainty or is it a range? (Column 3).
- Decide what the probability distribution is (Column 4).
- Write down the divisor corresponding to that distribution taken from Table 5.6 (Column 5).
- Calculate the standard uncertainty by dividing the magnitude by the divisor (Column 6).
- Add all of the standard uncertainties in quadrature to give the overall standard uncertainty (at the foot of Column 6) and then multiply by the required coverage factor to give the final expanded uncertainty.

It is useful to consider where these various sources of uncertainty are taken into account, as in Column 7. In many cases, this will be within the spectrum analysis program, although it is possible that the program

Table 5.7 An example of an uncertainty budget based on Equation (5.74)[a]

Source of uncertainty	Magnitude (%)	Describing	Probability distribution	Divisor	Standard uncertainty (%)	Where taken into account
Sample homogeneity	Negligible	—	—	1	0	Sample preparation agreed with client is assumed to produce a homogeneous sample
Net counts: $(C - B)$	Variable	—	Normal	1	Variable	Automatic in spectrum analysis program
Sample position	Negligible	—	—	1	0	Sample locator used
Sample height	5, based on ±2 mm in sample height	Half-range	Rectangular	$\sqrt{3}$	2.89	'Additional random' in software
Decay[b] $e^{-\lambda t}$	0.1	Standard uncertainty	Normal	1	0.1	'Additional random' in software
Random summing correction: R	0.2	Standard uncertainty	Normal	1	0.2	Accounted post-analysis
Self-absorption correction: S	0.3	Standard uncertainty	Normal	1	0.3	Automatic in spectrum analysis program – empirical allowance
Mass: m	0.08 on 50.00 g	Full-range	Rectangular	$2 \times \sqrt{3}$	0.023	'Additional random' in software
Live time: LT	Negligible	—	—	1	0	—
Efficiency: ε	3.5 from calibration	2 × Standard uncertainty	Normal	2	1.75	'Additional random' in software
Emission probability: P_γ	3, typical from nuclear data	Standard uncertainty	Normal	1	3	Automatic in spectrum analysis program – defined in library
			Overall distribution: Normal		4.53 %	
	Expanded uncertainty (95 % uncertainty; coverage factor = 2): 9.07 %					

[a] The data in this table are intended as examples only and should not be taken to be representative of any particular or general circumstances.
[b] Corrected.

does not handle particular items correctly. In Table 5.7, this is indicated in the case of the uncertainty of the random summing correction where, in this case, a post-analysis adjustment to the uncertainty turned out to be necessary. The budget quoted referred to a laboratory when GammaVision™ was used for spectrum analysis. The items designated as being taken into account by 'Additional Random' refer to the box within 'GammaVision™' into which optional extra amounts of uncertainty, which would not otherwise be taken into account, can be specified. In the example of Table 5.7, an amount of 3.38 % would be specified, representing all the 'Additional Random' items summed in quadrature.

The example shown is incomplete in that it does not take into account the degrees of freedom of each of the uncertainty items. Unless the overall number of degrees of freedom is infinite, calculation of the expanded uncertainty by multiplying by the factors derived from Table 5.1 would not be valid. It would not be appropriate to go into such matters here and for a fuller explanation of uncertainty budgets other sources should be consulted (e.g. Bell, 2001 and UKAS, 1997). It should be said that, for the majority of uncertainty items in a gamma spectrometry budget, infinite degrees of freedom can be assumed.

PRACTICAL POINTS

- The basic distribution underlying counting statistics is binomial in nature.
- In most practical circumstances, it is appropriate to assume a Poisson distribution, which, if the number of counts is large, can be approximated by a Normal (Gaussian) distribution.
- For a Poisson distribution, the following is true: $\mathrm{var}(n) = n$.
- The simple peak calculation area algorithms are:

$$A = \sum_{i=\mathrm{L}}^{\mathrm{U}} C_i - n\left(\sum_{i=\mathrm{L}-m}^{\mathrm{L}-1} C_i + \sum_{i=\mathrm{U}+1}^{\mathrm{U}+m} C_i\right)/2m$$

$$\mathrm{var}(A) = \sum_{i=\mathrm{L}}^{\mathrm{U}} C_i + n^2\left(\sum_{i=\mathrm{L}-m}^{\mathrm{L}-1} C_i + \sum_{i=\mathrm{U}+1}^{\mathrm{U}+m} C_i\right)/4m^2$$

- For lowest peak area uncertainty, the background region width (m) should be as large as possible under the particular circumstances. There would be little point in using more than about 10 channels.
- From the point of view of peak area uncertainty, the optimum spectrum size is 4096–8192 channels, depending upon the gamma-ray energy to be measured.
- Optimum sharing of counting time between samples and background is achieved when the ratio of count times equals the ratio of sample to background activity.
- Decision limits are calculated according to the following equations:

Limit	Single count	Peak area
L_C (95 %)	$2.33\sqrt{B}$	$1.645\sqrt{[B(1+n/2m)]}$
L_U (95 %)	$(N+)^*1.645$ $\sqrt{(N+2B)}$	$(A+)^*1.645$ $\sqrt{[A+B(1+n/2m)]}$
(*if $A<0$ or $N<0$, then that part of the equation in parentheses is ignored)		
Confidence limits (95 %)	$1.96\sqrt{(N+2B)}$	$1.96\sqrt{[A+B(1+n/2m)]}$
L_D (95 %)	$2.71+4.65\sqrt{B}$	$2.71+3.29\sqrt{[B(1+n/2m)]}$
L_Q (95 %, >10 %)	$50\{1+\sqrt{[1+B/12.5)]}$	$50\{1+\sqrt{[1+B(1+n/2m)/25]}\}$

- If the combined probability of decay and detection ($p\varepsilon$) is high, then Poisson statistics are inapplicable. The correct Binomial treatment provides the following – for a count n, the expected count is $E(n) = n + 1 - p\varepsilon$ and the variance, $\mathrm{var}(n) = (n+1)(1-p\varepsilon)$.
- If the number of counts is small (< 25) then the decision limits cannot be calculated from the Normal distribution but must be taken from statistical tables of the Poisson distribution.
- All results should be accompanied by a realistic uncertainty, taking into account all sources of uncertainty in the measurement. This is arrived at by constructing a complete uncertainty budget.
- When counts are judged against the critical limit and found to be 'not significant', the upper limit should be quoted, *not* the MDA.
- The MDA should be used when assessing the performance, or expected performance, of a method.
- It is important to remember that the MDA is *not* the Minimum Activity Detectable.

FURTHER READING

- General statistics. A very good general introduction to statistics is:

Moroney, M.J. (1990). *Facts from Figures*, Penguin, London, UK.
Miller, J.C. and Miller, J.N. (1993). *Statistics for Analytical Chemistry*, Ellis Horwood, New York, NY, USA.

- Counting statistics. Although published some time ago, the following volume of the series on Nuclear Science is of value:

Stevenson, P.C. (1966). *Processing of Counting Data*, NASNS 3109, National Academy of Sciences – National Research Council, Washington, DC, USA.

- Nomenclature in the chemical measurement process:

Currie, L.A. (1999). Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995), *Anal. Chim. Acta*, 391, 105–126.

- Experimental uncertainty and presentation of results:

Campion, P.J., Burns, J.E. and Williams, A. (1973). *A Code of Practice for the Detailed Statement of Accuracy*, National Physical Laboratory, HMSO, London, UK.

UKAS (1997). *The Expression of Uncertainty and Confidence in Measurement*, M2003, United Kingdom Accreditation Service, HMSO, London, UK.

- Decision limit (this article is essential reading):

Currie, L.A. (1968). Limits for qualitative detection and quantitative determination, *Anal. Chem.*, **40**, 586–593.

- Statistics of small counts:

Sumerling, T.J. and Darby, S.C. (1981). *Statistical Aspects of the Interpretation of Counting Experiments Designed to Detect Low Levels of Radioactivity*, NRPB R113, National Radiological Protection Board, HMSO, London, UK.

Strom, D.J. and MacLellan, J.A. (2001). Evaluation of eight decision rules for low-level radioactivity counting, *Health Phys.*, **81**, 27–34.

Currie, L.A. (2004). Detection and quantification limits: basic concepts, international harmonization and outstanding ('low-level') issues, *Appl. Radiat. Isotopes*, **61**, 145–149.

- Non-Poisson statistics due to live time correction:

Pomme, S., Robouch, P., Anana, G., Eguskiza, M. and Maguregui, M.I. (2000). Is it safe to use Poisson statistics in nuclear spectrometry? *J. Radioanal. Nucl. Chem.*, **244**, 501–506.

- Uncertainty and uncertainty budgets:

Bell, C. (2001). *A Beginner's Guide to Uncertainty of Measurement*, Issue 2, HMSO, London, UK (this is a very worthwhile introduction to the subject with many references to more substantial documentation).

United Kingdom Accreditation Service (1997). *The Expression of Uncertainty and Confidence in Measurement*, M3003, HMSO, London, UK.

National Physical Laboratory (1973). *A Code of Practice for the Detailed Statement of Accuracy*, HMSO, London, UK.