A further note on the variance of a background-corrected OSL count

Galbraith R.

Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK

(Received 6 March 2014; in final form 7 March 2014)

Abstract

Galbraith (2002) gave formulae for calculating the relative standard error of a background-corrected optically stimulated luminescence (OSL) count. This note extends those formulae to the case where the number of channels used to estimate the background is not necessarily a multiple of (and may indeed be less than) the number used to estimate the signal. The theoretical formulae are unchanged, but the estimate of the background is expressed in a more general notation and I comment further on how the Poisson over-dispersion may be estimated. I will use the same notation as in Galbraith (2002) and will repeat enough of the text there for this note to be self-contained.

Derivation

The usual scenario is as follows. Optical stimulation of an aliquot of quartz produces a series of counts —a number of recorded photons for each of N equal length consecutive time intervals (channels). The optically stimulated luminescence (OSL) "signal" is measured from the total count in the first n channels minus an estimate of the contribution to this count from "background" sources, which here are taken to include all sources other than the signal to be estimated. The background emission rate is assumed to be constant over the whole period used, and is estimated from counts later in this period, where the contribution from the signal is assumed to be negligible.

Mathematically, the above may be expressed as follows. Let y_i denote the OSL count from channel i, for $i = 1, 2, \dots, N$, and let $Y_0 = y_1 + y_2 + \dots + y_n$ be the total count over the first n channels. Write

$$Y_0 = S_0 + B_0$$

where S_0 and B_0 are the contributions to Y_0 from the signal (or source of interest) and background respectively. Of course S_0 and B_0 are not observed directly. Assume that S_0 and B_0 are independent random quantities with expectations μ_S and μ_B, and variances σ_S^2 and σ_B^2, respectively. Then the observed count Y_0 will have expectation $\mu_S + \mu_B$ and variance $\sigma_S^2 + \sigma_B^2$. An estimate of the signal μ_S is thus obtained by subtracting an estimate of μ_B from Y_0, i.e.,

$$\hat{\mu}_S = Y_0 - \hat{\mu}_B$$

We want to calculate the relative standard error of this estimate. An estimate of μ_B is usually obtained from the formula

$$\hat{\mu}_B = \frac{n}{m} Y_1 = \frac{1}{k} Y_1$$

where Y_1 denotes the total count over m later channels chosen so that the mean count per channel can be assumed to equal that for B_0, and $k = m/n$. It is desirable to choose m to be large enough to provide a sufficiently precise estimate of μ_B and it is convenient to choose $m = nk$ for some integer k that is large enough to assess any over-dispersion (with respect to Poisson variation) in total background counts made over different sets of n channels. However, k need not be an integer and the data used to estimate any over-dispersion need not be exactly the same as those used to estimate μ_B. Furthermore, k could be less than 1. That is, m could be chosen to be smaller than n, though the smaller m is, the less precise is the estimate of μ_B (other things being equal) and hence of μ_S. We here adapt the formulae of Galbraith (2002) to include this case explicitly.

Assume that all counts from individual channels that contribute to the background are independent random quantities from a distribution with mean μ_b and variance σ_b^2. This includes the y_is for the m channels that contribute to Y_1 as well as the unobserved counts that contribute to B_0. Then $\mu_B = n\mu_b$ and $\sigma_B^2 = n\sigma_b^2$ and the expected value and variance of Y_1 are $m\mu_b = k\mu_B$, and $m\sigma_b^2 = k\sigma_B^2$, respectively. So $\hat{\mu}_B$ has expectation μ_B and variance

$$\mathrm{var}(\hat{\mu}_B) = \left(\frac{1}{k}\right)^2 k\sigma_B^2 = \sigma_B^2 / k$$

Hence the variance of the estimated signal (corrected for background) is

$$\mathrm{var}(\hat{\mu}_S) = \mathrm{var}(Y_0) + \mathrm{var}(\hat{\mu}_B) = \sigma_S^2 + \sigma_B^2 + \sigma_B^2 / k \qquad (1)$$

and the relative standard error is

$$\mathrm{rse}(\hat{\mu}_S) = \frac{\sqrt{\sigma_S^2 + \sigma_B^2 + \sigma_B^2 / k}}{\mu_S} \qquad (2)$$

These formulae agree exactly with equations (1) and (2) of Galbraith (2002). In order to calculate this relative standard error in practice, we need estimates of σ_S^2 and σ_B^2 in addition to the estimate of μ_S.

If S_0, B_0 and Y_1 are all assumed to have Poisson distributions, then $\sigma_S^2 = \mu_S$ and $\sigma_B^2 = \mu_B$, and Equ. 1 becomes $\mathrm{var}(\hat{\mu}_S) = \mu_S + \mu_B + \mu_B / k$, which may be estimated as $Y_0 + \hat{\mu}_B / k = Y_0 + Y_1 / k^2$. Substituting these estimates into Equ. 2 gives the following estimated relative standard error:

$$\mathrm{rse}(\hat{\mu}_S) \approx \frac{\sqrt{Y_0 + Y_1 / k^2}}{Y_0 - Y_1 / k} \qquad (3)$$

This is the same as equation (3) of Galbraith (2002) with $\bar{Y}$ replaced by Y_1/k.

If the background counts do not have a Poisson distribution, but are over-dispersed, we may write, as in Galbraith (2002),

$$\sigma_B^2 = \mu_B + \sigma^2$$

for some positive value of σ^2 to be estimated. There are several possible ways to estimate σ^2. One is to use

$$\hat{\sigma}^2 = s_Y^2 - \bar{Y} \qquad (4)$$

provided that this is positive, where $\bar{Y}$ and s_Y denote the sample mean and standard deviation of total counts for sets of n channels that contribute to Y_1. (There are $[k]$ such total counts, where $[k]$ denotes the largest integer less than or equal to k.) This is the same as Equ. 4 of Galbraith (2002), but, as pointed out there, we would like an estimate based on a reasonable number of degrees of freedom, and it makes sense to obtain a pooled estimate from several different series. Such a pooled estimate of over-dispersion could be used for each series, while at the same time using separate estimates of background level.

Another method is to use

$$\hat{\sigma}^2 = n\left(s_y^2 - \bar{y}\right) \qquad (5)$$

provided that this is positive, where $\bar{y}$ and s_y denote the sample mean and standard deviation of the m counts for single channels that contribute to Y_1. This has the advantage of having more degrees of freedom, but the possible drawback of being based on variation between single channels rather than between sums over n channels. In theory, if the individual y_is are independent with constant variance σ_b^2, then sums of them over non-overlapping sets of n channels will be independent with variance $n\sigma_b^2$. However, in practice it is possible that the estimates based on Equ. 4 and Equ. 5 may not agree, and it is really Equ. 4 that is more relevant here. In any case, if m is small, a more reliable estimate of σ^2 may again be obtained by averaging the estimates for several series.

It is not so straightforward to obtain a corresponding estimate of σ_S^2 because the expected counts change rapidly at the start of the stimulation period. But there is perhaps a case for assuming that S_0 does have a Poisson distribution, while B_0 does not. The former comes from pure OSL emissions while the latter comes, at least partly, from other sources such as scattered light and instrument noise, which may not exhibit Poisson variation. Then we still have $\sigma_S^2 = \mu_S$ and the resulting estimated relative standard error is

$$\mathrm{rse}(\hat{\mu}_S) \approx \frac{\sqrt{Y_0 + Y_1 / k^2 + \hat{\sigma}^2 (1 + 1/k)}}{Y_0 - Y_1 / k} \qquad (6)$$

This is the same as equation (6) of Galbraith (2002) with $\bar{Y}$ replaced by Y_1/k.

Further comments

Li (2007) discussed the estimation of the error variance of background-corrected OSL counts, where he distinguished between two sources of background — namely, the "slow" component of the signal and "instrumental background". On the basis of laboratory experiments, he argued that it was just the counts from the latter source that were over-dispersed with respect to Poisson variation, and furthermore that such over-dispersion would be approximately the same for all analyses that used the same instrumental conditions. He therefore advocated the use of his Equation 6, which is equivalent to equation (6) of Galbraith (2002) and Equ. 6 above, with $\hat{\sigma}^2$ obtained from appropriate laboratory experiments and used for all analyses made under the same instrumental conditions. For his experiments, the estimate of over-dispersion based on variation between single channels was very similar to that based on variation between totals for sets of *n* channels.

Experiments to identify and measure specific sources of variation are to be encouraged, though it is also good practice to check that empirical estimates, such as those based on Equ. 4 or 5 or averages of them from several series, agree with those obtained from such experiments.

Finally, in practice an estimated background-corrected signal is usually divided by a similar background-corrected estimate obtained from a response to a test dose, in order to allow for a possible "sensitivity change" in the response to optical stimulation. The resulting quantity is usually denoted by L_x/T_x and its relative standard error may be calculated as the sum in quadrature of the separate relative standard errors of L_x and T_x. That is,

$$\mathrm{rse}\left(\frac{L_x}{T_x}\right) \approx \sqrt{[\mathrm{rse}(L_x)]^2 + [\mathrm{rse}(T_x)]^2}$$

where $\mathrm{rse}(L_x)$ and $\mathrm{rse}(T_x)$ are each given by $\mathrm{rse}(\hat{\mu}_S)$ in Equ 6 applied to the appropriate series of counts.

References

Galbraith R., (2002) A note on the variance of a background-corrected OSL count. Ancient TL, 20, 49–51.

Li B., (2007) A note on estimating the error when subtracting background counts from weak OSL signals. Ancient TL, 25, 9–14.

Reviewer

G.A.T. Duller