

Random walk and quantitative stratigraphical sequences

Fred L. Bookstein¹ and Richard A. Reymont²

¹Center for Human Growth and Development, University of Michigan, Ann Arbor, Michigan, USA; ²Paleontologiska Institutionen, Uppsala Universitet, Uppsala, Sweden

ABSTRACT

A sequence of digitized observations on short-normal resistivity determinations seems to show trend from higher to lower values. An appropriate statistical model proves it to have less range than expected on the distribution of its successive increments. On a two-tailed statistical procedure for testing deviations from a random walk, the series tends towards 'stasis' rather than trend. The random walk model is shown to be plausible for the problem considered.

Terra Nova, 4, 147–151

INTRODUCTION

If you toss a fair coin one hundred times, will it land about 50 times with Heads up and about 50 times with Heads down? It comes to most people as a shock to learn that this is not an adequate characterization of the history of the hundred trials by any means. The coin-tossing paradigm forms the basis of our note.

Series in time occur quite naturally in the Earth Sciences, owing to the role of time in our explanations of processes. Such series are often considered to have special diagnostic value in evolutionary studies. The most immediate application of geological time-sequences lies in the search for trends: values that become greater or smaller, curves that swing to the left or the right, etc. The visual message conveyed by plots of means against time, with or without confidence intervals can, however, deceive. One univariate statistical procedure occasionally suggested for testing step-by-step shifts for significance is a chain of *t*-tests. This is logically flawed in that successive changes could be due to any of a variety of causes – ecological, taphonomical, based in collecting bias or even instrumental. We argue below that the regression of successive means against time is no less flawed. It is not that its algebra is incorrect; rather, the problem is that it has not got its 'null' and 'alternate' hypotheses straight.

The situation reviewed here may be thought of as considering whether in long series the signs of these successive changes may be thought of as having an underlying tendency towards positive or negative values. In practice, one commonly finds that there is no such tendency, which means that time cannot be regarded as a 'cause' of the phenomenon in any statistical sense – regression coefficients upon time are unreliable across subseries, so that a regression analysis has no meaning. The slopes that are more likely to be meaningful are the regressions of changes in the outcome upon changes in its known causes as independently measured along the same time-series. The

cross-products underlying these regression coefficients will average correctly over the replications of the underlying causal process. This has the consequence that one outcome of the tests we shall sketch below might be the strong recommendation that an analysis be performed using change-scores rather than absolute levels. A similar argument leads to the use of change-scores in computing the loadings representing 'concomitant variation', the joint dependence of two outcomes upon a common cause.

In order to ascertain whether or not there is significant trend in a sequence, the question must be phrased in terms of an appropriate statistical model. In most of the applied literature, that 'null' hypothesis is another 'trend', with slope precisely zero. Such an hypothesis is far too easy to reject in practice, because it is not really 'null'; it asserts the existence of a mean slope, which is to say, a tendency of the sign of increments to be either predominantly positive or predominantly negative. But such a tendency cannot just be presumed; it must be verified for geological purposes.

The null model of symmetric random walk provides a more realistic conceptual framework. This model simulates all the features we find so striking in the geological record – jumps, trends and irregular cycles (Bookstein, 1987, 1988), while still allowing the 'trend of slope zero' to be itself a *finding*, not a nullity. That is, series characterized by 'stasis'¹ and series having a genuine trend are *both* rejections of such a null hypothesis; they deviate from random walk in opposite senses. This paper reviews a recent

¹Note the term *stasis* is used in a purely descriptive manner, as describing a sequence of observations having a range significantly constrained below that expected on an hypothesis of independent increments. We do not mean by this any specific form of explanation, biological or otherwise. We might as well have coined a term such as 'subvariability', to be contrasted with the 'supervariability' that corresponds to an hypothesis of *anagenesis*. We do not find such an excess of variability in any aspect of the borehole data considered in this article.

suggestion by which a classical statistic from the probabilistic literature on random walk may be converted into a straightforward test among the three possibilities.

These remarks are not to be taken as wholly prohibiting *ad hoc* intuitive interpretations of series in time. We urge only that before you proceed with the interpretation of a sequence of observations where the only other variable is some measure of time, direct or indirect, special steps to ensure statistical validity should be taken.

The implications of the random walk model deserve to be more widely known, since an awareness of the issues involved will serve to sharpen analyses of geological sequences in time.

STRATIGRAPHICAL SEQUENCES

Sequentially ordered stratigraphical data can occur in several forms:

- (1) As frequencies of species.
- (2) As sedimentary counts.
- (3) As measurements of the properties of sediments (electrical logs, radioactivity logs, sonic logs, etc. observed down a borehole).
- (4) As morphological means assessed on the shells of some fossil species.

There are doubtless other varieties that could be put forward here. The *X*-statistic to follow applies to series of types (3) and (4), for which the model of 'independent increments' is conceivable. It does not apply to data in the form of counts that begin from zero anew at each moment of the series.

A principal interest of the stratigrapher concerned with such data is to use them for producing a curve to characterize the sequence for analytical purposes. In borehole stratigraphy, palaeoecological reconstructions can sometimes be based on the behaviour of a log (Pirson, 1977; Reyment, 1980). For present purposes, we shall assume that there is only a single set of records over time, and that the observations are perfectly accurate. Our concern is then solely with the issue of underlying parameterization.

BOREHOLE LOGS

The various methods of borehole logging aim at producing a curve for typifying a sedimentary sequence. Minor peaks and troughs of the curve yielded by some logging procedure are likely to be due to chance fluctuations, resulting from such factors as instrumentation, minor variabilities in the sediment, the composition of electrolytes, etc. The logging engineer will be on the look-out for variations registered on a larger scale, such as spikes, or major excursions, caused by marker beds or sedimentary bands with well manifested physical properties: compact limestone is one example, a coal horizon another.

The geologist hopes that the curve derived from a logging procedure will display some useful pattern. In the absence of marker beds and the like, interest will be concentrated on recognizing trends. For example, according to Pirson (1977) a trend in a self-potential log might be interpretable as indicating evidence of a regression-transgression situation.

The problem we take up in this note concerns the question of ascertaining whether a directed sequence really is indicative of a significant trend in the sense that it could not have been produced by chance alone. Until you have been confronted with the problem, and given it some thought, the whole matter may seem trite to the point of absurdity. If your eyes clearly inform you that the values are becoming greater and greater over the page, any effort wasted on abstruse statistical manoeuvres can be no more than computing statistics for their own sake. We shall prove to you that this is not so and the acceptance on face-value of trends in geology can indeed lead to self-deception. It is far too simple a solution to fit regressions to apparent trends, cycles to apparent oscillations, and catastrophes to jumps. This is merely the psychology of the compulsive gambler.

THE ANALYSIS

For the purposes of illustrating what is involved, we have analysed a sequence of digitized short-normal resistivity log-values (Fig. 1) taken from an evaluation by the Schlumberger Company for the Shell/BP Petroleum Development Company of Nigeria Ltd. The data derive from a borehole drilled in coastal western Nigeria in late Maastrichtian to early Palaeocene clay-shale. Relevant details of the stratigraphy of the region are available in Reyment (1965). The observations are electrical resistivity observations over levels 6 to 74 in the borehole; they range from 111 to 12 ohm $m^{-2} m^{-1}$. (The first five levels, all above 200 ohm m^{-2} , have been deleted from this series, because of the clear heterogeneity with the 69 values that follow them.) For these 69 values, the correlation with sequence-number in the log is -0.340 , which is nominally significant at $P = 0.003$. One might thus be tempted to conclude that these values were trending downward, that is, falling off in value over time.

We shall now demonstrate for you an alternative mode of analysis of these same data arguing that the sequence, far from trending, actually hints at *stasis* over the series of observations.

The usual test for trend is based on a model of data that are generated by independent and identically distributed normal deviates about a trend-function, linear in level. This model has three parameters: (1) the slope and (2) intercept of that trend-line, and (3) the variance of the errors around the predicted values. Our alternative model has only one parameter, a variance for the increment or decrement between each value and the next in the series. This

alternative model turns out to be that of *random walk* or *Brownian motion* across the sequence of resistivity measurements.

An invaluable reference for probabilistic concepts is the two-volume work by Feller (1957). Early on in his text, Feller (1957, Chapter 3) introduced the 'coin-toss model' for random walk for purposes that are not only expository, but philosophical: he remarked that the first time they are encountered, the theoretical conclusions are not only unexpected but actually come as a shock to intuition and common sense (Feller, 1957, p. 67).

The underlying stochastic model for our study is unusually simple.

Let Y be any random variable symmetrically distributed about 0 and having finite variance σ^2 (and hence expected value 0).

Let Y_1, \dots, Y_k, \dots be an indefinitely long series of independent realizations of Y . A *symmetric random walk* is the series

$$S_0 = 0, S_k = S_{k-1} + Y_k, k = 1, \dots$$

By the Law of Large Numbers, the distribution of any single S_n approaches $N(0, n\sigma^2)$.

Our study will not be of such single cumulative sums, however, but of the entire series

$$\{S_0, S_1, \dots, S_n\}$$

from its start to some convenient termination that is not a function of the values of S .

While the statistic demonstrated here is actually derived from the range of the series S_k , it is useful to introduce the paradoxes of random walk by reference to a different quantity, namely, the expected location of the extremum (= the maximum deviation). To do this we require the *arcsine law* which, strangely enough, was discovered only as recently as 1939 by Lévy. Soon afterwards, the derivation proved accessible by elementary methods (see Feller, 1957).

Informally, the arcsine law tells us that the probability that a random walk S_k spends a fraction of time α on the *positive* side of its starting value is proportional to $(2/\pi)\sin^{-1}\sqrt{\alpha}$. From a graph of this function, it is plain that it is much more likely to be close to 0 or 1 than to the 'expected value' of 0.5. As an equally counter-intuitive corollary, the location of the *first maximum* of the sequence S is distributed in accordance with the same law.

An exact combinatoric formula yields the probability distribution of the time of the maximum state. For the first or last as maximum, a simple approximation reduces to the formula

$$\sqrt{\left(\frac{2}{\pi n}\right)}$$

for the probability that the maximum (or minimum) will be at the very first (or last) term. For the series of 69 terms in our example, this value is $P = 0.097$ for the circumstance that the series begins with its maximum. The series does, in fact, so begin, and the value contributes with remarkable influence to the computed trend over time.

Whenever an hypothesis of independent increments is plausible, then, as applied to the random walk which ensues, any test for apparent trend is greatly inflated by the predilection of extremes of value to align with extremes of index number under this model. We need to find a test for apparent trend over time which takes this stubborn bias

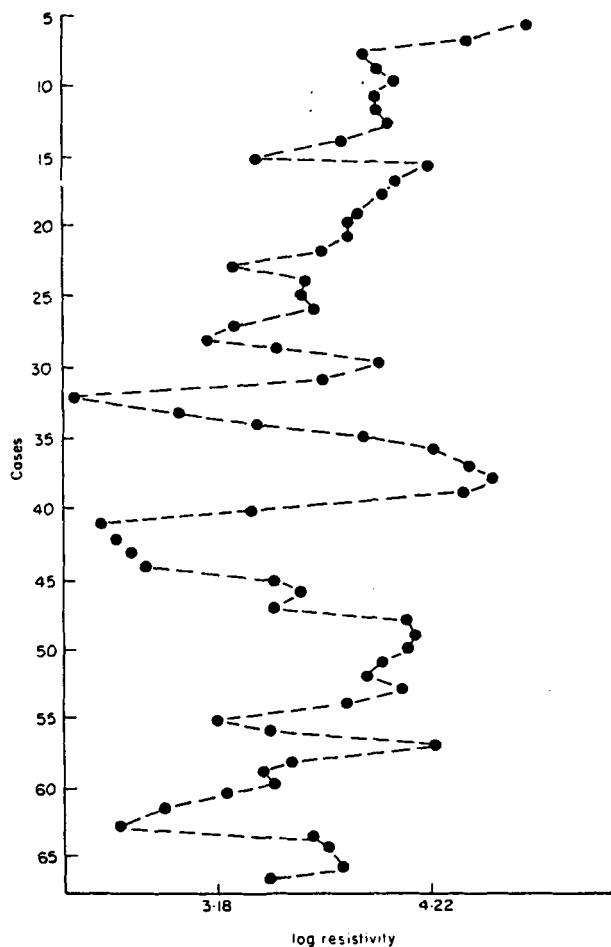


Fig. 1. Plot of 69 digitized points for the short-normal resistivity curve analysed in this paper (logarithmically transformed values).

into account, and we recommend one introduced by one of us (Bookstein, 1987, 1988, 1991) in the application to evolutionary time-series. This test is a simple modification of a distribution-theorem first suspected at the turn of the century (Bachelier's theorem of 1900), proved in 1946, and now presented as a homework exercise in graduate texts on probability theory.

Ironically, at its discovery this theorem was *rejected* by the renowned mathematician Poincaré as the subject of Bachelier's doctoral dissertation in 1900, perhaps because its inaugural application was to the fluctuation of prices on the Paris Bourse. This is neither the first nor last example of methodological break-throughs that have been thought too subversive to be considered academically acceptable; the cogency of the associated arguments is apparent only in century-long retrospect.

The significance of Bachelier's theorem can be understood intuitively by a study of Fig. 2. Along the X-axis we see a rather typical random walk - here the expression of the excess of cumulative Heads over cumulative Tails in 20,000 tosses of a fair coin. The *step-variance* of the underlying random walk is thus $\sigma^2 = 1$. We have erected a vertical scale of probability at the 20,000th step, and upon it indicated two distinct probability distributions.

To the left, according to the Law of Large Numbers, is the normal approximation to the value $S_{20,000}$ of the last step: here normal with mean 0 and variance 20,000. To the right of this vertical line is the distribution of a different quantity, the subject of Bachelier's theorem: the *maximum deviation* of the entire random walk from its starting value over the entire run. It is perhaps not obvious that such a distribution should even exist in any stable limiting form, independent of all aspects of the data except its step-variance σ^2 and the number of steps, but it does. The variable plotted here is conveniently scaled as the ratio of the

maximum excursion of the time series to the root-summed-squared step-variance. Denoting this range by

$$M = \max_k |S_k - S_0|$$

then the distribution is of the ratio

$$\frac{M}{\sqrt{(n\sigma^2)}}$$

as estimated by

$$X = \frac{M}{(\sum(S_k - S_{k-1})^2)^{1/2}}$$

regardless of the distribution of increments Y , as long as they have a variance, and dependent on the length of the series only through the extension of that summation in the denominator. In Fig. 2, the value of X is 1.35, corresponding to a tail-probability of about 35%. The walk shown is about as typical as they come, in spite of the tempting feature of a 'cut-point' between the first third and the rest, and all sorts of other 'features' all of which are meaningless. The curve drawn appears not to have any closed formula. It is constructed by summing and then differentiating an infinite series as in Bookstein (1978, 1988).

The distribution of this quantity X , the *Scaled Maximum Statistic*, may be considered to underlie a two-tailed test for deviation of empirical series from this null model in either of two interesting directions. Ranges which are large with respect to this expected excursion indicate series which deviate from random drift by way of having a *trend*, just as we thought when we were testing with the (entirely inappropriate) statistic of covariance with time. As the extremum might be found at any level along the walk, this might be considered a test for quadratic, cubic etc., shapes as well as linear. The threshold for such deviation of excursion from that expected under identically and independently distributed increments is $X > 2.25$ for the 95% significance level, 2.8 for the 99% significance level, and 3.5 for the 99.9% level of significance. (Each of these equates to the corresponding normal threshold of one-fourth the tail probability.) But there is also a tail of excursions that are *too low* to be plausible under the null model and thus represent processes that are *too highly constrained* to have risen from random drift. To these, we might well assign the label *stasis*, as they correspond to the theme of that name in the literature on evolutionary rates. The 5% tail for stasis is at a ratio 0.62 for Y ; the 1% tail, 0.49; the 0.1% tail, 0.41.

So as not to keep the reader in suspense, we turn now to the value of this statistic for our data. We compute a value of $X = 0.73$ for the series of 69 values from 111 down to 14. Inasmuch as the data are required to be positive, however, it is reasonable to convert to the logarithm as plotted. For that derived series, we find that $X \approx 0.63$, which is significant for stasis at about the 6% level. Note that this is

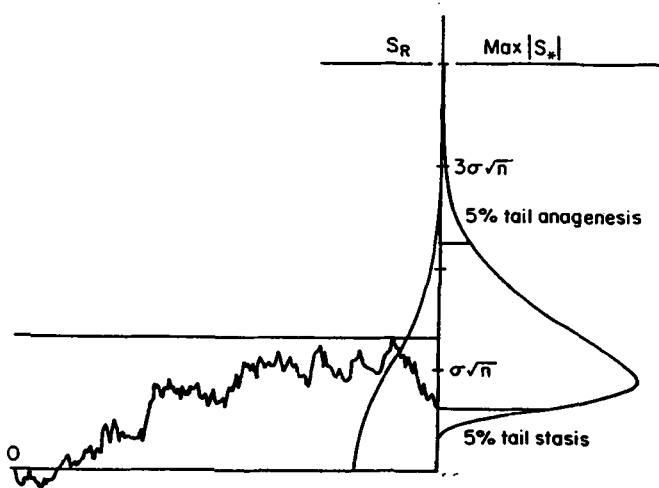


Fig. 2. The outcome of 20,000 tosses with a fair coin to illustrate the random formation of an apparent trend.

from a series that shows a regression on time which is significant at $P = 0.003$! This is a clear pointer to a stance of initial caution before announcing a finding of trend.

DISCUSSION

It is not our intention in this brief account to over-emphasize the import of this particular electrical log, nor to attempt to discover what aspects of the palaeoenvironment are responsible for its having drifted in this way. We argue only that the random walk model is plausible for these data. For example, over the last 69 terms of this series the correlation between successive increments is 0.183; for successive log-increments, 0.075. Neither of these achieves significance at even the 5% level, thus allowing us to put forward an hypothesis of independent increments. By contrast, if values of the resistivity log were actually derived from independent errors around a smooth time-sequence, these correlations would be expected to approach -0.5 , and we should be very likely to find them to be significant and negative even in a sample of 67 pairs of successive changes. Similarly, the distribution of absolute log changes (Fig. 3) appears consistent with an assumption of the stationary nature of either the changes or the log-changes over the series.

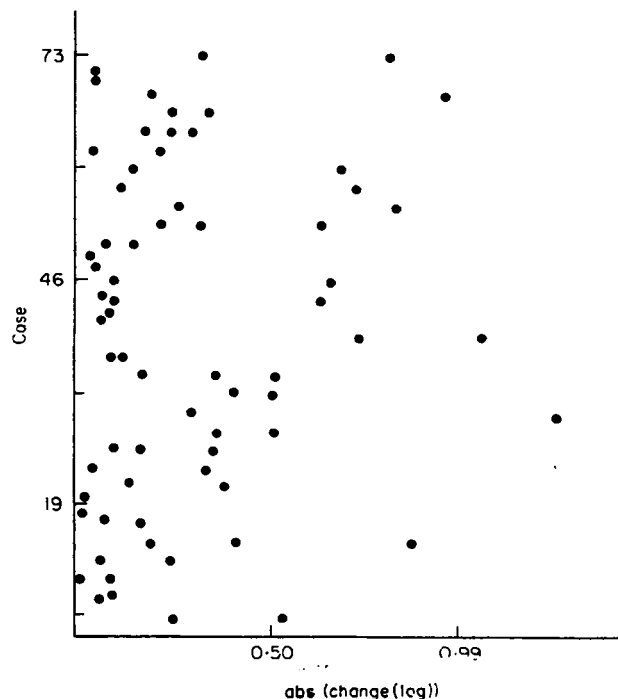


Fig. 3. Distribution of successive changes for log data. This is consistent with the property of stationarity.

We are left with only one valid interpretation: the series is *mildly* constrained, rather less trending than static. It would certainly be unwise to base inferences about palaeoecological or sedimentary regimes from its deviations.

The problem we have identified here is not restricted to explicitly observed time-trends, but extends to what Yule long ago identified as 'nonsense correlations between time series' even when the indexing is suppressed (Bookstein, 1991: pp. 398–400). Any covariance matrix among a suite of multivariate observations over time, such as the organisms of a thick series of strata or one that has been subjected to bioturbation, manifests principal components that *may* be expressing only the joint effects of these spurious time trends on more than one dimension at a time. For two outcome variables, for instance, the large-sample scatterplot of a bivariate random walk shows typical anisotropies in the range of 2.3:1, misleadingly suggesting that the long axis is reliable in its direction (Rudnick and Gaspari, 1987). Likewise, techniques of relative eigenanalysis of covariances over time-series, such as canonical variate analysis, will often be misled by these dimensions of wholly unreliable 'trend'. In our view, the only means of evading this fallacy is the concomitant observation of true causes – ecophenotypic, sedimentological, palaeoclimatological, geochemical – at the same time as the series of 'effects'. Otherwise, the treatment of non-replicable (historic) time as 'cause' cannot usually be given any coherent representation in multivariate statistical analysis. Regression slopes on time should be replaced by regressions on true causes. When patterns of apparent 'effect' are to be analysed without access to their causes, the hypothesis of independent increments must be rejected before any further haruspication should begin.

REFERENCES

- Bookstein F.L. (1987) Random walk and the existence of evolutionary rates, *Paleobiol.*, **13**, 446–464.
- Bookstein F.L. (1988) Random walk and the biometrics of morphological characters, *Evol. Biol.*, **23**, 369–398.
- Bookstein, F.L. (1991) *Morphometric Tools for Landmark Data*. Cambridge University Press, New York, xvii+435 pp.
- Feller W. (1957) *An Introduction to Probability Theory and its Applications*, Vol. 1, 2nd Edn. Wiley, New York.
- Pirson S.J. (1977) *Geologic Well Log Analysis*, 2nd Edn, Gulf Publishing Company, Houston, xiii+377 pp.
- Reyment R.A. (1965) *Aspects of the Geology of Nigeria*. Ibadan University Press, Nigeria, 145 pp.
- Reyment R.A. (1980) *Morphometric Methods in Biostratigraphy*. Academic Press, London, xi+175 pp.
- Rudnick J. and Gaspari G. (1987) The shapes of random walks, *Science*, **237**, 384–389.