Anomalous convergence of Lyapunov exponent estimates

James Theiler
Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545 and Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501

Leonard A. Smith
Mathematical Institute, University of Oxford, Oxford, OX1 3LB, United Kingdom

(Received 2 September 1994)

Numerical experiments reveal that estimates of the Lyapunov exponent for the logistic map $x_{t+1}=f(x_t)=4x_t(1-x_t)$ are anomalously precise: they are distributed with a standard deviation that scales as $1/N$, where N is the length of the trajectory, not as $1/\sqrt{N}$, the scaling expected from an informal interpretation of the central limit theorem. We show that this anomalous convergence follows from the fact that the logistic map is conjugate to a constant-slope map. The Lyapunov estimator is just one example of a "chaotic walk"; we show that whether or not a general chaotic walk exhibits anomalously small variance depends only on the autocorrelation of the chaotic process.

PACS number(s): 02.50.−r, 05.45.+b

Estimating the Lyapunov exponents for a dynamical system from a finite trajectory of length N is an important task in nonlinear time series analysis. These exponents characterize the average rate of divergence of nearby trajectories, and a positive exponent is a necessary condition for chaos. The estimate depends on the initial condition x_1 and the length N of the orbit; for one dimensional dynamics, we have

$$\hat{\lambda}(N,x_1)=\frac{1}{N}\sum_{t=1}^{N}\ln|f'(x_t)| \ . \qquad (1)$$

Oseledec's theorem [1] assures us that the limit $\lim_{N\to\infty}\hat{\lambda}(N,x)$ exists and is the same value (call it λ) for almost all x [2]. Our interest is in the rate at which $\hat{\lambda}(N,x)$ approaches λ as $N\to\infty$. In particular, we want to think of $\hat{\lambda}(N)$ as a random variable, and estimate the scaling of the variance $\mathrm{var}(\lambda(N))$ as a function of N.

Two sources of uncertainty are inherent in such an estimate: first, from a time series alone, one can only approximate the derivative $f'(x)$; second, the finite trajectory provides only a sample of the natural invariant measure. The error from both sources tends to decrease as N increases, but we will consider the second in isolation from the first, and assume that $f'(x)$ is known precisely at each point on the trajectory.

Note that if $f(x)$ is known, there may be more efficient ways to estimate the Lyapunov exponent than to just generate a length N trajectory and invoke Eq. (1). For instance, in Refs. [3], a cycle-expansion approach is used; these calculations are based on identifying and characterizing the unstable periodic orbits that are the skeleton of the attractor. If, in addition, the invariant measure is known, one may compute the exponents directly, at least for one-dimensional (1D) maps.

The central limit theorem of probability gives most physicists the intuition that the statistical (as opposed to systematic) error of an approximation based on N observations should scale as $1/\sqrt{N}$ or, equivalently, that the variance should scale as $1/N$. While this intuition is often valid for the Lyapunov exponent estimation, we will show that exceptional cases exist for which the variance exhibits $1/N^2$ scaling. We characterize a class of these exceptional cases, specificially those where the map is conjugate to a constant-slope map. We also show, for a larger class of estimators, that anomalous scaling occurs when a certain autocorrelation condition is met, and relate this result to the question of when chaotic walks mimic the properties of random walks.

We are aware of several instances in which chaos appears to violate the central limit theroem [4,5], and, in this paper, we will describe another situation in which this occurs. We are also aware of cases (such as in Ref. [6]) in which chaos successfully imitates the properties of random sequences; we will propose a necessary condition for this successful imitation.

Consider the logistic map

$$f(x)=rx(1-x) \ , \qquad (2)$$

where the parameter r is in the range $0\le r\le 4$, and estimate the Lyapunov exponent from Eq. (1) from a single orbit $x_1,x_2,\ldots,x_N$, where $x_t=f(x_{t-1})$.

The estimator of the Lyapunov exponents is an ordinary arithmetic average of values $y_t=\ln|f'(_t)|$. If the x_t (and therefore, the y_t) are independent identically distributed (IID) random variables, then it is straightforward to show that the variance should scale as $1/N$. If x_t arises from a time-autocorrelated process (such as chaos) where the autocorrelation decreases rapidly with time delay, then one might expect, from an informal interpretation of the central limit theorem, that the variance of $\hat{\lambda}(N)$ should still scale as $1/N$, possibly with a different prefactor. While this is often true, we will see that there are cases in which this prefactor is zero.

In particular, Fig. 1 shows that, for $r=4$, the $\mathrm{var}(\hat{\lambda}(N))\sim 1/N^2$. This anomalously fast convergence occurs only for special values of the map parameter r; in general, as shown in Fig. 1, the variance scales as $1/N$ for large N. Also, if some dynamical noise is added to the map, as in Fig. 2, then the ordinary $1/N$ scaling is again recovered. We note that the choice $r=4$ is not a

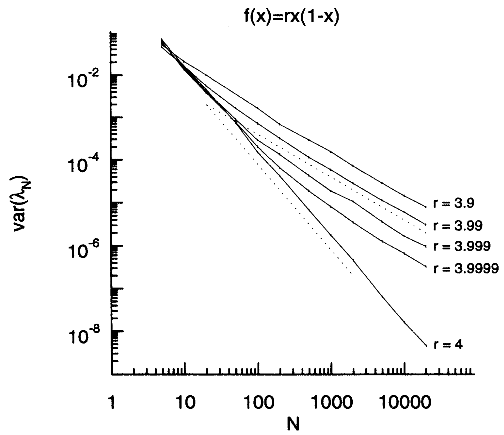

FIG. 1. Variance of an estimator of the Lyapunov exponent as a function of the number N of points in the trajectory. These are numerical estimates based on 1000 trials. Shown are various parameters r. The dotted lines exhibit slopes corresponding to $1/N$ and to $1/N^2$ scalings.

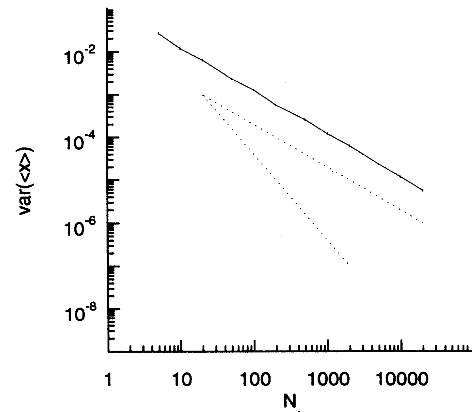

FIG. 3. Variance vs N of the estimator for $\langle x \rangle$ for a time series given by the logistic map with $r=4$. The dotted lines have slopes corresponding to $1/N$ and $1/N^2$ scalings.

sufficient condition for this special scaling; it requires a special choice of the observable y as well. For instance, we see in Fig. 3 that the estimator for $\langle x \rangle$ has a variance which scales in ordinary fashion with N.

In general, suppose that $f(x)$ is conjugate to the map $g(\theta)$, and the absolute value of the derivative $|g'|$ is constant. That is, there exists a homeomorphism h which maps θ to x such that

$$h \circ f = g \circ h : \theta \mapsto x \ . \tag{3}$$

We will further assume the h is differentiable and that the derivative is nonzero for most values of θ.

From the chain rule,

$$f'(x_t) = \frac{dx_{t+1}}{dx_t} = \left[\frac{dx_{t+1}}{d\theta_{t+1}}\right]\left[\frac{d\theta_{t+1}}{d\theta_t}\right]\left[\frac{d\theta_t}{dx_t}\right] = h'(\theta_{t+1})g'(\theta_t)/h'(\theta_t) \ . \tag{4}$$

For maps f in this class, the Lyapunov estimator is given by Eq. (1):

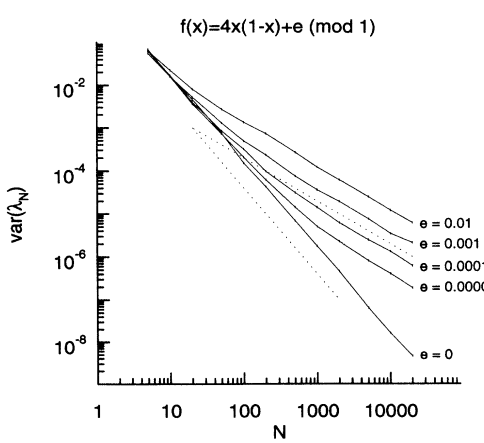

FIG. 2. Variance of an estimator of the Lyapunov exponent as a function of the number N of points in the trajectory. The map was iterated with Gaussian noise of rms amplitude e. Again, the dotted lines exhibit slopes corresponding to $1/N$ and to $1/N^2$ scalings.

$$\begin{aligned}\hat{\lambda}(N) &= \frac{1}{N}\sum_{t=1}^{N} \ln|f'(x_t)| \\ &= \frac{1}{N}\sum_{t=1}^{N} \ln|h'(\theta_{t+1})g'(\theta_t)/h'(\theta_t)| \\ &= \frac{1}{N}\sum_{t=1}^{N} \ln|g'(\theta_t)| + \frac{1}{N}\sum_{t=1}^{N} \ln|h'(\theta_{t+1})/h'(\theta_t)| \\ &= \ln|g'| + (1/N)\ln|[h'(\theta_{N+1})/h'(\theta_1)]| \ . \end{aligned} \tag{5}$$

It follows that $\ln|g'|$ is the Lyapunov exponent, and that as long as $\mathrm{var}(\ln|h'(\theta)|)$ is bounded, the variance of $\hat{\lambda}(N)$ will scale as $1/N^2$. Note that the restriction on $\mathrm{var}(\ln|h'(\theta)|)$ is not trivial: there are other homeomorphisms that conjugate the logistic map to an equivalent tent map, but are not necessarily smooth, so $\mathrm{var}(\ln|h'(\theta)|)$ is not necessarily bounded.

Not only do such maps exhibit anomalously fast convergence of the Lyapunov exponent estimator for chaotic trajectories, but a further remarkable property is that if the Lyapunov exponent is estimated from an unstable periodic orbit of the chaotic map with period p, then $\theta_1 = \theta_{p+1}$ and the Lyapunov exponent estimator given by Eq. (5) will be exact [since $h'(\theta_1) = h'(\theta_{N+1})$ for N any multiple of p].

In the case of the logistic map with $r=4$, the celebrated transform of Ulam and von Neumann [7] uses

$$h(\theta) = \sin^2\theta \tag{6}$$

to obtain

$$g(\theta) = h^{-1}\{f[h(\theta)]\} = h \circ f \circ h^{-1} \tag{7}$$

$$= \begin{cases} 2\theta & \text{for } 0 \le \theta < \pi/4 \\ \pi - 2\theta & \text{for } \pi/4 \le \theta \le \pi/2. \end{cases} \tag{8}$$

That is, $g(\theta)$ is the tent map on the interval $[0,\pi/2]$.

In this case, $h'(\theta) = 2\sin\theta\cos\theta = \sin(2\theta)$ and $|g'(\theta)| = 2$, so Eq. (5) gives

$$\hat{\lambda}(N,\theta_1) = \ln 2 + \frac{1}{N}[\ln|\sin(2^{N+1}\theta_1)| - \ln|\sin(2\theta_1)|] \ . \tag{9}$$

We see immediately that for most values of θ, $\hat{\lambda}(N,\theta)$ will be near ln2 for large N. The exceptions will be those values near $\theta_* = k\pi/2^{N+2}$ for integer k. The points θ_* are preimages of the unstable fixed point at the origin; as such they have different asymptotic dynamics, and illustrate the restriction to "almost all" x in Oseledec's theorem.

Treating the initial $\theta \in [0, \pi/2]$ as a uniform random variable, the expectation value of $\hat{\lambda}(N)$ is given by

$$\langle \hat{\lambda}(N) \rangle = \ln 2 + \frac{1}{N} \left[\frac{2}{\pi} \int_0^{\pi/2} \ln|\sin(2^{N+1}\theta)| d\theta - \frac{2}{\pi} \int_0^{\pi/2} \ln|\sin(2\theta)| d\theta \right] \quad (10)$$

or

$$\langle \hat{\lambda}(N) \rangle = \ln 2 + \frac{1}{N} [\langle \ln|\sin(2^{N+1}\theta)| \rangle - \langle \ln|\sin(2\theta)| \rangle] , \quad (11)$$

where we use the notation $\langle f(\theta) \rangle$ to denote the average $(2/\pi)\int_0^{\pi/2} f(\theta) d\theta$. Using $\theta' = 2^n \theta$, we see

$$\begin{aligned} \langle \ln|\sin(2^n\theta)| \rangle &= \frac{2}{\pi} \int_0^{\pi/2} \ln|\sin(2^n\theta)| d\theta \\ &= \frac{2}{\pi} \frac{1}{2^n} \int_0^{2^n\pi/2} \ln|\sin\theta'| d\theta' \\ &= \frac{2}{\pi} \int_0^{\pi/2} \ln|\ln\theta'| d\theta' = \langle \ln|\sin(\theta)| \rangle , \end{aligned} \quad (12)$$

which implies from Eq. (11) that $\langle \hat{\lambda}(N) \rangle = \ln 2$. The variance is given by

$$\begin{aligned} &\langle [\hat{\lambda}(N) - \ln 2]^2 \rangle \\ &= \frac{1}{N^2} \langle [\ln|\sin(2^{N+1}\theta)| - \ln|\sin(2\theta)|]^2 \rangle \quad (13) \\ &= (1/N^2)[\langle [\ln|\sin(2^{N+1}\theta)|]^2 \rangle \\ &\quad + \langle [\ln|\sin(2\theta)|]^2 \rangle \\ &\quad - 2\langle [\ln|\sin(2^{N+1}\theta)|][\ln|\sin(2\theta)|] \rangle] \quad (14) \\ &= (2/N^2)[\langle \ln|\sin\theta|)^2 \rangle \\ &\quad - \langle [\ln|\sin(2^{N+1}\theta)|][\ln|\sin(2\theta)|] \rangle] . \quad (15) \end{aligned}$$

effectively independent, and we can approximate the average of the product with a product of averages. Then

$$\begin{aligned} \mathrm{var}(\hat{\lambda}(N)) &= (2/N^2)[\langle (\ln|\sin\theta|)^2 \rangle - \langle \ln|\sin\theta| \rangle^2] \quad (16) \\ &= \frac{4}{\pi N^2} \left[\left[\int_0^{\pi/2} (\ln \sin\theta)^2 d\theta \right] - \frac{2}{\pi} \left[\int_0^{\pi/2} \ln \sin\theta \, d\theta \right]^2 \right] \quad (17) \\ &= \frac{\pi^2}{6N^2} \approx \frac{1.645}{N^2} , \quad (18) \end{aligned}$$

in agreement with the numerical simulations in Fig. 1.

Example 1. Given an arbitrary map $f(x)$, one can devise special statistics which will be estimated with anomalously small variance. A trivial construction is

$$y = f(x) - x . \quad (19)$$

It is straightforward to see that the estimator of $\langle y \rangle$,

$$\hat{y}(N) = \frac{1}{N} \sum_{t=1}^{N} y_t = \frac{1}{N} [f^N(x_1) - x_1] , \quad (20)$$

has a variance which scales like $1/N^2$.

Example 2. You do not need chaos to find a counterexample to the intuitive notion that the central limit theorem should apply to series with a finite correlation time. Let $x_0, x_1, \ldots, x_N$ be IID random variables, and let $y_t = g(x_t) - g(x_{t-1})$ for any function g. Then it is clear that y_s and y_t will be independent as long as $|t-s| \geq 2$. However,

$$\hat{y}(N) = \frac{1}{N} \sum_{t=1}^{N} y_t = \frac{1}{N} [g(x_N) - g(x_0)] \quad (21)$$

will have anomalously small variance.

The estimators we have spoken of can be rephrased in terms of general (not necessarily random) walks. These are dynamical systems of the form

$$s_t = s_{t-1} + y_t , \quad (22)$$

where $s_0 = 0$ and the y_t arises from some stationary process; they may be IID random numbers, autocorrelated random numbers, or even a chaotic time series (this last situation is referred to as a "chaotic walk" in Ref. [6]). We can rewrite the walk equation [Eq. (22)] as

$$s_t = y_1 + \cdots + y_t , \quad (23)$$

and then the estimator for $\langle y_t \rangle$ is given by $\hat{y}(N) = s_N/N$. Whenever the walk has anomalously small variance, the estimator will be anomalously precise.

When y_t are independent random numbers, then Eq. (22) is an ordinary random walk, and its variance is

$$\mathrm{var}(s_N) = N \, \mathrm{var}(y) , \quad (24)$$

so

$$\mathrm{var}(\hat{y}(N)) = \mathrm{var}(y)/N . \quad (25)$$

However, if the y_t are nontrivially autocorrelated (regardless of whether they were generated by a deterministic or a stochastic process), then this result has to be modified. In particular, if

$$A(n) = \langle y_t y_{t+n} \rangle / \langle y_t^2 \rangle , \quad (26)$$

where $\langle \; \rangle$ is an average over t, then one can write

$$\begin{aligned} \mathrm{var}(s_N) &= \langle (y_1 + \cdots + y_N)^2 \rangle \quad (27) \\ &= \sum_{i,j=1}^{N} \langle y_i y_j \rangle \quad (28) \\ &= \sum_{i,j=1}^{N} \langle y^2 \rangle A(i-j) . \quad (29) \end{aligned}$$

Of the N^2 pairs (i,j), with $1 \leq i,j \leq N$, there are $N - |n|$ of them for which $i - j = n$. Thus

$$\mathrm{var}(s_N) = \mathrm{var}(y) \left[N \sum_{n=-N}^{N} A(n) - \sum_{n=-N}^{N} |n| A(n) \right] . \quad (30)$$

We will take the $N \to \infty$ limit, and consider only those cases for which the autocorrelation of y decays rapidly with lag time. In particular, we will demand that there exist constants α and $\beta > 2$ such that $|A(n)| < \alpha |n|^{-\beta}$.

For many processes, including chaos (as long as it is mixing) and most autoregressive moving average (ARMA) [8] processes, the autocorrelaton decays exponentially, so this condition is easily met. The condition implies that the sums

$$\tau_N \equiv \sum_{n=-N}^{N} A(n) , \qquad (31)$$

$$\tau'_N \equiv \sum_{n=-N}^{N} |n| A(n) \qquad (32)$$

are bounded in the $N \to \infty$ limit. Then,

$$\text{var}(s_N) = (N\tau_N - \tau'_N)\text{var}(y) . \qquad (33)$$

Let τ and τ' be the large N limits of τ_N and τ'_N, respectively. If, in addition, τ is nonzero, then $\text{var}(s_N) \sim N\tau\,\text{var}(y)$ asymptotically for large N, and

$$\text{var}(\hat{y}(N)) \sim \text{var}(y)/(N/\tau) . \qquad (34)$$

When $\tau \gg 1$, we can interpret τ as an effective autocorrelation time inasmuch as Eq. (34) is the same as Eq. (25) with N replaced by N/τ. In other words, it is as if there were N/τ independent terms.

On the other hand, if $\tau = 0$, that is,

$$\lim_{N \to \infty} \sum_{n=-N}^{N} A(n) = 0 , \qquad (35)$$

then anomalous scaling will be observed. In particular,

$$\text{var}(\hat{y}(N)) \sim \text{var}(y)/(N^2/\tau') . \qquad (36)$$

Anomalous scaling is observed whenever the condition in Eq. (35) holds.

For the logistic map, we see in Fig. 4 that the sum which defines τ is approaching zero. Shown in the inset is the estimated autocorrelation function itself.

We have only shown how the variance scales with N. An extension of the full central limit theorem (which says that the normalized distribution will be Gaussian) to this situation is described by Dianada [9]. A vivid example of a chaotic approach to a Gaussian distribution is shown by Farmer [10] (see Fig. 2 of that paper).

The anomalous precision of the Lyanpunov exponent estimator for the logistic map at $r = 4$ arises from the conjugacy of this map to a constant-slope map. Maps which are conjugate to constant-slope maps have the additional property that Lyapunov exponent estimators obtained from (unstable) periodic orbits are exact. An extended version of the usual statement of the central limit theorem implies that the autocorrelation function of the time series of log slopes of such maps are constrained to have a correlation time of zero.

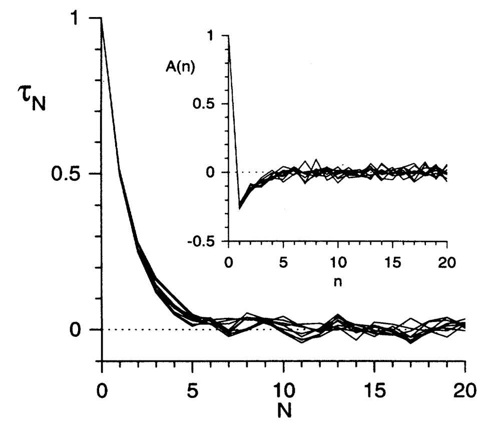

FIG. 4. The partial sum $\tau_N = 1 + 2[A(1) + \cdots + A(N)]$ vs N. Note that as N becomes large, $\tau_N \to 0$, and an anomalous scaling of the variance is expected. The inset is the autocorrelation function. Shown are results based on ten realizations of a time series with 50 000 points.

Heagy, Platt, and Hammel [6] observed experimentally that chaotic walks gave the same scaling laws as random walks; while they were not concerned particularly with the variance of the walk variable, we would suggest that those processes for which the condition in Eq. (35) holds will provide anomalous results for their system as well. Let us briefly mention that the general topic of anomalous diffusion has a rich literature; the review of fractal Brownian motions by Mandelbrot and Van Ness [11] is a classic. Geisel, Zacherl, and Radons [12] provide a recent example of a simple chaotic system that exhibits anomalous diffusion.

We are grateful to Colin Sparrow, Ronnie Mainieri, and Dean Prichard for useful commments, and to Jim Heagy for providing his manuscript prior to publication. An anonymous referee cleared up several errors in the original version of this paper. Work by J.T. is performed under the auspices of the U. S. Department of Energy, and is partially funded by a grant from the National Institute of Mental Health (No. 1-R01-MH47184); L.A.S. is supported by Pembroke College, Oxford.

[1] V. I. Oseledec, Trudy. Mosk. Mat. Obsc. [Moscow Math. Soc.] **19**, 197 (1968); cited in J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[2] Oseledec's theorem is relevant for the more general situation in which the $f'(x_t)$ are matrices instead of scalars.

[3] R. Artuso *et al.*, Nonlinearity **3**, 325 (1990). See also the focus issue on periodic orbit theory in Chaos **2**, 1 (1992); of particular interest are the articles by P. Cvitanović *et al.*, p. 85; and by R. Mainieri, p. 91.

[4] K. Kaneko, Phys. Rev. Lett. **65**, 1391 (1990); **66**, 243 (1990).

[5] J. Theiler, Phys. Rev. A **41**, 480 (1990); J. Theiler and T. Lookman, Int. J. Bif. Chaos **3**, 765 (1993).

[6] J. Heagy *et al.*, Phys. Rev. E **49**, 1140 (1993).

[7] S. Ulam and J. von Neumann, AMS Bull. **53**, 1120 (1947); cited in S. M. Ulam, *Analogies Between Analogies: the Mathematical Reports of S. M. Ulam and his Los Alamos Collaborators,* edited by A. R. Bednarek and F. Ulam (University of California Press, Berkeley, 1990), p. 337.

[8] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control* (Holden-Day, San Francisco, 1970).

[9] P. H. Diananda, Proc. Cambridge Philos. Soc. **49**, 239 (1953). We are grateful to Dean Prichard for pointing out this article to us.

[10] J. D. Farmer, Phys. Rev. Lett. **47**, 179 (1981).

[11] B. B. Mandelbrot and J. W. Van Ness, SIAM Rev. **10**, 422 (1968).

[12] T. Geisel *et al.*, Z. Phys. B Condens. Matter **71**, 117 (1988).