JAAS

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Statistical bias in isotope ratios

Cite this: *J. Anal. At. Spectrom.*, 2013, **28**, 52

Received 15th July 2011
Accepted 20th September 2012

DOI: 10.1039/c2ja10205f

www.rsc.org/jaas

Christopher D. Coath,*[a] Robert C. J. Steele†[a] and W. Fred Lunnon[b]

This paper presents the mathematics of the systematic bias in the expected value of the ratio of two noise-corrected Poisson-distributed variables, such as ion counting measurements. Such bias can lead to the reporting of incorrect ratios and, in some cases, systematic correlations with other measurements which can impact the scientific interpretation. We describe a novel method of treating such measurements which results in a negligible, exponentially small bias. We also re-examine the conventional approach deriving an exact expression for the bias including the noise correction explicitly.

1 Introduction

Analysts measuring a time series of ion beam intensities are faced with choices of how to process the data in order to estimate intensity ratios. Is it preferable, for example, to calculate the mean of the ratios, the ratio of the means or is there a better approach entirely? At the heart of the problem is the fact that the mean value a ratio estimator returns, over the long term, is *biased* relative to the true ratio being estimated. A simple example serves to illustrate the phenomenon of bias as follows. Suppose we write a 1 and a 3 on the faces of coin A and a 3 and 5 on the faces of coin B. Clearly, the mean value of a coin flip is 2 and 4 for coins A and B respectively. We now flip the two coins and record the ratio, B/A, as indicated by the upturned faces. In such an experiment, there are four equally probable outcomes: 3/1, 3/3, 5/1, and 5/3. Hence, the *expectation value* of the ratio, $E(B/A) = (3 + 1 + 5 + 5/3)/4$ or 8/3. The expectation value of B/A is the mean value of B/A over the long term. If we had hoped to estimate the 'true' ratio equal to the ratio of the means for each coin, $4/2 = 2$, then clearly the estimator B/A is biased. We define the bias as the relative difference between $E(B/A)$ and the true ratio. In this example, therefore, the bias is $(8/3-2)/2 = 1/3$.

The situation with ratios of ion-counting signals in isotope ratio measurements is directly analogous. The number of ions counted is accurately modelled by Poisson statistics which states that the probability of detecting X ions is given by

$$\mathrm{Pois}(X;\mu_x) = e^{-\mu_x}\mu_x^X/X!, \tag{1}$$

where μ_x is the *population mean* or expectation value of X. Note that this single parameter, μ_x, characterises the Poisson distribution. An analyst may wish to calculate an estimate of μ_x/μ_y, the ratio of the expectation values of two such signals, X and Y. More generally, it is often the case that analyte signals are superposed on 'noise', whose origin may be in the detector (dark noise), spectral or contamination during sample preparation. Regardless of origin the mean noise is often assumed to be constant, is determined empirically and the measured signals corrected for its contribution. Let the mean noise contribution to X and Y be μ_{x_0} and μ_{y_0} respectively. The ratio to be estimated, R, is, therefore, given by

$$R = \frac{\mu_x - \mu_{x_0}}{\mu_y - \mu_{y_0}}. \tag{2}$$

In principal, any function, $f(X,Y)$, may have it's bias relative to R determined by deriving the expectation valve, $E(f)$, given by the double summation over all X and Y of $f(X,Y)\cdot\mathrm{Pois}(X;\mu_x)\cdot\mathrm{Pois}(Y;\mu_y)$. For example, the commonly used ratio estimator,

$$r_0 = \frac{X - \mu_{x_0}}{Y - \mu_{y_0}}, \tag{3}$$

is considered in Section 2 but, even without detailed analysis, r_0 is clearly problematic for any integer value of μ_{y_0}, including zero, as $E(r_0)$ does not exist due to the non-zero probability of events where the denominator, $Y - \mu_{y_0}$, is zero. In Section 3 a novel, *quasi*-unbiased ratio estimator is proposed, which is well behaved for all Y and μ_{y_0}.

1.1 Previous work

Coakley *et al.*[1] and Ogliore *et al.*[2] have shown that X/Y is a *biased* estimate of R. Similar approaches by both[1,2] yield

$$E(X/Y) \approx \mu_x/\mu_y(1 + 1/\mu_y + 2/\mu_y^2) \tag{4}$$

in the limit of large μ_y. As we shall show in Section 2, this approximation holds providing $Y = 0$ events are excluded from the distribution. Noise correction is not considered explicitly but, since the method used by these authors is rather general, it

[a]University of Bristol, School of Earth Sciences, Wills Memorial Building, Queens Rd., Bristol, UK. E-mail: chris.coath@bristol.ac.uk; Fax: +44 (0)117 925 3385; Tel: +44 (0) 117 954 5370

[b]Department of Computer Science, National University of Ireland, Maynooth, co., Kildare, Republic of Ireland

† Present address: University of California, Los Angeles, Department of Earth and Space Sciences, 595 Charles Young Dr E., Los Angeles, CA 90095-1567, USA.

is straightforward to include it as a modification of the distribution. Note that eqn (4) is not a solution to the bias problem, it is only a means of estimating its magnitude. If it can be assumed that μ_y is constant then μ_y, and hence the bias, may be estimated and the ratio corrected but such a procedure lacks generality.

Ogliore *et al.*[2] compare ratio estimators $r_1 = \bar{x}/\bar{y}$ (the 'ratio of the means') and $r_2 = \overline{(x/y)}$ (the 'mean of the ratios') where $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$, and $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$ are the means of samples x_i and y_i, and $\overline{(x/y)} = n^{-1}\sum_{i=1}^{n}(x_i/y_i)$ by deriving expressions for $E(r_1)$ and $E(r_2)$. Note, however, that these reduce to a single problem as follows. Let x_i and y_i be instances of independent random variables X and Y respectively where X has a Poisson distribution with mean μ_x, which we shall write $X \sim \text{Pois}(\mu_x)$, and similarly $Y \sim \text{Pois}(\mu_y)$. Since the sum of Poisson variables is itself a Poisson variable, we have $n\bar{x} \sim \text{Pois}(n\mu_x)$ and similarly for $n\bar{y}$. With this nomenclature, $E(r_2) = E(X/Y)$ and is given by approximation 4 above and $E(r_1)$ has the same form with μ_x and μ_y replaced by $n\mu_x$ and $n\mu_y$ respectively (see Ogliore *et al.*[2] eqn (19) and (22)). The biases in r_2 and r_1 are, therefore, $O(\mu_y^{-1})$ and $O\{(n\mu_y)^{-1}\}$ respectively. Note that both r_1 and r_2 have, to first order, biases proportional to the reciprocal of the mean number of counts in the denominator *when the ratio is taken*.

Ogliore *et al.*[2] also consider Beale's ratio estimator, r_3, given by

$$r_3 = r_1\left(\frac{1 + \text{cov}(x, y)/(n\bar{x}\,\bar{y})}{1 + \text{var}(y)/(n\bar{y}^2)}\right) \tag{5}$$

where cov and var return the sample covariance and variance respectively. Beale's estimator reduces the bias to $O\{(n\mu_y)^{-2}\}$ but, in common with r_1, all n data are reduced to a single value so any true within-analysis variations in R, which may be of interest, are obscured.

2 The ratio of noise-corrected Poisson variables

We will assume that the distribution of the ion counts, X and Y, obeys Poisson statistics. Furthermore, we will consider only cases where X and Y are independent. This latter restriction is not so severe as it might seem since by far the most important correlated variations in ion counting signals are common mode 'intensity' fluctuations, that is, proportional changes in both μ_x and μ_y which largely cancel by taking the ratio. Of course, common mode variations will not be completely rejected on account of any intensity dependence of the bias.

Let R be defined as given by eqn (2) and let Z be distributed like Y but truncated at y_0, that is, the probability distribution function, $\Pr(Z)$, is given by

$$\Pr(Z; \mu_y, y_0) = \begin{cases} 0 & \text{for } Z \leq y_0 \\ N\text{Pois}(Z; \mu_y) & \text{for } Z > y_0 \end{cases} \tag{6}$$

where N is a normalisation constant and y_0 is an integer. To satisfy $\sum \Pr(Z) = 1$ the normalisation constant required is

$$N = y_0!/\gamma(y_0 + 1, \mu_y) = 1/P(y_0 + 1, \mu_y)$$

where γ is the incomplete gamma function and P the normalised incomplete gamma function.[3] We shall choose y_0 to be sufficiently large to ensure $Z - \mu_{y_0}$ cannot be zero or negative, *i.e.*

$$y_0 \geq \lfloor \mu_{y0} \rfloor.$$

Note that, in cases of large signal to noise ratio, the probability of $Y \leq \mu_{y_0}$ may be so small as to make excluding these events notional in practice. We shall now derive an expression for the expectation value of

$$r = \frac{X - \mu_{x_0}}{Z - \mu_{y_0}}, \tag{7}$$

which is the conventional expression for the noise-corrected ratio but with rejections to avoid zeroes or negative values in the denominator.

For independent random variables we can separate them thus

$$E(r) = E(X - \mu_{x_0})E\left(1\Big/\left(Z - \mu_{y_0}\right)\right)$$
$$= (\mu_x - \mu_{x_0})E\left(1\Big/\left(Z - \mu_{y_0}\right)\right)$$

reducing the problem to one in a single variable, Z, so we can drop the subscript y, *i.e.* $\mu_{y_0} \rightarrow \mu_0$ and $\mu_y \rightarrow \mu$.

The Taylor series expansion of $1/(Z - \mu_0)$ about $\mu_0 = 0$ is,

$$1/(Z - \mu_0) = (1/Z)\sum_{k=0}^{\infty}(\mu_0/Z)^k. \tag{8}$$

Note that the truncation of the distribution ensures $|\mu_0/Z| < 1$, guaranteeing convergence. To take the expectation value of the right-hand side (RHS) of eqn (8) requires an expression for the expectation value of $1/Z^k$. This problem is addressed in the appendix and given by eqn (25). Substituting into eqn (8) gives,

$$E\left(\frac{1}{Z - \mu_0}\right) = N\sum_{k=0}^{\infty}\mu_0^k \sum_{j=k+1}^{\infty}\frac{a_{k+1}(j)}{\mu^j}P(y_0 + j + 1, \mu),$$

where coefficients $a_k(j)$ are given by eqn (26) (or, more conveniently for computational purposes, using $a_1(j) = (j-1)!$, $a_j(j) = 1$ and the recursion 11). Changing the order of the summation,

$$E\left(\frac{1}{Z - \mu_0}\right) = N\sum_{j=1}^{\infty}\frac{P(y_0 + j + 1, \mu)}{\mu^j}\sum_{k=0}^{j-1}a_{k+1}(j)\mu_0^k. \tag{9}$$

The first few values of $a_k(j)$ are,

$j =$	1	2	3	4	5
$a_1(j)$	1	1	2	6	24
$a_2(j)$		1	3	11	50
$a_3(j)$			1	6	35
$a_4(j)$				1	10

From the exact form eqn (9), we can derive an asymptotic form in the limit as $\mu \rightarrow \infty$ and bounded μ_0 by replacing the infinite sum with a finite sum, and using $P(y_0 + j + 1, \mu) \rightarrow 1$, which is true so long as y_0 and j are also bounded. Hence,

$$E\left(\frac{1}{Z - \mu_0}\right) = \sum_{j=1}^{n-1}\frac{1}{\mu^j}\sum_{k=0}^{j-1}a_{k+1}(j)\mu_0^k + O(1/\mu^n) \tag{10}$$

Define the bias, $B(\mu,\mu_0)$, as the relative difference between $E(1/(Z - \mu_0))$ and $1/(\mu - \mu_0)$, *i.e.*

$$B(\mu,\mu_0) = (\mu - \mu_0)E(1/(Z - \mu_0)) - 1.$$

Substituting from eqn (10) and after some cancellation we have

$$B(\mu,\mu_0) = \frac{1}{\mu} + \sum_{j=2}^{n-1}\frac{1}{\mu^j}\left(j! + \sum_{k=1}^{j-1}\left(\mu_0^k(a_{k+1}(j+1) - a_k(j))\right) + O(1/\mu^n)\right).$$

From eqn (27) it follows that

$$a_{k+1}(j+1) - a_k(j) = ja_{k+1}(j), \tag{11}$$

hence

$$B(\mu,\mu_0) = \frac{1}{\mu} + \sum_{j=2}^{n-1}\frac{1}{\mu^j}\left(j! + j\sum_{k=1}^{j-1}a_{k+1}(j)\mu_0^k\right) + O(1/\mu^n).$$

For example, putting $n = 3$ we have

$$B = 1/\mu + (2 + 2\mu_0)/\mu^2 + O(1/\mu^3) \tag{12}$$

in agreement with Coakley *et al.*[1] and Ogliore *et al.*[2] for the case $\mu_0 = 0$.

3 *Quasi*-unbiased ratios

Here we present an alternative ratio estimator which reduces the bias to a factor which is exponentially small or *quasi*-unbiased. Furthermore, the method does not require truncation of the distribution so all the measured data can be used.

Let Y be distributed as before. We define a new random variable,

$$Y'(Y,\mu_0) = (Y + 1)/M(1, Y + 2,\mu_0) \tag{13}$$

where $M(a,b,z)$ is the Kummer confluent hypergeometric function. The series expansion of M is[4]

$$M(a,b,z) = \sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k k!}z^k \tag{14}$$

where $(a)_k$ is the rising factorial or Pochhammer's symbol,

$$(a)_k = \begin{cases} 1 & \text{for } k = 0 \\ a(a+1)(a+2)\ldots(a+k-1) & \text{for } k = 1,2,\ldots \end{cases}$$

Explicitly,

$$E(1/Y') = (\mu - \mu_0)^{-1}(1 - e^{-\mu + \mu_0}).$$

Proof. From eqn (13) the expectation value is given by

$$E(1/Y') = e^{-\mu}\sum_{y=0}^{\infty}\frac{\mu^y}{y!}\cdot\frac{M(1,y+2,\mu_0)}{(y+1)} = e^{-\mu}\sum_{y=0}^{\infty}\frac{(1)_y\mu^y}{(2)_y y!}M(1,y+2,\mu_0).$$

A summation of this form is given by Prudnikov *et al.*[5] which is

$$\sum_{y=0}^{\infty}\frac{(b-a)_y\mu^y}{(b)_y y!}M(a,y+b,\mu_0) = e^{\mu}M(a,b,\mu_0-\mu).$$

Substituting $a \to 1$ and $b \to 2$ gives,

$$\begin{aligned}\sum_{y=0}^{\infty}\frac{(1)_y\mu^y}{(2)_y y!}M(1,y+2,\mu_0) &= e^{\mu}M(1,2,\mu_0-\mu)\\ &= e^{\mu}(M(1,1,\mu_0-\mu)-1)/(\mu_0-\mu)\\ &= e^{\mu}(1-e^{\mu_0-\mu})/(\mu-\mu_0)\end{aligned}$$

where we have used Gradshteyn and Ryzhik[6] equation 9.212 and $M(1,1,z) = e^z$.

Let B' be the bias of $E(1/Y')$ relative to $(\mu - \mu_0)^{-1}$,

$$B'(\mu,\mu_0) = (\mu - \mu_0)E(1/Y') - 1 = -e^{-\mu+\mu_0}.$$

Therefore, noise-corrected *quasi*-unbiased ratios, r', can be computed from measurements of Poisson events X and Y with,

$$r'(X,Y) = \frac{(X - \mu_{x_0})}{(Y+1)}\cdot M\left(1, Y+2, \mu_{y_0}\right). \tag{15}$$

Using eqn (15) a ratio can be calculated at each cycle of data collection and, if required, statistics on these such as the mean and standard deviation. This contrasts using either the ratio of the means, r_1, or Beale's estimator, r_3, both of which return a single ratio from a set of n measurements.

Fig. 1 compares conventional and *quasi*-unbiased ratios from 128 cycles of measurements on boron isotopes ^{10}B and ^{11}B. The ion counts have been summed in 'blocks' of p cycles, the ratio taken for each block and the mean over all blocks plotted. The value of p, therefore, controls integration period and the total number of blocks, $q = 128/p$. Let the total counts in each block be x_i and y_i for the two isotopes, where $i = 1,2,\ldots q$. We use $p = 1,2,4,8$ and 16 and plot the means of $r'(x_i,y_i)(t_y/t_x)$ (red) and $(x_i/y_i)(t_y/t_x)$ (blue) against mean of $1/y_i$, where t_x and t_y are the cycle integration times for the two isotopes. The 'ratio of the means', equivalent to putting $p = 128$, is also shown and is indistinguishable from Beale's estimator. The noise is negligible for these data and has been set to zero. Both $^{11}B/^{10}B$ and its reciprocal are plotted to demonstrate that the effect is independent of which isotope is chosen as the denominator. The plots show that, to first order and, as predicted by theory, the conventionally processed data, shown in blue, plot as a straight line with equal slope and intercept. The y-intercept, which corresponds to counts $\to \infty$, should equal the unbiased ratio. Our novel *quasi*-unbiased ratio estimator (eqn (15)), shown by the red data have, as expected, no bias regardless of the number of counts.

Fig. 2 shows the results of a Monte-Carlo simulation with $\mu_x = \mu_y = 20$ and noise from 0 to 15. The simulation shows the noise-corrected ratio calculated in four ways: (i) conventionally without data rejection (eqn (3)), (ii) conventionally with rejection when the denominator is zero or negative (eqn (7)), (iii) Beale's estimator, and (iv) *quasi*-unbiased (eqn (15)) all with $\mu_{x_0} = \mu_{y_0} = b$. The mean of 10^6 ratio estimations is plotted for each value of the noise, which is incremented in steps of 0.01.

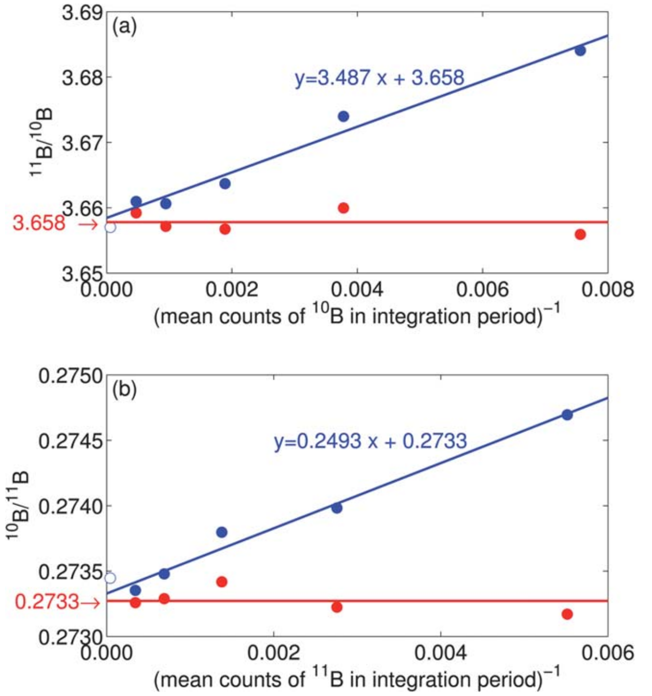

Fig. 1 Boron isotope ratios (a) $^{11}B/^{10}B$ and (b) $^{10}B/^{11}B$ as a function of the mean number of counts per integration period showing the conventionally computed ratios in blue and the novel *quasi*-unbiased ratios (eqn (15)) in red. Ratios are calculated for each integration period in the analysis and the mean value plotted. Note that the *quasi*-unbiased ratios (red) show no trend and give the desired result regardless of the integration period. All plotted points are computed from the same 128 cycles of data by summing the counts in blocks of p adjacent cycles for p = 1,2,4,8, and 16, dividing the analysis into 128/p integration periods. The greatest bias in the blue data corresponds to $p = 1$ which plots at the far right. The 'ratio of the means' is also shown (open blue symbol), which is equivalent to p = 128, close to Beale's estimator (not shown) which lies below the open symbol in both (a) and (b) by 0.0056% and 0.0035% respectively. See main text for further details. Data are raw secondary-ion mass spectrometry (SIMS) data from an analysis of a foraminifera using a CAMECA IMS 1270.

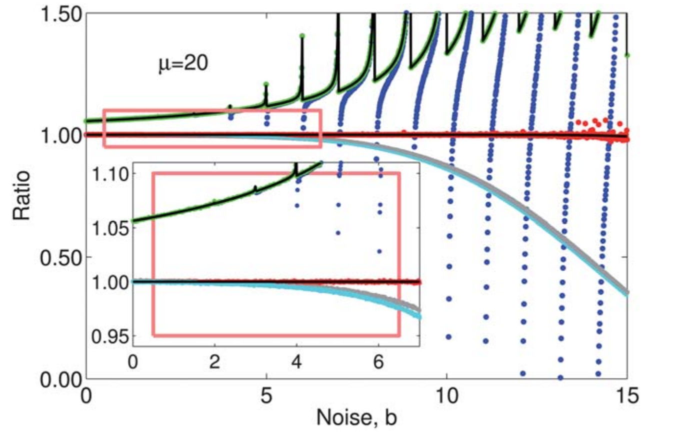

Fig. 2 Monte-Carlo simulation of $(X - b)/(Y - b)$ (blue), $(X - b)/(Z - b)$ (green), Beale's estimator (grey, see main text), simplified Beale's estimator r'_3 (cyan, eqn (18)) and the novel *quasi*-unbiased ratio, r' (red, eqn (15)) with $\mu_{x_0} = \mu_{y_0} = b$, where X and Y are independent Poisson variables with means, μ_x and μ_y, equal to 20 and $Z = Y$ for $Y > b$ and rejected otherwise. For each data point the simulation computes the mean over 10^6 samples. Rare single events where Y is small can significantly shift the mean r' for large values of b giving rise to the observed scatter in the red data for $b \gtrsim 10$. Black lines show the theoretical behaviour in the cases of the green and red data. Much of the blue data are obscured behind the green. The noise, b, is incremented in steps of 0.01.

The simulated *quasi*-unbiased ratios are closer to unity (no bias) than either of the conventional ratios or Beale's estimator over the entire range of b, although significant scatter does occur for large b (see caption). Beale's estimator (eqn (5)) on the noise-corrected data requires, for each of the 10^6 simulated ratio estimations, the variance, covariance and mean intensities to be calculated. We calculate these statistics from a dataset of n simulated data pairs, $\{x_i,y_i; i = 1\ldots n\}$, and to make the comparison with the other estimators fair we draw these from a Poisson distribution with mean value $20/n$, from which noise of b/n is subtracted. Beale's estimator is now

$$r_3 = \frac{n^{-1}\sum_{i=1}^{n} x'_i}{n^{-1}\sum_{i=1}^{n} y'_i}\left(\frac{1+\dfrac{\operatorname{cov}(x_i,y_i)}{n^{-1}\sum_{i=1}^{n} x'_i\sum_{i=1}^{n} y'_i}}{1+\dfrac{\operatorname{var}(y_i)}{n^{-1}\left(\sum_{i=1}^{n} y'_i\right)^2}}\right), \tag{16}$$

where

$$x'_i = x_i - b/n \tag{17}$$

and similarly for y'_i. A value of $n = 10$ is used for the simulation of Fig. 2. A simplified variation on Beale's estimator for Poisson data, r'_3, which is independent of n, may be obtained by substituting $n \to 1$, $\operatorname{cov}(x_i,y_i) \to 0$, $\operatorname{var}(y_i) \to Y$, $x'_i \to X - b$ and $y'_i \to Y - b$ giving

$$r'_3 = \left(\frac{X-b}{Y-b+\dfrac{Y}{Y-b}}\right). \tag{18}$$

These substitutions are arrived at by noting that, for independent Poisson distributions, $E\{\operatorname{cov}(X,Y)\} = 0$ and $E\{\operatorname{var}(Y)\} = E(Y)$. Simulations of both r_3 and r'_3 are included in Fig. 2. Note

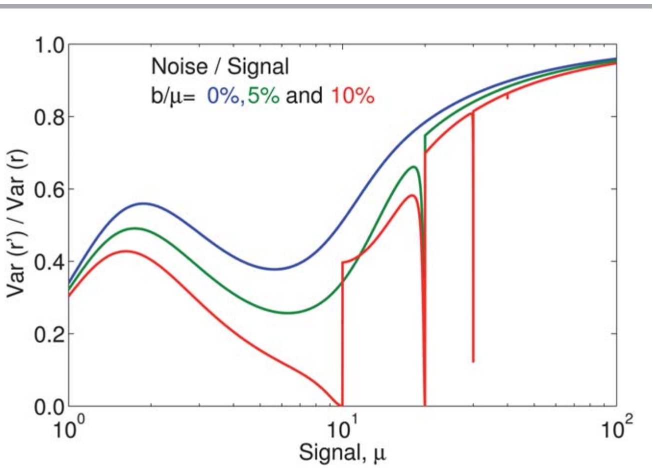

Fig. 3 Variance of *quasi*-unbiased ratio, r' (eqn (15)), divided by the variance of $r = (X - b)/(Z - b)$ as a function of mean signal, $\mu_x = \mu_y = \mu$, for relative noise, b/μ = 0, 5% and 10% showing that the variance in r' is always smaller than that of the conventional ratio, r, over this range of parameters.

that for zero noise, $r'_3 = X/(Y+1)$, *i.e.* identical to r' (eqn (15)), and for $b > 0$ the (absolute) bias of r'_3 is marginally greater (more negative) than that of r_3.

Fig. 3 compares the variance of r' (eqn (15)) with that of r (eqn (7)) as a function of mean signal, $\mu_x = \mu_y = \mu$, for three different noise to signal ratios. Over the plotted parameter range the variance of r' is always the smaller (var(r')/var(r) < 1) and, therefore, more *efficient* ratio estimator.

4 Discussion

Quasi-unbiased ratios offer advantages over conventional methods of calculating noise-corrected ratios of ion-counting measurements, namely, (i) an exponentially small statistical bias, (ii) no need to sacrifice within-analysis information by summing counts over entire analysis before taking the ratio, (iii) insensitivity to common-mode changes in signal intensity, (iv) no mathematical singularities, and (v) good stability even with low signal to noise ratios ($\gtrsim 2$).

Ratio bias has increasing importance in isotope ratio measurements since the scientific demands lead researchers to strive for ever higher precisions on small quantities of sample. This is particularly the case in secondary-ion mass spectrometry (SIMS) where, because of low blanks, very low count rates are acceptable. In studies on short-lived radionuclides (SLRs) (*e.g.* see ref. 7–10) the bias is particularly insidious as it can give rise to an apparent linear relationship between measurements of the daughter nuclide and a proxy for the parent nuclide on a so-called 'isochron' plot. To illustrate, consider the SLR ^{60}Fe which decays to ^{60}Ni. A suite of measurements are made on phases with a range of Fe/Ni ratios and an isochron plot is made of the daughter, ^{60}Ni, against the parent element, Fe. Both are plotted as ratios using some stable isotope of Ni as the denominator. A straight line with positive slope (the isochron) indicates that ^{60}Fe was present at the time the Fe and Ni were fractionated between the phases. It is sometimes the case that most or all of the variation in Fe/Ni is controlled by the Ni content which, since it appears in the denominator on both axes, will give rise to a linear relationship in the data as a consequence of the bias, adding to the positive slope due to in-growth of ^{60}Ni. Huss *et al.*[11] have recently reported this problem with some of their own published data on the Fe–Ni system concluding that in some, but not all, samples their published estimates of initial ^{60}Fe can no longer be distinguished from zero. They also discuss the likely size of any corrections to other published work on ^{10}Be, ^{26}Al and ^{53}Mn concluding that any changes to the published conclusions are small except for one older study on the Mn–Cr system in pallasites.

Note that, where internally normalised ratios are calculated, *i.e.* where a third isotope is used to correct for mass-bias, the magnitude of the statistical bias can increase or decrease, depending on the relative mass differences, by propagation of the bias of the normalising ratio. For example and assuming a linear mass bias law, in the case of ^{60}Ni/^{61}Ni normalised to ^{62}Ni/^{61}Ni the statistical bias on ^{60}Ni/^{61}Ni increases by a factor of two, whereas using ^{62}Ni as the denominator the statistical bias on ^{60}Ni/^{62}Ni stays the same magnitude but changes sign as it does in the case of ^{26}Mg/^{24}Mg normalised to ^{25}Mg/^{24}Mg. For ^{53}Cr/^{52}Cr normalised to ^{50}Cr/^{52}Cr the statistical bias increases by a factor of 1.5.

Studies where mass-bias correction is made by sample – standard bracketing are potentially susceptible to statistical bias in cases where there are differences in analyte concentration (ion count rate) between sample and standard. Standards are usually chosen to have analyte concentrations high enough that good precision can be achieved in a short time under the same analytical conditions employed on the sample. Where ion counts are higher on standards than samples and ratios are calculated conventionally, the statistical bias will result in systematically high ratios reported on samples corrected by sample-standard bracketing. In short, sample-standard bracketing does not necessarily eliminate the bias.

Huss *et al.*[11] rightly point out that ratio bias is a problem that the community will have to be aware of to avoid this source of systematic error in future work. However, we disagree that the best solution is necessarily to sum the counts over the entire analysis before taking the ratio (with or without using Beale's ratio estimator), or to correct for the bias based upon eqn (4) or (12), for the reasons (i)–(v) given at the beginning of this discussion, but rather to use eqn (15) to compute the ratio r' at each measurement cycle.

It may seem laborious to have to evaluate the Kummer confluent hypergeometric function for every measurement cycle but this should not be particularly so if (i) a good library of special functions is available to the software developers or (ii) the signal to noise ratio is sufficiently high to be able to truncate the infinite series of eqn (14) to yield an approximate value for $M(1, Y+2, \mu_0)$. The truncation error, ε_n, using an upper summation limit of $n-1$ in eqn (14) is given by,

$$\varepsilon_n = \sum_{k=n}^{\infty} \frac{\mu_0^k}{(Y+2)_k}. \tag{19}$$

Let $\alpha = \mu_0/(Y+2)$ be subject to the constraint $0 \le \alpha < 1$. This constraint is not severe: it is sufficient only that the signal is at least as large as the mean noise (and both are positive). Since $(Y+2)_k \ge (Y+2)^k$ we may write,

$$\varepsilon_n \le \sum_{k=n}^{\infty} \alpha^k \tag{20}$$

$$\le \frac{\alpha^n}{1-\alpha}. \tag{21}$$

Making n the subject of the inequality,

$$n \ge \frac{\log\{\varepsilon_n(1-\alpha)\}}{\log(\alpha)}. \tag{22}$$

Therefore, if we wish to calculate $M(1, Y+2, \mu_0)$ with a truncation error no larger than ν we can use

$$n = \left\lceil \frac{\log\{\nu(1-\alpha)\}}{\log(\alpha)} \right\rceil \tag{23}$$

and

$$M(1, Y+2, \mu_0) \approx \sum_{k=0}^{n-1} \frac{\mu_0^k}{(Y+2)_k}. \tag{24}$$

e.g. $\nu = 10^{-4}$, $Y = 10$ and $\mu_0 = 0.5$ gives $n = 3$ and truncation error, $\varepsilon_n = 6 \times 10^{-5}$, smaller than ν as required.

Appendix

A.1 The expectation value of $1/Z^l$

For Z distributed as a truncated Poisson distribution (eqn (6)) we have

$$E\left(1/Z^l\right) = N \sum_{j=l}^{\infty} \frac{a_l(j) P(y_0 + j + 1, \mu)}{\mu^j} \tag{25}$$

where,

$$P(a,z) = \gamma(a,z)/(a-1)! \quad a = 1,2,\ldots$$

is the normalised incomplete gamma function and γ is the incomplete gamma function.[3] The coefficients $a_l(j)$ are given by $l - 1$ nested summations

$$a_l(j) = j!(1/j) \sum_{j_1=l-1}^{j-1} (1/j_1) \sum_{j_2=l-2}^{j_1-1} (1/j_2) \ldots \sum_{j_{l-1}=1}^{j_{l-2}-1} (1/j_{l-1}) \tag{26}$$

or, alternatively, by $a_l(j) = j! b_l(j)$ and the recursion

$$b_{l+1}(j) = \begin{cases} 1/j & \text{for } l = 0; \\ (1/j)\sum_{k=l}^{j-1} b_l(k) & \text{for } l \geq 1. \end{cases} \tag{27}$$

The proof of eqn (25) can be subdivided into proofs of

$$1/Z^l = \sum_{j=l}^{\infty} \frac{a_l(j)}{(Z+1)_j} \tag{28}$$

and

$$E\left(1/(Z+1)_j\right) = \frac{N}{\mu^j} \cdot P(y_0 + j + 1, \mu) \tag{29}$$

where we have used Pochhammer's symbol, $(Z + 1)_j = (Z + 1)(Z + 2)\cdots(Z + j)$. The proof of eqn (29) yields easily as follows. From the probability distribution function defined in eqn (6) it follows that

$$\begin{aligned} E\left(1/(Z+1)_j\right) &= N \mathrm{e}^{-\mu} \sum_{z=y_0+1}^{\infty} \frac{\mu^z}{(z+1)_j z!} \\ &= \frac{N \mathrm{e}^{-\mu}}{\mu^j} \sum_{z=y_0+1}^{\infty} \frac{\mu^{z+j}}{(z+j)!} \\ &= \frac{N}{\mu^j} \left(1 - \mathrm{e}^{-\mu} \sum_{z=0}^{y_0+j} \frac{\mu^z}{z!}\right) \\ &= \frac{N}{\mu^j} \cdot P(y_0 + j + 1, \mu) \end{aligned}$$

where we have used Arfken[12] equation 10.70 in the final step completing the proof.

Eqn (28) yields as follows. Denote

$$\{p\}_j = 1/(p)_j.$$

Given integers $s \geq 0$ and $t \geq s$ we have

$$\sum_{s \leq z < t} \{z+1\}_j = \left(\{s+1\}_{j-1} - \{t+1\}_{j-1}\right) \Big/ (j-1) \tag{30}$$

Proof: for $s = t$, obviously RHS = LHS = 0. For $t > s$, by induction on $t - s$ using

$$\begin{aligned} \mathrm{RHS}(s,t) - \mathrm{RHS}(s,t-1) &= \left(\{t\}_{j-1} - \{t+1\}_{j-1}\right) \Big/ (j-1) \\ &= \left((t+j-1)\{t\}_j - t\{t\}_j\right) \Big/ (j-1) \\ &= \{t\}_j \\ &= \mathrm{LHS}(s,t) - \mathrm{LHS}(s,t-1). Q.E.D. \end{aligned}$$

Define

$$h_l(j) = \sum_{j_{l-1}=j+1}^{\infty} (1/j_{l-1}) \sum_{j_{l-2}=j_{l-1}+1}^{\infty} (1/j_{l-2}) \ldots \sum_{j_0=j_1+1}^{\infty} (1/j_0) \{j_0+1\}_n, \tag{31}$$

that is, l nested summations where each lower limit is the next outer summation variable. We have explicitly

$$h_l(j) = 1/n^l \{j+1\}_n \tag{32}$$

Proof: by induction on l. From the definition (eqn (31)) follows the recursion

$$h_l(j) = \begin{cases} \{j+1\}_n & \text{for } l = 0; \\ \sum_{k=j+1}^{\infty} h_{l-1}(k)/k & \text{for } l \geq 1. \end{cases} \tag{33}$$

The result is trivial for $l = 0$; for $l > 0$, substituting the RHS value for $h_{l-1}(k)$ (eqn (32)) into the recursion (eqn (33)),

$$\begin{aligned} h_l(j) &= \sum_{k=j+1}^{\infty} 1/n^{l-1} \{k+1\}_n / k = 1/n^{l-1} \sum_{k=j+1}^{\infty} \{k\}_{n+1} \\ &= 1/n^l \{j+1\}_n \end{aligned}$$

using eqn (30) with $t \to \infty$, $s \to j$, and $j \to n + 1$. *Q.E.D.*

Changing the order of summation and combining eqn (31) and (32) gives

$$\{j+1\}_n / n^l = \sum_{j_0=j+l}^{\infty} (1/j_0)\{j_0+1\}_n \sum_{j_1=j+l-1}^{j_0-1} (1/j_1) \sum_{j_2=j+l-2}^{j_1-1} (1/j_2) \ldots \sum_{j_{l-1}=j+1}^{j_{l-2}-1} (1/j_{l-1}).$$

with $j \to 0$ and substituting using the identity $n!\{j_0+1\}_n \equiv j_0!/(n+1)_{j_0}$,

$$1/n^l = \sum_{j_0=l}^{\infty} j_0!(1/j_0)/(n+1)_{j_0} \sum_{j_1=l-1}^{j_0-1} (1/j_1) \sum_{j_2=l-2}^{j_1-1} (1/j_2) \ldots \sum_{j_{l-1}=1}^{j_{l-2}-1} (1/j_{l-1}).$$

With $n \to Z$ this completes the proof of eqn (28) and, hence, also of eqn (25).

Acknowledgements

The SIMS data in Fig. 1 were collected at the Edinburgh Ion Microprobe NERC Facility on a sample provided by Dr Laura

Foster of the University of Bristol. Dr Coath is supported in this work by STFC grant ref. ST/F002734/1.

References

1 K. Coakley, D. Simons and A. Leifer, *Int. J. Mass Spectrom.*, 2005, **240**, 107–120.
2 R. C. Ogliore, G. R. Huss and K. Nagashima, *Nucl. Instrum. Methods Phys. Res., Sect. B*, 2011, **269**, 1910–1918.
3 *Digital Library of Mathematical Functions*, National Institute of Standards and Technology, 2010, ch. 8.
4 *Digital Library of Mathematical Functions*, National Institute of Standards and Technology, 2010, ch. 13.
5 A. Prudnikov, Y. Brychkov and O. Marichev, in *More Special Functions*, Gordon and Breach, 1986, vol. 3, p. 409.
6 I. Gradshteyn and I. Ryzhik, in *Tables of Integrals, Series and Products*, Academic Press, 7th edn, 2007, p. 1023.
7 K. McKeegan, M. Chaussidon and F. Robert, *Science*, 2000, **289**, 1334–1337.
8 E. Kurahashi, N. T. Kita, H. Nagahara and Y. Morishita, *Geochim. Cosmochim. Acta*, 2008, **72**, 3865–3882.
9 P. Hoppe, D. Macdougall and G. W. Lugmair, *Meteorit. Planet. Sci.*, 2007, **42**, 1309–1320.
10 S. Tachibana and G. Huss, *Astrophys. J.*, 2003, **588**, L41–L44.
11 G. Huss, R. Ogliore, K. Nagashima, M. Telus and C. Jilly, *42nd Lunar and Planetary Science Conference*, #2608, 2011.
12 G. Arfken, in *Mathematical Methods for Physicists*, Academic Press, Inc., 3rd edn, 1985, p. 566.