World Journal of Microbiology & Biotechnology **18**: 285–288, 2002.

Technical communication

Evaluation of expected absolute error affecting the maximum specific growth rate for random relative error of cell concentration

Sabyasachi Bhattacharya[1,*], Ashis SenGupta[2] and Tapas Kumar Basu[1]
[1]*Biometry Research Unit, Indian Statistical Institute, 203, B.T. Road, Calcutta 700 035, India*
[2]*Applied Statistics Unit, Indian Statistical Institute, Calcutta 700 035, India*
**Author for correspondence: Tel.: +91-33-5778085, Fax: +91-33-5776680, E-mails: ddf_pj@isical.ac.in, sabyasachi_1999@yahoo.com*

Received 16 August 2001; accepted 4 December 2001

Keywords: Expected absolute error, maximum specific growth rate, random relative error

Summary

Borzani's [(1994) *World Journal of Microbiology and Biotechnology* **10**, 475–476] idea of evaluation of absolute error affecting the 'maximum specific growth rate' (ESGR), calculated on the basis of the first and the last time points of the entire experimental time period, is generalized to the real-life situations where the relative errors of cell concentration cannot be assumed to be constant during the experiment. Visualizing the entire experimental time period as to comprise of several successive, mutually exclusive and exhaustive time intervals, we compute specific growth rates (SGRs) for each of these time intervals. Defining maximum of these SGR values as MSGR in contrast to Borzani's ESGR our aim is to study the effect of the expected absolute error on SGRs of different intervals. This will reveal the discrepancy between the true and observed MSGRs. Assuming the relative error distribution on (0,1) to be rectangular and symmetric truncated normal with mean at 0.5 and suitable variance, the expected values of the absolute errors are evaluated and numerically tabulated using the software packages MATHEMATICA and S-PLUS. Our results thus hold for situations involving varying relative errors where Borzani's results cannot be applied. A discussion with a concrete numerical example on the misidentification of the MSGR interval due to the effect of the random relative measuremental errors reveals to an experimental biologist that ignorance of this fact may lead to his/her entire experiment being futile.

Introduction

In most biological experiments involving measurements, e.g. those of cell mass concentration of yeast, longitudinal growth of fish, etc., it is not possible for the experimenter to obtain the true reading of the experimental samples. Such measurements will inevitably include errors.

Borzani (1994) evaluated the absolute error that affects the maximum specific growth (MSGR) rate (it is better to say, specific growth rate over the entire experimental time period (ESGR)) when the relative measuremental errors of cell concentration for his experiments was assumed to be constant during the exponential growth phase. To calculate absolute error affecting the ESGR he used the famous specific growth rate (SGR) formula introduced and rigorously motivated by Fisher (1921) and subsequently applied in zoological studies involving the growth of brown trout by Ball & Jones (1960). This is expressed as

$$\mathrm{SGR} = \mu = \frac{1}{\Delta t}\ln\frac{X_2}{X_1}$$

For his derivation Borzani defined

X_2 = size at last experimental point

X_1 = size at first experimental point

He then evaluated the expression for absolute error affecting his ESGR as

$$\frac{1}{2\Delta t}\ln\left[\frac{(1+\alpha)(1+\beta)}{(1-\alpha)(1-\beta)}\right]$$

Under this same assumption, related work done very recently by Gupthar *et al.* (2000), uses an alternative definition of this 'maximum specific growth rate' and calculates the amount of absolute error affecting it.

First of all we can relax the exponential growth phase assumption easily. According to Fisher (1921) whatever form a growth curve may take (may not be exponential) the mean value of the relative growth rate over a period between two experimental time points t_1, t_2 can be treated analogously to that of the original SGR, which is expressed as

$$\frac{1}{\Delta t}\int_{t_1}^{t_2}\left(\frac{1}{X_t}\frac{dX_t}{dt}\right)dt = \frac{1}{\Delta t}\int_{t_1}^{t_2} d\ \ln(X_t) = \frac{1}{\Delta t}(\ln(X_{t_2}) - \ln(x_{t_1})) \quad (1)$$

Note that Equation (1) is the same as that of the expression of SGR mentioned by Borzani (1994) under the exponential growth phase assumption between two (first and last) time points.

Note that we are dealing with the problem in a somewhat different way as compared to the approach of Borzani (1994). We first visualize our entire experimental time period to comprise of several successive, mutually exclusive and exhaustive time intervals. For each of these time intervals we compute its corresponding SGRs. Finally we obtain the maximum of these SGR values and define it as MSGR. Rather than calculating the absolute error affecting the ESGR (as done by Borzani 1994) our aim is to study the effect of the expected absolute error on SGRs of different intervals.

We know that errors committed are fully unknown to us and random in nature. Thus it is desirable to relax the above assumption of constancy of relative errors (as defined by Borzani 1994) to some probability model for their distribution, say $R(0,1)$, the rectangular distribution on (0,1). The choice of $R(0,1)$ is motivated by its simplicity and serves only as a starting point. However, probability of committing high and low errors are the same in case of $R(0,1)$. But in practical situations, the probability of committing large errors is usually lower than that for committing small errors. So to encompass this fact, we next further relax the assumption of $R(0,1)$. We adopt STN (0.5, σ^{*2}), the symmetric truncated normal density with parameters (0.5, σ^{*2}) over (0,1), obtained by truncating $N(0.5, \sigma^2)$ from the left at 0 and from the right at 1, as a model for the relative error density.

The purpose of this present communication is to find the expectation of the absolute error affecting the MSGR. The magnitude of these absolute errors reflects the extent of the reliability of the experiment in terms of its effects on the calculations of the MSGR. Such knowledge of the magnitude of absolute errors will thus be useful for the experimental scientists to infer on the plausible range of values of this growth rate based on the measurements that have been made.

Materials and methods

Nomenclature

- X_1 cell concentration at the first experimental time points
- X_1' random lower value of X_1 due to experimental errors
- X_1'' random higher value of X_1 due to experimental errors
- X_2 cell concentration at the second experimental time points
- X_2' random lower value of X_2 due to experimental errors
- X_2'' random higher value of X_2 due to experimental errors
- U random relative error that affects X_1
- V random relative error that affects X_2
- Δt exponential growth stage duration
- $\Delta\mu$ random absolute error that affects μ
- μ MSGR
- M_1 random lower value of μ due to experimental errors
- M_2 random higher value of μ due to experimental errors
- E ($\Delta\mu$) expected absolute error that affects μ
- $E(M_1)$ expected lower value of μ due to experimental errors
- $E(M_2)$ expected higher value of μ due to experimental errors

We know that the SGR formula (Ball & Jones 1960) assuming exponential for two time points (mentioned above) is defined as

$$\mu = \frac{1}{\Delta t}\ln\frac{X_2}{X_1}$$

Following Borzani, let us define random variables M_1 and M_2 as

$$M_1 = \frac{1}{\Delta t}\ln\frac{X_2'}{X_1''}, \qquad M_2 = \frac{1}{\Delta t}\ln\frac{X_2''}{X_1'}$$

where

$$X_1' = X_1(1-U)$$
$$X_1'' = X_1(1+U)$$
$$X_2' = X_2(1-V)$$
$$X_2'' = X_2(1+V)$$

With the above values of M_1 and M_2, $\Delta\mu$ is defined as

$$\Delta\mu = \frac{1}{2\Delta t}\ln\frac{(1+U)(1+V)}{(1-U)(1-V)} = \frac{1}{2\Delta t}\left[\ln\frac{(1+U)}{(1-U)} + \ln\frac{(1+V)}{(1-V)}\right]$$

Taking expectations of both sides, we get

$$E(\Delta\mu) = \frac{1}{2\Delta t}\left[E\left(\ln\frac{(1+U)}{(1-U)}\right) + E\left(\ln\frac{(1+V)}{(1-V)}\right)\right].$$
$$= \frac{1}{2\Delta t}[E_1 + E_2] \quad \text{say}$$

Case 1. *(U, V)* $\sim$ iid *R(0, 1)*

As (U, V) are iid $R(0,1)$, $\ln((1+U)/(1-U))$ and $\ln((1+V)/(1-V))$ are identically distributed with same expectations.

So, $E_1 = E_2$ implies

$$E(\Delta\mu) = \frac{1}{\Delta t}(E_1)$$

Now

$$E_1 = E\left(\ln\frac{(1+U)}{(1-U)}\right) = E(\ln(1+U)) - E(\ln(1-U))$$

$$= \int_0^1 \ln(1+u)\mathrm{d}u - \int_0^1 \ln(1-u)\mathrm{d}u \quad (\text{as } U \sim R(0,1))$$

$$= [z\ \ln z - z]_1^2 - (-1) = 2\ln 2 \quad \text{where } z = 1+u$$

$$\Rightarrow E(\Delta\mu) = \frac{2\ \ln 2}{\Delta t} = \frac{1.3863}{\Delta t}$$

So expected absolute error depends on the length of the time interval.

E_1 is reminiscent of Borzani's expression for constant absolute error as $(2\alpha/\Delta t)$ when α is less than 0.20.

Case 2. $U, V \sim$ iid STN$(0.5, \sigma^{*2})$, $0 < U, V < 1$

As in Case 1, here also E_1 and E_2 are equal.

$$\text{Now, } E_1 = E\left(\ln\frac{(1+U)}{(1-U)}\right)$$

The p.d.f. of U is

$$f^*(u) = \frac{\frac{1}{\sigma}\phi(\frac{u-0.5}{\sigma})}{2\Phi(\frac{0.5}{\sigma})-1}, \quad 0 \le u \le 1$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the p.d.f. and c.d.f. of a standard normal variate respectively, and σ^2 denotes the variance of the original normal distribution $N(0.5, \sigma^2)$ from which the above STN distribution is obtained.

The c.d.f. of u is defined as

$$F^*(u) = \frac{\frac{1}{\sigma}}{2\Phi(\frac{0.5}{\sigma})-1}\int_0^u \phi\left(\frac{t-0.5}{\sigma}\right)\mathrm{d}t = \frac{\Phi\left(\frac{u-0.5}{\sigma}\right) - \Phi\left(\frac{-0.5}{\sigma}\right)}{2\Phi(\frac{0.5}{\sigma})-1}$$

The c.d.f. of $Y = \ln[(1+U)/(1-U)]$ is

$$G^*(y) = \frac{\Phi\left(\frac{\frac{e^y-1}{e^y+1}-0.5}{\sigma}\right) - \Phi(\frac{-0.5}{\sigma})}{2\Phi(\frac{0.5}{\sigma})-1}$$

$$g^*(y) = \left(\frac{1}{2\Phi(\frac{0.5}{\sigma})-1}\right)\frac{\mathrm{d}}{\mathrm{d}y}\left[\Phi\left(\frac{\frac{e^y-1}{e^y+1}-0.5}{\sigma}\right) - \Phi\left(\frac{-0.5}{\sigma}\right)\right],$$

$$0 \le y \le \infty$$

$$= \frac{1}{\sigma(2\Phi(\frac{0.5}{\sigma})-1)}\frac{2e^y}{(e^y+1)^2}\phi\left(\frac{\frac{e^y-1}{e^y+1}-0.5}{\sigma}\right)$$

Now $E(Y) = E_1^* = E(\Delta\mu)\Delta t$

$$= \frac{1}{\sigma(2\Phi(\frac{0.5}{\sigma})-1)}\int_0^\infty \frac{2ye^y}{(e^y+1)^2}\phi\left(\frac{1}{\sigma}\frac{e^y-1}{(e^y+1)}\right)\mathrm{d}y$$

Usually in such experiments as above, scientists would have some *a priori* knowledge about the variance of the underlying relative error distribution. Accordingly, we furnish below (see Table 1) the numerical magnitude of the expected absolute errors for various plausible values of the variances of the relative error distributions that one is usually expected to encounter in real-life experimentations.

We know that if a normal density with mean μ and variance σ^2 is truncated from the left at a and from the right at b, then the mean and variance of the truncated normal distribution are respectively given by,

$$\mu^* = \mu + \sigma\frac{\phi(\frac{a-\mu}{\sigma}) - \phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}$$

$$\sigma^{*2} = \sigma^2\left[1 + \frac{(\frac{a-\mu}{\sigma})\phi(\frac{a-\mu}{\sigma}) - (\frac{b-\mu}{\sigma})\phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})} - \left(\frac{\phi(\frac{a-\mu}{\sigma}) - \phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}\right)^2\right]$$

In our present case $\mu = 0.5$, $a = 0$ and $b = 1$. So the resulting expressions are,

$$\mu^* = 0.5$$

$$\sigma^{*2} = \sigma^2\left[1 + \frac{\frac{1}{\sigma}\phi(\frac{0.5}{\sigma})}{2\Phi(\frac{0.5}{\sigma})-1}\right] \tag{2}$$

Using Equation (2), and exploiting the S-PLUS software, we calculated the values of σ for some standard *a priori* values of σ^*. These values of σ then were used to yield the corresponding values of E_1^*. Table 1 thus exhibits the influence of the variability on the expected absolute error.

Discussion

When the relative error distribution is $R(0, 1)$, the amount of absolute error that affects the MSGR is 1.3869. For STN $(0.5, \sigma^{*2})$ distribution, the maximum amount of such absolute error is attained at the $\sigma = 3.0$ (the largest value in its considered range) and this error

Table 1. E_1^* values for various σ^* and σ.

σ^*	0.5	1.0	1.5	2.0	2.5	3.0
σ	0.39455	0.63735	0.84980	1.06225	1.24435	1.42645
E_1^*	1.3114	1.3552	1.3684	1.3748	1.3778	1.3798

The calculations in the above table were done by using the MATHEMATICA software.

Table 2. Table for true and observed SGRs.

Time intervals (h)	True SGR	Erroneous values of observed SGR	Comments
(0–3)	0.4622	0.4622 − 0.4621 = 0.0001	True interval not identified
(3–6)	0.3211	0.3211 + 0.4622 = 0.7832	Wrongly identified interval
(6–9)	0.2192	0.2192 + 0.4621 = 0.6813	

is 1.3798, (see Table 1). This maximum absolute error is smaller than the absolute error for $R(0,1)$. The present paper reflects that the selection of the maximum specific growth interval is much less affected under the symmetric truncated normal as compared to the rectangular distribution.

The interval corresponding to our MSGR is of prime importance to the underlying study. Since this time interval produces the maximum growth rate, a misidentification of such an interval can lead to the entire experiment being futile. Below we show a concrete numerical example that such may be the case in real life experiments.

From Table 2 above, we see the drastic effect of real-life random errors on the identification of the vital interval with MSGR value.

This paper attempts to expose biologists to and make them aware of the dangers of ignoring the effect of random relative measuremental errors. A further deeper analysis in the future would demand the enhancement of probabilistic/statistical ideas for the development of confidence interval for MSGR, which in turn will yield the set of associated possible time intervals.

Acknowledgements

We are thankful to a Referee for a careful reading of the original version of our paper and encouragements, which have improved the presentation in this version.

References

Ball, J.N. & Jones, J.W. 1960 On the growth of brown trout of Llyn Tegid. *Proceedings of the Zoological Society of London* **134**, 1–41.

Borzani, W. 1994 A general equation for the evaluation of the errors that affects the maximum specific growth rate. *World Journal of Microbiology and Biotechnology* **10**, 475–476.

Fisher, R.A. 1921 Some remarks on the methods formulated in a recent article on 'The quantitative analysis of plant growth'. *Annals of Applied Biology* **7**, 367–372.

Gupthar, A.S., Bhattacharya, S.B. & Basu, T.K. 2000 Evaluation of maximum specific growth rate of a yeast indicating non-linear growth trend in batch culture. *World Journal of Microbiology and Biotechnology* **7**, 613–616.