STOCHASTIC FORMULATION OF LOGISTIC DYNAMICS AND PARAMETER ESTIMATION BY THE LARGEST ORDER STATISTICS

Roberto N. Padua
https://orcid.org/0000-0001-5592-8867
Liceo de Cagayan University, Cagayan de Oro City, Philippines

Dionisel Y. Regalado
https://orcid.org/0000-0002-4413-8154
University of Science and Technology of Southern Philippine
Cagayan de Oro City, Philippines

Merliza F. Libao
https://orcid.org/0000-0002-7380-7951
Caraga State University
Butuan City, Philippines

ABSTRACT

The dynamical system represented by the logistic map is formulated as a stochastic process based on its periodic points both in the chaotic and non-chaotic cases. An estimator for the parameter of the logistic map is given by $\hat{\theta} = 4\, max_i \{X_i\}$. The performance of the estimator is assessed through Monte Carlo simulation although theoretical results are likewise presented in the paper.

Keywords: *beta distribution, geometric distribution, largest order statistics, logistic map*

1.0 Introduction

Applications of the evolution of dynamical systems abound in control theory such as in signal processing and control of chaos (Ott, Grebogi and Yorke, 1990). Hao and Godbole (2014) suggested using the properties of dynamical systems in predicting the largest earthquake in a given region in a year where the number of earthquakes N is random. In particular, the dynamics of the simple logistic map were studied by May (1976) in the context of the population of fruitflies:

$$X_{n+1} = \theta X_n(1 - X_n), \;\; 0 < X_n < 1, \;\;\; 0 < \theta \leq 4 \qquad (1)$$

The parameter θ is referred to as the *biotic potential* of the environment supporting the biological organism. Hayes (1997) provided an alternative stochastic formulation of (1):

$$X_{n+1} = \theta X_n(1 - X_n) + \varepsilon_n \qquad (2)$$

where $\varepsilon_i{}'s$ are iid errors. He essentially estimated θ through a weighted average of:

$$\theta_i = \frac{X_i}{X_{i-1}(1-X_{i-1})} \qquad ,i = 1, 2, \dots, N \qquad (3)$$

For many technological applications, however, it is important to provide a quick and accurate estimate of the parameter θ. A simple non-parametric estimate of this parameter is suggested in this paper using the largest order statistics of the observations. Use of the proposed estimator was tried by Padua and Lapinig (2017) in the context of rapid ecosystems appraisal of the marine fishing grounds in the Philippines.

2.0 Analytic Properties of the Logistic Map

We begin by defining a fixed point of a dynamical map.

Definition 1. (Devaney, 1997) A point x* is a fixed point of $X_{n+1} = f(x_n)$ iff $x^* = f(x^*)$.

The map (1) has fixed points at:

$$x^* = 0 \text{ and } x^* = \frac{\theta-1}{\theta} \qquad (4)$$

The stability of the fixed points depends on the Jacobian:

$$|f'(x)| = |\theta - 2\theta x| \qquad (5)$$

The fixed point $x = 0$ is an *attracting fixed point (stable)* if:

$$|f'(0)| = \theta \leq 1 \qquad (6)$$

while it *repels (unstable)* points if:

$$|f'(0)| = \theta > 1$$

Starting from any initial value x_0, then $x_1, x_2, \dots, x_n \to 0$ if $\theta \leq 1$. On the other hand, the fixed point $x = \frac{\theta-1}{\theta}$ is an attracting fixed point if:

$$\left|f'\left(\frac{\theta-1}{\theta}\right)\right| = \left|\theta - 2\theta\left(\frac{\theta-1}{\theta}\right)\right| < 1$$

$$= |-\theta + 2| < 1 \text{ or } 1 < \theta \leq 3 \qquad (7)$$

and is a repelling fixed point if:

$$\left|f'\left(\frac{\theta-1}{\theta}\right)\right| = \left|\theta - 2\theta\left(\frac{\theta-1}{\theta}\right)\right| > 1$$

$$= \theta > 3$$

Starting from any initial value, then $x_1, x_2, \ldots, x_n \to \frac{\theta-1}{\theta}$ provided $1 < \theta \leq 3$ otherwise, if $\theta \leq 1$, $x_1, x_2, \ldots, x_n \to 0$.

In the range $0 < \theta \leq 3$, the logistic map has two(2) period 1 orbits; as θ increases in the range $3 < \theta \leq 3.44904$.. the stable period 1 orbits lose their stability and new period 2 orbits take their place. Periods 4, 8, 16, 32…2^n orbits are observed as θ is progressively increased. At $\theta = 3.57$ ..., the onset of chaos, in which periodic points of all orders begin to appear, the motion of the logistic map becomes unpredictable. When $\theta = 4$, the set of all aperiodic attractors constitute a dense subset of [0,1].

Definition 2. A subset S of [0, 1] is a **dense** subset of [0, 1] iff for each $x_p \in S$ there is a point $x \in [0, 1]$ such that $|x_p - x| < \varepsilon$ for all $\varepsilon > 0$.

In a chaotic system, every point visits any sub – interval of [0, 1] a finite number of times and has periodic points of all orders including infinity.

Regardless of the value of θ > 1, however, the logistic map is maximum at $x = \frac{1}{2}$ and has the maximum value $x = \frac{\theta}{4}$. If at the nth iterate, $x_n = \frac{\theta}{4}$, then the next iterate x_{n+1} gives the minimum value of the system, that is,

$$x_n = \frac{\theta}{4} \text{ then } x_{n+1} = \frac{\theta^2}{16}(4-\theta),\ n = 0, 1, 2, \ldots \qquad (8)$$

Hence, starting from any initial value $x_0\,\epsilon\,(0,1)$, the logistic map goes through a series of transients and gets locked in the range $\frac{\theta^2}{16}(4-\theta) \leq x_n \leq \frac{\theta}{4}$. In the special case when $\theta = 4$, then the range of the iterates fall between:

$$\frac{\theta^2(4-\theta)}{16} = 0 \text{ and } \frac{\theta}{4} = 1 \qquad (9)$$

3.0 Stochastic Models and Invariant Measures

Since we cannot expect to know the chaotic dynamics precisely, we need a statistical description. From the bifurcation map it appears that for some parameter values the iterated points cover intervals of the line with some density or probability distribution. We can use this to define the "invariant measure" of the attractor. We begin by proving:

Lemma 1. Let $x_{n+1} = 4x_n(1 - x_n)$, $x_n \in [0, 1]$. Then, starting from a set of initial values $x_0 \in (0, 1)$,

$$x_n \underset{\sim}{d} \; Beta\left(\frac{1}{2}, \frac{1}{2}\right), \text{ n = 1,2,3,....}$$

Proof. Assume n is large and let

$$y = 4x(1 - x) \;\; , x \in (0, 1), \quad x \underset{\sim}{d} \cup (0, 1).$$

Let $x = \frac{1}{2}(1 - cos\pi x)$ and so:

$$Y = (1 - cos\pi x)\,(1 + cos\pi x) = sin^2\pi x. \tag{10}$$

Thus,

$$x = \frac{1}{\pi} sin^{-1}\left(\sqrt{y}\right) \tag{11}$$

The Jacobian of the transformation is:

$$J = \pm \frac{1}{2\pi\sqrt{y(1-y)}} \tag{12}$$

It follows that

$$\rho(y) = 2.\,|J| = \frac{1}{\pi\sqrt{y(1-y)}}, \;\; y \in (0, 1) \tag{13}$$

Equation (16) is the beta density with $\alpha = \frac{1}{2}, \beta = \frac{1}{2}$ ■

Lemma 1 shows that if we have a time series whose values are between 0 and 1 and $x_1 \neq x_2 \neq \cdots \neq x_n$ for all n with probability 1, then the beta density (13) reasonably approximates its probability distribution. Moreover, Lemma 1 assumes that the initial values x_0 come from dense subset S of the interval [0, 1]. The initial values constitute the set of periodic attractor of this logistic map.

We can generalize Lemma 1 as follows:

Theorem 1. Let $x_{n+1} = \theta x_n(1 - x_n)$, n = 1,2, 3…. Then, starting from a set of initial values $x_0 \in (0, 1)$, there exists an N > 0 such that for all $n \geq N$

$$x_n \underset{\sim}{d} \frac{4}{\theta} Beta\left(\frac{1}{2}, \frac{1}{2}\right) \text{ and } 0 < x_n \leq \frac{\theta}{4}$$

Proof. The periodic attractors of the logistic map lie on the range $0 < x_n \leq \frac{\theta}{4}$ for $1 < \theta \leq 4$, hence, there exists an iterate N for which all successive iterates of the logistic map are locked within this interval. Set

$$y = \theta x(1 - x), \qquad x\varepsilon\ (0,1)$$

$$x = \frac{1}{2}\,(1 - \cos\frac{\pi\theta}{4}\,x)$$

from which we obtain:

$$y = \frac{\theta}{4}\,sin^2(\frac{\pi\theta}{4}\,x)$$

or

$$x = \frac{4}{\pi\theta}\,sin^{-1}\left(\sqrt{\frac{4y}{\theta}}\right). \tag{14}$$

The Jacobian of the transformation (14) is:

$$J = \frac{8}{\pi\theta^2}\frac{1}{\sqrt{\frac{4y}{\theta}\,(1 - \frac{4y}{\theta}}}$$

from which:

$$f(y) = \frac{16}{\pi\theta^2}\frac{1}{\sqrt{\frac{4y}{\theta}\,(1 - \frac{4y}{\theta}}}, \qquad 0 < y \leq \frac{\theta}{4} \qquad (15).$$

If we let $z = \frac{4y}{\theta}$, then we can write (15) as:

$$g(z) = \frac{4}{\theta}\frac{1}{\pi\sqrt{z(1-z)}},\ 0 < z < 1, or\ the\ beta\ (.5,.5) density. \blacksquare$$

Note: Total aperiodicity is required and is implicitly assumed above. This requirement tells us roughly that the orbit smoothly fills some area, and is not concentrated at a few points. Then the measure is in some sense smooth as well. However, this is not satisfied for the logistic map when periodic orbits of finite orders are present. We require instead an invariant probability mass function as given in the next theorem.

Theorem 2. Let $\{x_t\}_{t=1}^{N}$ be a sequence of uniformly distributed random variables on (0, 1) and N be a non – negative integer – valued random variable. Suppose that $y_i = \theta x_i(1 - x_i)$, $0 < \theta < 4$.

$$g_y(y_i) = \frac{1}{N}\frac{\left(\frac{y_i}{b}\right)^{-\frac{1}{2}}\left(1-\frac{y_i}{b}\right)^{-\frac{1}{2}}}{\sum_{i=1}^{N}\left(\frac{y_i}{b}\right)^{-\frac{1}{2}}\left(1-\frac{y_i}{b}\right)^{-\frac{1}{2}}} \quad , 0 < y_i \leq b \qquad (18)$$

where $b = \frac{\theta}{4}$.

Proof. Note that

$$f_{x,i}(x_i, i) = f(x_i|i = k).P(i = k) = 1 \cdot \frac{1}{N} = \frac{1}{N} \quad ,$$
$$i = 1, 2, \dots, k$$
$$k = 1, 2, \dots, N$$

Let

$X_i = \frac{1}{2}(1 - cos\ kx_i)$ where k is a normalizing constant

$$Y_i = b\ sin^2\ kx_i$$

$$x_i = \frac{1}{k} sin^{-1}\sqrt{\frac{Y_i}{b}}$$

$$\frac{dX_i}{dY_i} = \frac{1}{2k} \cdot \frac{1}{\sqrt{\frac{Y_i}{b}\left(1-\frac{Y_i}{b}\right)}}$$

Hence,

$$g(Y_i) = \frac{1}{N} \cdot \frac{1}{k}\left(\frac{Y_i}{b}\right)^{-\frac{1}{2}}\left(1 - \frac{Y_i}{b}\right)^{-\frac{1}{2}}$$

Put $k = \sum_{i=1}^{N}\left(\frac{Y_i}{b}\right)^{-\frac{1}{2}}\left(1 - \frac{Y_i}{b}\right)^{-\frac{1}{2}}$ to obtain:

$$g(Y_i) = \frac{1}{N}\frac{\left(\frac{Y_i}{b}\right)^{-\frac{1}{2}}\left(1-\frac{Y_i}{b}\right)^{-\frac{1}{2}}}{\sum_{i=1}^{N}\left(\frac{Y_i}{b}\right)^{-\frac{1}{2}}\left(1-\frac{Y_i}{b}\right)^{-\frac{1}{2}}} \quad , 0 < Y_i \leq b. \blacksquare$$

Corollary: If $\theta = 4$, then

$$g(y) = \frac{1}{\pi\sqrt{Y(1-Y)}} \qquad , 0 < y < 1.$$

Proof: If $\theta = 4$, then $b = 1$ and $N \to \infty$ so:

$$\frac{1}{N}\sum_{i=1}^{N} Y_i^{-\frac{1}{2}}(1-Y_i)^{-\frac{1}{2}} \to \int_0^1 y^{-\frac{1}{2}}(1-y)^{-\frac{1}{2}}\,dy = \pi.$$

Hence:

$$g(y) = \frac{1}{\pi\sqrt{Y(1-Y)}} \qquad , 0 < y < 1 \qquad \blacksquare$$

The net effect of Theorem 1 is to put discrete probability mass functions to the stable fixed points of the logistic map prior to the onset of chaos.

3.0 Largest Order Statistics

For the logistic map, we have established that $0 < x_n \le \frac{\theta}{4}$ for any n. Hence, it is natural to estimate the parameter θ by:

$$\hat{\theta} = 4max_{1\le i \le n}\{x_i\} \qquad (19)$$

In turn, we require the probability distribution of:

$$y = \max\{x_1, x_2 \ldots x_n\} \qquad (20)$$

It is shown in elementary texts that .$F_Y(x) = F^n(x)$ but we deviate from this theory since we are considering that the number N of $x_i's$ is random.

Following Hao and Godbole (2014), we set up the problem as follows: Suppose $X_1, X_2 \ldots, X_n$. are i.i.d. random variables following a continuous distribution on [0, 1] with probability density and distribution functions given by f(x) and F(x) respectively. N is a random variable following a discrete distribution on {1, 2, . . .} with probability mass function given by P(N = n) = p(n), n = 1, 2, ... Let Y be given by (20) . Then the p.d.f. is derived as follows: Since

$\mathrm{P}(\mathrm{Y} \le \mathrm{y}|\mathrm{N} = \mathrm{n}) = \mathrm{F}^{\mathrm{n}}(\mathrm{y})$, we see that

$$g(y|N = n) = nF^{n-1}(y)f(y)$$

Consequently, the marginal pdf of Y is:

$$g(y) = \sum_{n=1}^{\infty} g(y|N = n)p(n) \qquad (21)$$

$$= f(y)\sum_{n=1}^{\infty} nF^{n-1}(y)p(n) \quad (22)$$

We make use of (22) under two conditions: When we know nothing about the distribution of the iterates x_n corresponding to the assumption of a uniform distribution on [0,1] and a geometric distribution on the random number N of random variables X_i. The random variable N corresponds to the "waiting time" until a periodic point of the logistic map is observed. The second condition assumes that the random variables X_i obeys the beta distribution $x_n \underset{\sim}{d} \frac{4}{\theta} Beta\left(\frac{1}{2},\frac{1}{2}\right)$. We state without proof a Theorem due to Hao and Godbole (2014):

Theorem 3 (Hao and Godbole). Let $X \sim U(0,1) and\ N \sim Geo(\rho). Let\ Y =$ $\max\{X_i\}$. Then,

$$g(Y) = \frac{\rho}{[1-(1-\rho)y]^2}$$

The probability of success ρ is proportional to the length of the interval $[\frac{\theta^2(4-\theta)}{16}, \frac{\theta}{4}]$, hence, in the case of the logistic map

$$\rho = \frac{\theta(\theta-2)^2}{16} \text{ for } \theta > 2.$$

Corollary 2. *The random variable* Y *has mean and variance given, respectively, by*

$$E(Y) = \frac{\rho(\ln\rho + \frac{1}{\rho} - 1)}{(1-\rho)^2}$$

$$V(Y) = \frac{\rho^3 - 2\rho^2 - \rho^2\, ln^2(\rho) + \rho}{(1-\rho)^4}$$

Proof. Evaluate the integrals:

$E(Y) = \int_0^1 yg(y)dy$ and

$E(Y^2) = \int_0^1 y^2 g(y)dy$

$$V(Y) = E(Y^2) - E(Y)^2$$

to obtain the results. ■

We now consider the second stochastic formulation where the inputs are random variables from a beta distribution.

Theorem 4. (Hao and Godbole, 2014)

Suppose $X \sim B\left(\frac{1}{2}, \frac{1}{2}\right)$ $and\ N \sim Geo(\rho)\ Let\ Y = \max\{X_i\}$. Then,

$$g(Y) = \frac{\rho\pi^{-1}[y(1-y)]^{-1/2}}{\left[1-(1-\rho)\frac{2}{\pi}arcsin\sqrt{y}\,\right]^2} \quad (23)$$

Proof. Substitute the appropriate probability distributions in (22) and perform the algebra.■

The mean and variance of (23), however, will be verified by simulation.

4.0 Simulation

We set up the simulation experiments as follows: Choose the parameter θ to represent the following situations: (a.) start of period doubling bifurcation $\theta = 3$; (2.) start of period 4 orbits $\theta = 3.5$; (3.)onset of chaos $\theta = 3.6$; and (4.) chaotic regime $\theta = 4$. Although the sample size N required for the two models depends on ρ, we selected equi-spaced values starting from 10 until N = 100 with different starting values $x_0 = .01, .02, .03, \ldots, .99$. One hundred simulation runs were performed for each sample size. The mean and standard deviation of the maximum order statistics were then tabulated.

Case 1: Inputs: Uniform $[\frac{\theta^2(4-\theta)}{16}, \frac{\theta}{4}]$ or $X \sim U(0,1) and\ Geo(\frac{\theta}{4} - \frac{\theta^2(4-\theta)}{16})$

Table 1. Mean and Standard Deviation of the Maximum Order Statistics for Various N

$\theta = 3$ $min = .5625$ $max = .75$			$\theta = 3.5$ $min = .3828125$ $max = .875$			$\theta = 3.6$ $min = .324$ $max = .9$			$\theta = 4$ $min = 0$ $max = 1$		
n	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**
10	.73444	.01510	10	.83482	.03413	10	.84461	.05400	10	.90956	.07190
20	.74111	.00870	20	.85058	.02260	20	.87561	.02450	20	.94225	.05861
30	.74423	.00479	30	.86003	.01366	30	.88337	.01519	30	.96809	.02789
40	.74439	.00568	40	.86208	.01298	40	.88754	.01269	40	.97451	.02375
50	.74549	.00418	50	.86538	.01071	50	.88971	.01062	50	.98204	.01585
60	.74748	.00249	60	.86632	.00861	60	.88983	.00926	60	.98531	.01475
70	.74744	.00233	70	.86875	.00554	70	.89257	.00780	70	.98630	.01290
80	.74760	.00228	80	.86957	.00581	80	.89318	.00673	80	.98741	.01348
90	.74794	.00177	90	.86987	.00485	90	.89377	.00685	90	.99114	.00905
100	.74818	.00149	100	.86931	.00631	100	.89441	.00554	100	.99003	.01017

For parameter values prior to the onset of chaos, the standard deviation of the maximum order statistics decreases at a rate of $n^{-\alpha}$ where $1 < \alpha \leq 1.5$ while under a chaotic regime, the standard deviation decreases proportional to $\frac{1}{n}$.

Table 2. Mean, Standard Deviation and Squared Bias of the Estimator of the Biotic Potential

$\theta = 3$ $min = .5625$ $max = .75$			$\theta = 3.5$ $min = .3828125$ $max = .875$			$\theta = 3.6$ $min = .324$ $max = .9$			$\theta = 4$ $min = 0$ $max = 1$		
n	**Sq. Bias**	θ	**n**	**Sq. Bias**	θ	**n**	**Sq. Bias**	θ	**n**	**Sq. Bias**	θ
10	0.0038738	2.93776	10	0.0258309	3.33928	10	0.0490888	3.37844	10	0.130870	3.63824
20	0.0012645	2.96444	20	0.0095414	3.40232	20	0.0095180	3.50244	20	0.053361	3.76900
30	0.0005327	2.97692	30	0.0035856	3.44012	30	0.0044249	3.53348	30	0.016292	3.87236
40	0.0005036	2.97756	40	0.0026708	3.44832	40	0.0024840	3.55016	40	0.010396	3.89804
50	0.0003254	2.98196	50	0.0014807	3.46152	50	0.0016941	3.55884	50	0.005161	3.92816
60	0.0001016	2.98992	60	0.0012055	3.46528	60	0.0016549	3.55932	60	0.003453	3.94124
70	0.0001049	2.98976	70	0.0006250	3.47500	70	0.0008833	3.57028	70	0.003003	3.94520
80	0.0000922	2.99040	80	0.0004718	3.47828	80	0.0007442	3.57272	80	0.002536	3.94964
90	0.0000679	2.99176	90	0.0004211	3.47948	90	0.0006210	3.57508	90	0.001256	3.96456
100	0.0000530	2.99272	100	0.0005180	3.47724	100	0.0005000	3.57764	100	0.001590	3.96012

$\theta = 3$ $min = .5625$ $max = .75$			$\theta = 3.5$ $min = .3828125$ $max = .875$			$\theta = 3.6$ $min = .324$ $max = .9$			$\theta = 4$ $min = .5625$ $max = .75$		
n	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**
10	.74413	.00955	10	.85052	.04765	10	.88020	.03745	10	.96212	.06351
20	.74827	.00271	20	.87237	.00402	20	.89649	.00486	20	.99174	.01684
30	.74905	.00177	30	.87272	.00334	30	.89737	.00385	30	.99540	.00917
40	.74963	.000597	40	.87378	.00238	40	.89848	.00302	40	.99787	.00388
50	.74950	.00105	50	.87417	.00152	50	.89884	.00168	50	.99810	.00329
60	.74976	.000468	60	.87437	.00116	60	.89926	.00165	60	.99905	.00192
70	.74984	.000249	70	.87453	.000729	70	.89956	.000887	70	.99914	.00165
80	.74986	.000259	80	.87463	.000712	80	.89965	.000771	80	.99940	.00135
90	.74987	.000288	90	.87476	.000450	90	.89962	.000734	90	.99956	.000856
100	.74991	.000196	100	.87477	.000580	100	.89972	.000590	100	.99943	.000937

References

Aharonov D., Devaney R.L., Elias U. *The Dynamics of Piecewise Linear Map and Its Smooth Approximation.* International Journal of Bifurcation and Chaos.

Aiello, B., F. Chung, and L. Lu. (2001). “A random graph model for power law graphs.” *Experimental Mathematics* 10, 53–66.

Chung, F., L. Lu, and V. Vu. (2003). “Eigenvalues of random power law graphs.” *Annals of Combinatorics* 7, 21–33.

Durrett, R. (1991). *Probability: Theory and Examples.* Wadsworth and Brooks/Cole, Pacific Grove.

Hao J., Godbole A. (2014). *Distribution of the Maximum and Minimum of Random Number of Bounded Random Variables.* http://www.arXiv:1403.1320V1[math.ST]

Johnson, N. S. Kotz, and N. Balakrishnan. (1995). *Continuous Univariate Distributions, Vol. 2*. Wiley, New York.

Kotz, S. and J. Van Dorp. (2004). *Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications.* World Scientific Publishing Co., Singapore.

Leadbetter, M. G. Lindgren, and H. Rootz´en. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer Verlag, NewYork.

May R.M. (1976). “Patterns of Dynamical Behavior in single – species Populations”. The Journal of Animal Ecology.

Ott E., Grebogi C. Yorke J. (1990).“*Controlling Chaos*”. Phys. Rev. Lett 64, 2837

Padua, R., Lapinig V. , Bolante J. and Rivera, R. (2017). *Using the dynamics of the logistic map as tool for Rapid Ecosystem Assessment (RAS)*. Journal of Higher Education Research Disciplines. Northwestern Mindanao State College of Science and Technology, Vol.2. Issue 1.