ditional on X but we don't object to the notation suggested by Doctors Laska and Meisner. It is useful to note that operationally T_j can be defined as the number of subjects whose response has not terminated or been censored at time t_j.

Chao W. Chen
Robert T. O'Neill
Food and Drug Administration
Department of Health, Education and Welfare
Rockville, Maryland 20857
U.S.A.

From: Thomas B. L. Kirkwood
Statistics Section
National Institute for Biological Standards and Control
Holly Hill
London NW3 6RB, United Kingdom

To the Editor of Biometrics:

Geometric Means and Measures of Dispersion

It is exactly one hundred years since Francis Galton (1879) observed to the Royal Society of London that many variables in the life sciences, as well as other disciplines, are subject to multiplicative rather than additive variation. At the same time, Galton's colleague, Donald McAlister (1879), showed that such variables tend to follow the log-normal distribution, the mathematical properties of which have since been so comprehensively reviewed by Aitchison and Brown (1957) that there would seem little left to add. Nevertheless, there is a basic practical difficulty in the presentation of analytical results, which still has not been satisfactorily resolved. This is the need for a standard terminology allowing concise expression, in a manner easily understood by non-statisticians, of the size of the multiplicative variation.

Common practice in analysing log-normal data is to use a logarithmic transformation, so that standard normal-theory methods may be used and problems of heteroscedasticity are minimized. However, the conclusions from such analyses must be converted back into the original scales of measurement if they are to be widely understood. No particular difficulty arises with measures of location, the natural measure for a log-normal distribution being the geometric mean, equal to the anti-logarithm of the simple or arithmetic mean of the transformed distribution; this is also the median of both the transformed and untransformed distributions. But, unfortunately, there is no analogous measure of log-normal dispersion. The standard deviation of the untransformed distribution is unsatisfactory since ranges of a given number of standard deviations either side of the mean (geometric or arithmetic) are not equi-probable and do not adequately reflect the multiplicative nature of the variation. Therefore, I propose a new terminology which I have already found useful in my own work and which has been readily accepted by my non-statistical colleagues.

Following Aitchison and Brown's notation, let X be a $\Lambda(\mu, \sigma^2)$ log-normal variate, so that $Z = \log X$ is $N(\mu, \sigma^2)$. The mean, μ, and standard deviation, σ, of Z suffice to convey all necessary information about its distribution to statistician and non-statistician, alike. The geometric mean of X is e^{μ}, and let us define the *geometric standard deviation* (GSD) to be e^{σ}. (Note that it would make no difference if logarithms and anti-logarithms to base 10, or any other base, were used instead of natural logarithms, provided they were used consistently throughout.) The GSD is then a multiplicative factor such that a range for Z of $\mu \pm \sigma$ is directly equivalent to the range $(e^{\mu \pm \sigma})$ for X that is obtained by dividing and multiplying the geometric mean by the GSD. Similarly, we can define the *geometric standard error* (GSE) for a log-normally distributed estimator to be the anti-logarithm of the standard error of its log. In practice, difficulties may sometimes arise in the use of these terms, due mainly to the confusion that multiplicative factors seem to generate in those who have little facility with numbers. In these situations, it is preferable to use a term analogous to Pearson's (1896) coefficient of variation, which is the standard deviation expressed as a percentage of the mean; for Z, this is $\mu/\sigma \times 100$. Let us, therefore, define the *geometric coefficient of variation* (GCV) of X to be (GSD $-$ 1) $\times$ 100. In words, the geometric coefficient of variation of a log-normal variate X is the standard deviation of $\log X$ expressed anti-logarithmically as a percentage. A value of the GCV $= \alpha\%$ means that a range for Z of $\pm\sigma$ about μ is equivalent to a range for X of $100 \times 100/(100 + \alpha) - (100 + \alpha)\%$ about e^{μ}, or equivalently, that the limits of

$\bar{G}$ $\sqrt[n]{n_i x_i}$ $CV = \sqrt{e^{\sigma^2} - 1}$

...anal GSD = 1

$GSD = e^{SD}$ $CV = \sqrt{\exp(SD)^2 - 1} \times 100$

the range are obtained by multiplying and dividing by $(100 + \alpha)/100$. For example, $\alpha = 25$ gives a range 80 − 125%.

The advantage of these new terms is that they facilitate the presentation of analytical results for log-normal variates, while preserving awareness of the fundamental difference between them and normal ones. Non-statisticians are already familiar with means, standard deviations, standard errors and coefficients of variation, and it should be relatively simple to extend this familiarity to their geometric counterparts. It is, of course, a little more complicated to perform the geometric calculations than the corresponding arithmetic ones, but this ought not to be a serious obstacle; nor is there any way such a difficulty can be wholly avoided if the multiplicative nature of the variation is to be acknowledged.

References

Aitchison, J. and Brown, J. A. C. (1957). *The Lognormal Distribution.* Cambridge University Press, Cambridge.

Galton, F. (1879). The geometric mean, in vital and social statistics. *Proceedings of the Royal Society 29*, 365–367.

McAlister, D. (1879). The law of the geometric mean. *Proceedings of the Royal Society 29*, 367–376.

Pearson K. (1896). Regression, heredity and panmixia. *Philosophical Transactions of the Royal Society, Series A 187*, 253–318.

From: Nicholas Griffin
Department of Philosophy
McMaster University
Hamilton, Ontario
CANADA

To the Editor of Biometrics:

Bertrand Russell and Causality

I was interested to read, in a recent issue of *Biometrics*, Professor Kempthorne's critical remarks on Bertrand Russell's theories of causality (Kempthorne 1978, p. 8) and the subsequent debate between Kempthorne (1978a) and Dr. G. R. Dolby (1978). Since Russell's theory of causality grew out of a philosophical tradition it might be appropriate for a philosopher to offer a few (hopefully elucidatory) comments, notwithstanding Kempthorne's sweeping condemnation of philosophy (1978a, p. 718). The ironical thing about Professor Kempthorne's criticism of Russell is that Russell was one of the few philosophers who early this century insisted that philosophy take account of the work of scientists—an opinion for which he has frequently been criticized by followers of Wittgenstein, who elicits Professor Kempthorne's approval.

On the traditional view, as Russell (1921, p. 93) claimed, '*A* causes *B*' was interpreted as '*B* necessarily follows *A*'. As Professor Kempthorne (1978a, p. 717) notes, Russell was unable to give any meaning to 'necessary' on this theory (since it is plainly not logical necessity that is involved). What Kempthorne misses is that Russell agrees with him on this, and rejects the traditional view precisely because it's central notion of causal necessity was capable of neither logical nor empirical elucidation. The classic criticism of the traditional view had been made along exactly these lines by Hume (1739), who argued, from a strongly empiricist position, that if there was to be some necessary connection between *A* and *B* then either it must be a logical connection, or some connection empirically observable. But, as Hume went on the argue, the only *observable* connection between *A* and *B* was that *B* follows *A*. After Hume's critique all that was left of the traditional view was that '*A* causes *B*' was to be interpreted as '*B* invariably follows *A*'. Russell (1912, pp. 132–134) presented an updated version of the essentials of Hume's argument.

Russell, however, pressed Hume's critique further. When we say '*B* invariably follows *A*', *A* and *B* are plainly no longer particular events but event-types (e.g., the striking of a match) which may recur. In specifying *A* and *B*, therefore, we have to use a fairly loose description which allows for their recurrence. It follows that *A* and *B* cannot be events which require a description of the entire universe at a