Counting statistics

1. Essential equations

All of what follows can be derived from two equations:

1) For M radioactive decay events (*ie*. M counts), the standard deviation is $\sqrt{M}$, in other words, the observed number of counts is:

$$\boxed{M \pm \sqrt{M}} \tag{1}$$

2) For a sum (or a difference), the absolute errors add up in quadrature, *ie*

$$\boxed{Err^2(a \pm b) = Err^2(a) + Err^2(b)} \tag{2}$$

Applying these two principles to a sample for which one measures T counts (Total counts) and B counts for the background in the same time interval Δt, N the net counts due to the sample is:

$$N = T - B \tag{3}$$

Then the standard deviation on N:

$$Err^2 N = Err^2 T + Err^2 B = T + B \tag{4}$$

or

$$Err(N) = \sqrt{T + B} \tag{5}$$

or expressed as relative error (in three different ways)

$$\%err(N) = \frac{Err(N)}{N} \times 100 = \frac{\sqrt{T+B}}{T-B} \times 100 = \frac{\sqrt{N+2B}}{N} \times 100 = \sqrt{\frac{(S+1)}{B(S-1)^2}} \times 100 \tag{6a}$$

where $S = \dfrac{T}{B}$ (6b)

Note that Eqs. (5) and (6) give the value of *one* standard deviation, or 67% confidence level interval. In many disciplines, the error reported is the 95% confidence level, which corresponds to *two* standard deviations. For this matter, it is this last quantity which is reported as % error in the ouput of Liquid Scintillation Counters.

The quantity S in Eq. (6a) and (6b) is related to the so-called signal-to-noise ratio; it describes by how much the sample count “sticks out” of the noise (the background), If $S = 1$ (*ie*., $T = B$), there is no signal. If $S >> 1$ ($T >> B$), there is plenty of signal and life is easy. A problem arises when $S > 1$ while $S \approx 1$. In this case, given that there are experimental uncertainties attached to the values of T, B and consequently of S, one may be faced with the question: is really $S > 1$ and if so, by how much? In other words, does my sample contain significant activity?

2. Meaning of the *two* standard deviations (or 95% confidence level).

Reporting an experimental result with its 95% conf. level assumes that the particular measurement obeys a normal distribution which in turn means that if the experiment were to be repeated, there is a 95% chance that the new result would agree with the previous result within the quoted interval (or to state it as pollsters like to do it: the result is considered accurate within the quoted interval 19 times out of 20).

Example of radioactive counting.

If one collects 100 counts in a given time interval, the 95% conf. level is $2x100^{1/2} = 20$.
If one were to recount for the same duration the same sample over and over, 95% of the measurements (19 out of 20) would fall in the interval 80 – 120 counts (after appropriate correction for decay if the half-life is “short”).

3. For how long should I count?

For how long should I count to get the net *cpm* (net *cpm* = gross *cpm* – blank *cpm*) of my sample within a predetermined percent error?

This question arises mostly for low activity samples. It is totally related to the question: is this cpm significantly higher than my background, or is there really some activity in this particular sample? High activity samples (> 1000 cpm), since they stand out of the background, do not present a problem in general.

From Eq. (6b), $S = \frac{T}{B} = \frac{cpm_{Gross}}{cpm_{Bkg}}$. Let *%err* be the desired percent error (95% conf.).

Then, the minimum counting time $\Delta t_{\min}$ to get net *cpm* of sample within *%err* is obtained by rearranging Eq. (6a):

$$\boxed{\Delta t_{\min} = \frac{1}{cpm_{Bkg}} \frac{S+1}{\left(\frac{\%err}{100}\right)^2 (S-1)^2}} \qquad (7)$$

Example:

After a quick count (say, 2 min per sample), the following is obtained: *Bkg* = 21 cpm and *Gross* = 34 cpm for the sample. Under these conditions, the percent error on these two pieces of data are respectively 31% and 24% (95% conf. level). This means that
Bkg = 21 ± 7 cpm and *Gross* = 34 ± 6 cpm. The net *cpm* is then:
Net = *Gross* – *Bkg* = 34 – 21 = 13 cpm and the error on *Net* , using elementary error propagation calculation of Eq. (5),

$$Err_{Net} = \sqrt{Err_{Gross}^2 + Err_{Bkg}^2} = \sqrt{6^2 + 7^2} = 9$$

i.e., *Net* = 13 ± 9 cpm (69%error!), meaning that I can be 95% confident that *Net* is anywhere between 4 and 22 cpm.
If I want to know the net *cpm* within 20%, using Eq. (7) above, the minimum counting time is calculated as ≈33 min (≈ ½ hour!!). One can verify that for such a length of time, the results would be (using the same cpm values):
Bkg = 21 ± 1.6 cpm (8% error) and *Gross* = 34 ± 2.1 cpm (6% error)
Then again:
Net = *Gross* – *Bkg* = 34 – 21 = 13 cpm and

$$Err_{Net} = \sqrt{Err_{Gross}^2 + Err_{Bkg}^2} = \sqrt{2.1^2 + 1.6^2} = 2.6$$

i.e., *Net* = 13.0 ± 2.6 cpm (20%error!), Now I can say with a 95% confidence level that my *Net* is between 10.4 and 15.6 cpm.

4. For a given counting time duration, how low a cpm can I measure with a given percent error?

For example. I routinely count my samples for two minutes; what is the lowest net cpm I can trust accepting a 20% error on this value.

This is called sometimes the minimum detectable activity (MDA), for a given counting duration and a given expected error.

The formula to use is a rearrangement of equation (7):

$$cpm_{Net}^{(\min)} \equiv MDA = \frac{1}{2 \times \Delta t \times \left(\frac{\%err}{100}\right)^2} \left(1 + \sqrt{1 + 8 \times \Delta t \times cpm_{Bkg} \times \left(\frac{\%err}{100}\right)^2}\right) \qquad (8)$$

where Δt is the counting time duration in minutes, and $\%err$ the desired percent error.

Example:

I count my swipe test samples for 1 min. To decide whether some surface contamination is present I must be able to detect reliably (± 20%) 100 cpm. Is a 1 min count sufficient? One must determine the Bkg cpm, then use formula (8) above with Δt = 1 min and $\%err$ = 20. Let set *Bkg* = 20 cpm (typical value). One gets *MDA* = 130 cpm, a bit on the tight side. 1.5 min counting duration would be ideal.