

Weighted estimation and tracking for Bienaymé Galton Watson processes with adaptive control

B. Bercu*

Laboratoire de Statistiques, Batiment 425 Mathématiques, Université de Paris-Sud, 91 405 Orsay Cedex, France

Received November 1997

Abstract

For Bienaymé–Galton–Watson processes in adaptive tracking situations, we establish asymptotic results for weighted estimators of the mean and the variance of the offspring distribution. © 1999 Elsevier Science B.V. All rights reserved

MSC: 60J80; 93E24; 60F05; 60F15

Keywords: Branching processes; Least squares; Central limit theorem; Law of iterated logarithm

1. Introduction

Consider the Bienaymé–Galton–Watson (BGW) process with adaptive control

$$X_{n+1} = \sum_{i=1}^{X_n + U_n} Y_{n,i}, \quad (1)$$

where the offspring distribution $(Y_{n,i})$ is a sequence of i.i.d. nonnegative, integer-valued random variables with finite mean m and finite variance σ^2 . The BGW process without control has been extensively studied in the literature (see e.g. Guttorp, 1991, and the references therein). The purpose of this note is to add an adaptive control to the BGW process in order to estimate the unknown parameters m and σ^2 . By the same token, the adaptive control also forces the BGW process to track, step by step, a given reference trajectory.

Let $\mathbb{F} = (\mathcal{F}_n)$ be the natural filtration of the model. The control (U_n) is a sequence of integer-valued random variables adapted to \mathbb{F} such that, whatever the value of $n \in \mathbb{N}$, $X_n + U_n \geq 1$. The initial variables X_0 and U_0 can be arbitrarily chosen. Our only assumption is that they are square-integrable and independent of $(Y_{n,i})$. Rewrite (1) as the autoregressive form

$$X_n = ma_n + \varepsilon_n, \quad a_n = X_{n-1} + U_{n-1}, \quad (2)$$

* E-mail: bernard.bercu@math.u-psud.fr.

where $\varepsilon_n = X_n - ma_n$. Clearly, (ε_n) is a martingale difference sequence adapted to \mathbb{F} with $E[\varepsilon_n^2 | \mathcal{F}_{n-1}] = \sigma^2 a_n$. In order to estimate m , we propose the weighted least-squares estimator \hat{m}_n that minimizes the quadratic criterion

$$\sum_{k=0}^n a_k^{-1} (X_k - ma_k)^2.$$

Consequently, we obviously have

$$\hat{m}_n = A_n^{-1} \sum_{k=1}^n X_k \quad \text{where } A_n = \sum_{k=1}^n a_k. \tag{3}$$

Set

$$b_n = \frac{a_n}{A_n} \quad \text{and} \quad B_n = \sum_{k=1}^n b_k.$$

We first give three useful lemmas which immediately follow from the strong law of large numbers for martingales. See Neveu (1972, Proposition 7.2.4), Dufló (1997, Theorem 1.3.24 and Theorem 4.3.16) and Wei (1987, Theorem 1).

Lemma 1. *Assume that*

$$\lim_{n \rightarrow \infty} A_n = +\infty \quad a.s. \tag{4}$$

Then, \hat{m}_n is a strongly consistent estimator of m and for all $\gamma > 0$

$$|\hat{m}_n - m|^2 = o\left(\frac{(\log A_n)^{1+\gamma}}{A_n}\right) \quad a.s.,$$

$$\sum_{k=1}^n a_k (1 - b_k) (\hat{m}_{k-1} - m)^2 = o((\log A_n)^{1+\gamma}) \quad a.s.$$

In addition, if $(Y_{n,i})$ has a finite moment of order > 2 , then

$$|\hat{m}_n - m|^2 = O\left(\frac{\log A_n}{A_n}\right) \quad a.s., \tag{5}$$

$$\frac{1}{B_n} \left(A_n (\hat{m}_n - m)^2 + \sum_{k=1}^n a_k (1 - b_k) (\hat{m}_{k-1} - m)^2 \right) \rightarrow \sigma^2 \quad a.s. \tag{6}$$

Remark 1. The assumption (4) always holds since $a_n \geq 1$ so that $A_n \geq n$.

2. Adaptive tracking

The goal of adaptive tracking is to find a control sequence (U_n) that forces the BGW process (X_n) to follow a given reference trajectory (x_n) . We assume that (x_n) is a predictable sequence of nonnegative, integer-valued random variables. We have from (2)

$$X_n - x_n = \pi_n + \varepsilon_n, \tag{7}$$

where $\pi_n = ma_n - x_n$. The performance of the tracking can be evaluated by the average weighted cost sequence (C_n) defined by

$$C_n = \frac{1}{n} \sum_{k=1}^n a_k^{-1} (X_k - x_k)^2. \tag{8}$$

The adaptive tracking is said to be optimal on average if C_n converges a.s. to σ^2 . In order to estimate the variance σ^2 , we can also propose

$$\Gamma_n = \frac{1}{n} \sum_{k=1}^n a_k^{-1} (X_k - \hat{m}_{k-1} a_k)^2, \quad \Sigma_n = \frac{1}{n} \sum_{k=1}^n a_k^{-1} (X_k - \hat{m}_k a_k)^2.$$

Lemma 2. *If $\Delta_n = \frac{1}{n} \sum_{k=1}^n a_k^{-1} \varepsilon_k^2$, then we always have*

$$|C_n - \Delta_n| = O\left(\frac{1}{n} \sum_{k=1}^n a_k^{-1} \pi_k^2\right) \quad \text{a.s.} \tag{9}$$

In addition, if $(Y_{n,i})$ has a finite moment of order > 2 , then

$$\liminf_{n \rightarrow \infty} C_n \geq \sigma^2 \quad \text{a.s.} \tag{10}$$

with equality if and only if $\sum_{k=1}^n a_k^{-1} \pi_k^2 = o(n)$ a.s.

Lemma 3. *Assume that (4) is satisfied with $\limsup_{n \rightarrow \infty} (a_n/A_n) < 1$ a.s. Then, for all $\gamma > 0$*

$$|\Gamma_n - \Delta_n| = o\left(\frac{(\log A_n)^{1+\gamma}}{n}\right), \quad |\Sigma_n - \Delta_n| = o\left(\frac{(\log A_n)^{1+\gamma}}{n}\right) \quad \text{a.s.}$$

In addition, if $(Y_{n,i})$ has a finite moment of order > 2 , then

$$|\Gamma_n - \Delta_n| = O\left(\frac{\log A_n}{n}\right), \quad |\Sigma_n - \Delta_n| = O\left(\frac{\log A_n}{n}\right) \quad \text{a.s.} \tag{11}$$

Remark 2. *If $(Y_{n,i})$ has a finite moment of order $\beta > 2$, then we have by Chow’s Theorem (Chow, 1965) together with Kronecker’s Lemma*

$$|\Delta_n - \sigma^2| = o\left(\frac{\lambda_n}{n}\right) \quad \text{a.s.}$$

where (λ_n) is a positive deterministic sequence such that

$$\sum_{n=1}^{\infty} \left(\frac{1}{\lambda_n}\right)^{\beta/2} < +\infty.$$

We can choose, for example, $\lambda_n = n^\delta$ with $2\beta^{-1} < \delta < 1$. Moreover, if $(Y_{n,i})$ has a finite moment of order > 4 , then

$$|\Delta_n - \sigma^2|^2 = O\left(\frac{\log n}{n}\right) \quad \text{a.s.}$$

The choice of the adaptive control sequence (U_n) is crucial. From relation (7), if the parameter m were known, we would choose U_n such that π_{n+1} be as close as possible to zero i.e. $U_n = P(m^{-1}x_{n+1}) - X_n$ where P denotes

the projection operator on \mathbb{N} . Therefore, we propose to make use of the adaptive tracking control

$$U_n = \begin{cases} 1 - X_n & \text{if } P(\hat{m}_n^{-1}x_{n+1}) = 0, \\ P(\hat{m}_n^{-1}x_{n+1}) - X_n & \text{otherwise.} \end{cases} \tag{12}$$

3. Main results

Theorem 4. Assume that $(Y_{n,i})$ has a finite moment of order > 2 and that (x_n) converges a.s. to an integer $x \geq 0$. If we use the adaptive control given by (12), then \hat{m}_n is a strongly consistent estimator of m

$$|\hat{m}_n - m|^2 = O\left(\frac{\log n}{n}\right) \quad a.s. \tag{13}$$

In addition, if $\alpha = \max(1, P(m^{-1}x))$, then we have the central limit theorem

$$\sqrt{n}(\hat{m}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\sigma^2}{\alpha}\right), \tag{14}$$

the law of iterated logarithm

$$\limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n}\right) (\hat{m}_n - m)^2 = \frac{\sigma^2}{\alpha} \quad a.s. \tag{15}$$

and the quadratic strong law

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n (\hat{m}_k - m)^2 = \frac{\sigma^2}{\alpha} \quad a.s. \tag{16}$$

Theorem 5. Assume that $(Y_{n,i})$ has a finite moment of order > 2 and that (x_n) converges a.s. to an integer $x \geq 0$. If we use the adaptive control given by (12), then Γ_n and Σ_n are both strongly consistent estimators of σ^2

$$|\Gamma_n - \Delta_n| = O\left(\frac{\log n}{n}\right), \quad |\Sigma_n - \Delta_n| = O\left(\frac{\log n}{n}\right) \quad a.s. \tag{17}$$

More precisely

$$\lim_{n \rightarrow \infty} \frac{n}{\log n} (\Gamma_n - \Delta_n) = \sigma^2, \quad \lim_{n \rightarrow \infty} \frac{n}{\log n} (\Sigma_n - \Delta_n) = -\sigma^2 \quad a.s. \tag{18}$$

Therefore, assume that $(Y_{n,i})$ has a finite moment of order > 4 . Denote by τ^4 the fourth-order centered moment of $(Y_{n,i})$ and set $\rho = \alpha^{-1}\tau^4 + (2 - 3\alpha^{-1})\sigma^4$. Then we have, for (Γ_n) as for (Σ_n) , the central limit theorem

$$\sqrt{n}(\Gamma_n - \sigma^2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \rho), \tag{19}$$

the law of iterated logarithm

$$\limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n}\right) (\Gamma_n - \sigma^2)^2 = \rho \quad a.s. \tag{20}$$

and the quadratic strong law

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n (\Gamma_k - \sigma^2)^2 = \rho \quad a.s. \tag{21}$$

Remark 3. First, if $(Y_{n,i})$ has only a finite moment of order 2, then (13) and (17) hold replacing the convergence rates $O(\log n/n)$ by $o((\log n)^{1+\gamma}/n)$ for all $\gamma > 0$. Next, the tracking is residually optimal since (C_n) converges a.s. to $\sigma^2 + \xi$ where $\xi = \alpha^{-1}(m\alpha - x)^2$ which differs from zero except for $x \in m\mathbb{N}^*$. Therefore, in order to obtain the tracking optimality, we have to require more on (x_n) .

Theorem 6. Assume that $(Y_{n,i})$ has a finite moment of order > 2 . Choose the reference trajectory (x_n) such that

$$x_n = \hat{m}_{n-1}z_n, \tag{22}$$

where (z_n) is a predictable sequence of positive integer-valued random variables which converges a.s. to a positive integer z . If we use the adaptive control given by (12), then \hat{m}_n is a strongly consistent estimator of m and relations (13)–(16) hold replacing α by z . In addition, (17) and (18) are also valid and the tracking is optimal

$$|C_n - A_n| = O\left(\frac{\log n}{n}\right) \quad \text{a.s.} \tag{23}$$

More precisely

$$\lim_{n \rightarrow \infty} \frac{n}{\log n} (C_n - A_n) = \sigma^2 \quad \text{a.s.} \tag{24}$$

Finally, assume that $(Y_{n,i})$ has a finite moment of order > 4 . Then, relations (19)–(21) hold for (C_n) , (Γ_n) and (Σ_n) replacing α by z in the definition of ρ .

4. Proofs

Proof of Theorem 4. We always have $a_n \geq 1$ so that $A_n \geq n$ and (\hat{m}_n) converges a.s. to m . Therefore, (a_n) and (A_n/n) both converge a.s. to $\alpha = \max(1, P(m^{-1}x))$. Thus, (13) immediately follows from (5). Moreover, assume that $(Y_{n,i})$ has a finite moment of order $\beta > 2$. From (1) and Rosenthal’s inequality (see Petrov, 1995, Theorem 2.12), we have for all $n \geq 1$, $E[|\varepsilon_n|^\beta | \mathcal{F}_{n-1}] = O(a_n^{\beta/2})$ so that

$$\sum_{k=1}^n E[|\varepsilon_k|^\beta | \mathcal{F}_{k-1}] = O(n) \quad \text{a.s.}$$

Consequently, the Lindeberg’s condition is satisfied (see Duflo, 1997, p. 48). Hence, we deduce (14) from the martingale central limit Theorem (see Brown, 1971, Theorem 2). In addition, (15) follows from the martingale law of iterated logarithm (see Stout, 1970, Theorems 1 and 2 and Duflo, 1997, Corollary 6.4.25). Furthermore, (nb_n) converges a.s. to 1 so that $B_n \sim \log n$ a.s. Therefore, since

$$A_n(\hat{m}_n - m)^2 = o(\log n) \quad \text{a.s.}$$

we obtain from (6) that

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n a_k(\hat{m}_{k-1} - m)^2 = \sigma^2 \quad \text{a.s.} \tag{25}$$

which directly implies (16). \square

Proof of Theorem 5. We have seen that (A_n/n) converges a.s. to α . Hence, (11) immediately implies (17). On the one hand, we have

$$X_n - \hat{m}_{n-1}a_n = a_n(m - \hat{m}_{n-1}) + \varepsilon_n.$$

Consequently,

$$n(\Gamma_n - \Delta_n) = \sum_{k=1}^n a_k(\hat{m}_{k-1} - m)^2 + R_n, \tag{26}$$

where $R_n = o(\log n)$ a.s. by (25). Therefore, again by (25) and (26), we obtain the first convergence of (18). On the other hand, we also have

$$X_n - \hat{m}_n a_n = a_n(1 - b_n)(m - \hat{m}_{n-1}) + (1 - b_n)\varepsilon_n.$$

Consequently,

$$n(\Sigma_n - \Delta_n) = \sum_{k=1}^n a_k(1 - b_k)^2(\hat{m}_{k-1} - m)^2 - 2P_n + Q_n + R_n \tag{27}$$

where $R_n = o(\log n)$ a.s. by (25) and

$$P_n = \sum_{k=1}^n a_k^{-1} b_k \varepsilon_k^2, \quad Q_n = \sum_{k=1}^n a_k^{-1} b_k^2 \varepsilon_k^2.$$

From Chow’s Theorem (Chow, 1965) together with Kronecker’s Lemma, we have

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} P_n = \sigma^2 \quad \text{a.s.} \tag{28}$$

In addition, (Q_n) converges a.s. to a finite positive random variable. Therefore, since $b_n = o(1)$ a.s., the second convergence of (18) follows from (25), (27) and (28). Furthermore, set

$$M_n = n(\Delta_n - \sigma^2) = \sum_{k=1}^n W_k, \quad W_n = a_n^{-1} \varepsilon_n^2 - \sigma^2.$$

From (1), $E[\varepsilon_n^4 | \mathcal{F}_{n-1}] = a_n \tau^4 + 3a_n(a_n - 1)\sigma^4$ so that $E[W_n^2 | \mathcal{F}_{n-1}] = a_n^{-1} \tau^4 + (2 - 3a_n^{-1})\sigma^4$ where τ^4 is the fourth order centered moment of $(Y_{n,i})$. Thus, (M_n) is a square-integrable martingale with increasing process $(\langle M \rangle_n)$ such that

$$\lim_{n \rightarrow \infty} \frac{\langle M \rangle_n}{n} = \rho \quad \text{a.s.},$$

where $\rho = \alpha^{-1} \tau^4 + (2 - 3\alpha^{-1})\sigma^4 > 0$. If $(Y_{n,i})$ has a finite moment of order $\beta > 4$, we also have from (1) and Rosenthal’s inequality

$$\sum_{k=1}^n E[|W_k|^{\beta/2} | \mathcal{F}_{k-1}] = O(n) \quad \text{a.s.}$$

Consequently, as in the proof of Theorem 4, Lindeberg’s condition is satisfied. Hence, we obtain from the martingale central limit Theorem

$$\frac{1}{\sqrt{n}} M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \rho), \quad \sqrt{n}(\Delta_n - \sigma^2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \rho). \tag{29}$$

Relation (19) clearly follows from (17) and (29). In addition, by the martingale law of iterated logarithm

$$\limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right) (\Delta_n - \sigma^2)^2 = \rho \quad \text{a.s.} \tag{30}$$

Thus, we directly obtain (20) by (17) and (30). Finally, from Duflo (1997, Theorem 4.3.16) together with (30), we find that

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left(\frac{M_k}{\langle M \rangle_k} \right)^2 = \rho^{-1}, \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left(\frac{M_k}{k} \right)^2 = \rho \quad \text{a.s.}$$

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n (\Delta_k - \sigma^2)^2 = \rho \quad \text{a.s.} \quad (31)$$

Therefore, (18) and (31) imply (21) completing the proof of Theorem 5. \square

Proof of Theorem 6. We always have from (12) together with (22), $a_n = z_n$ so that (a_n) and (A_n/n) both converge a.s. to z . Thus, we prove Theorem 6 exactly as before except for relations involving (C_n) . In fact, from (2) and (22), those relations have been already proved since $X_n - x_n = X_n - \hat{m}_{n-1} a_n$ so that $C_n = \Gamma_n$. \square

References

- Brown, B.M., 1971. Martingale central limit theorems. *Ann. Math. Statist.* 42, 59–66.
- Chow, Y.S., 1965. Local convergence of martingale and the law of large numbers. *Ann. Math. Statist.* 36, 552–558.
- Duflo, M., 1997. *Random Iterative Models*. Springer, Berlin.
- Guttorp, P., 1991. *Statistical Inference for Branching Processes*. Wiley, New York.
- Neveu, J., 1972. *Martingales à Temps Discret*. Masson, Paris.
- Petrov, V.V., 1995. *Limit Theorems of Probability Theory*. Oxford University Press, Oxford.
- Stout, W.F., 1970. A martingale analogue of Kolmogorov's law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie* 15, 279–290.
- Wei, C.Z., 1987. Adaptive prediction by least squares predictors in stochastic regression models with applications to time series. *Ann. Statist.* 15, 1667–1682.