Available online at www.sciencedirect.com
SCIENCE DIRECT®

C. R. Acad. Sci. Paris, Ser. I 339 (2004) 663–666

COMPTES RENDUS MATHEMATIQUE

http://france.elsevier.com/direct/CRASS1/

Statistics/Probability Theory

Estimation of the offspring mean of a supercritical or near-critical size-dependent branching process

Nadia Lalam [a], Christine Jacob [b]

[a] *EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*
[b] *Unité de biométrie, INRA, 78352 Jouy-en-Josas cedex, France*

Received 20 January 2004; accepted after revision 1 September 2004

Presented by Paul Deheuvels

Abstract

We consider a single-type supercritical or near-critical size-dependent branching process $\{N_n\}_n$ such that the offspring mean converges to a limit $m \geqslant 1$ with a rate of convergence of order N_n^α as the population size N_n grows to ∞ and the variance may change at the rate N_n^β, where $\alpha > 0$ and $-1 \leqslant \beta < 1$. The offspring mean depends on an unknown parameter θ_0 that we estimate on the non-extinction set by using the conditional least squares method. We prove the strong consistency of the estimator of θ_0 as $n \to \infty$ under some general conditions on the asymptotic behavior of the process. We also give its asymptotic distribution for a certain class of size-dependent branching processes. ***To cite this article: N. Lalam, C. Jacob, C. R. Acad. Sci. Paris, Ser. I 339 (2004).***

Résumé

Estimation de la descendance moyenne d'un processus de branchement taille-dépendant supercritique ou presque-critique. On considère un processus de branchement taille-dépendant unitype $\{N_n\}_n$ supercritique ou presque-critique tel que sa descendance moyenne converge vers une limite $m \geqslant 1$ à une vitesse de l'ordre de N_n^α lorsque l'effectif de la population N_n tend vers l'infini et tel que sa variance évolue à la vitesse N_n^β où $\alpha > 0$ et $-1 \leqslant \beta < 1$. La descendance moyenne dépend d'un paramètre inconnu θ_0 que l'on estime sur l'ensemble de non-extinction du processus à l'aide de la méthode des moindres carrés conditionnels. On démontre la consistance forte de l'estimateur de θ_0 quand $n \to \infty$ sous des hypothèses générales concernant le comportement asymptotique du processus. On donne aussi sa distribution asymptotique pour une certaine classe de processus taille-dépendants. ***Pour citer cet article : N. Lalam, C. Jacob, C. R. Acad. Sci. Paris, Ser. I 339 (2004).***

E-mail addresses: lalam@eurandom.tue.nl (N. Lalam), cj@banian.jouy.inra.fr (C. Jacob).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2004.09.001

1. Introduction

We are concerned with the estimation of the offspring mean of a supercritical or near-critical size-dependent single-type branching process in discrete time. For a survey of inference for branching processes, see [1]. The size N_{n+1} of the population at generation $n+1$ is defined by

$$N_{n+1}=\sum_{i=1}^{N_n} Y_{n+1,i}, \tag{1}$$

where $Y_{n+1,i}$ is the offspring size at generation $n+1$ of the i-th individual from generation n. The initial size N_0 is given and we assume that, conditionally to $F_n=\sigma(N_0,\dots,N_n)$, the $\{Y_{n+1,i}\}_i$ are independently and identically distributed (i.i.d.) with $0<p_{N0}+p_{N1}<1$, for all $N\geqslant 1$, where $\{p_{Nj};\ j\in\mathbb{N}\}$ is the offspring distribution, when the population size equals N, and with the offspring mean $m(N)$ and variance $\sigma^2(N)$ when the population size equals N satisfying:

A1: for all $N\geqslant 1$, $m(N)=m+f(N)$; $|f(N)|\leqslant KN^{-\alpha}$, $K<\infty$, $\alpha>0$, $m\geqslant 1$;

A2: for all $N\geqslant 1$, $\sigma^2(N)\leqslant\sigma^2N^{\beta}$; $0<\sigma^2<\infty$, $-1\leqslant\beta<1$;

A3: when $m=1$, for all $N\geqslant 1$, $m(N)>1$, the function $N\mapsto m(N)$ decreases to 1, $N\mapsto\sigma^2(N)/(N^2(m(N)-1))$ is ultimately decreasing and satisfies $\int_1^\infty\sigma^2(x)/(x^2(m(x)-1))\,\mathrm{d}x<\infty$.

The size-dependent branching process is called supercritical when $m>1$ and near-critical when $m=1$.

We assume that $m(\cdot)$ depends on a finite dimensional and identifiable parameter θ_0 that we estimate and on a nuisance part depending on a parameter ν_0 that may be of infinite dimension. Denote $m(N)$ by $m_{\theta_0,\nu_0}(N)$.

The estimator $\hat\theta_{h,n,\gamma,\nu}$ will minimize over θ the conditional least squares based on $n-h+1$ successive observations:

$$\widetilde S_{h,n,\gamma,\nu}(\theta)=\sum_{k=h+1}^{n}\bigl(N_k-m_{\theta,\nu}(N_{k-1})N_{k-1}\bigr)^2N_{k-1}^{-\gamma}, \tag{2}$$

where $m_{\theta,\nu}(\cdot)$ is the offspring mean in which θ is unknown, ν has a given value that may depend on n, and $\gamma\in\mathbb{R}$. One of the advantage of conditional least squares methodology is that it does not require the knowledge of the exact law of $\{Y_{n,i}\}_{i,n}$ unlike methods based on the likelihood. The asymptotic properties of $\{\hat\theta_{h,n,\gamma,\nu}\}_n$ will be studied on the set of non-extinction $\varepsilon_\infty=\{\overline{\lim}_{n\to\infty}N_n\neq 0\text{ a.s.}\}$ by increasing n to ∞, with either h or $n-h$ fixed.

From Eq. (1), $\{N_n\}_n$ follows the stochastic nonlinear regressive model $N_{n+1}=m(N_n)N_n+\eta_{n+1}$, where the noise $\eta_{n+1}=\sum_{i=1}^{N_n}[Y_{n+1,i}-m(N_n)]$ is a martingale difference satisfying $E(\eta_{n+1}^2|F_n)\leqslant\sigma^2N_n^{1+\beta}$. Estimation in this setting has been studied by Lai [5] and Skouras [8] but their methods require conditions that are difficult to check in practice. In order to get the consistency, we will rather rely on a minimum contrast method [9] and the rate of convergence will be obtained thanks to a central limit theorem [7].

2. Strong consistency of $\{\hat\theta_{h,n,\gamma,\nu}\}_n$ in the general model (A1, A2, A3)

Let $(\theta_0,\nu_0)\in\mathring{\Theta}\times\mathbb{R}^{d_2}$, where Θ is a compact set in $\mathbb{R}^{d_1}$, $d_1\in\mathbb{N}$, $d_2\in\mathbb{N}$ when ν is of finite dimension or $d_2=\mathbb{N}$ when ν is of infinite dimension.

Let $B_\delta^c=\{\theta=(\theta_1,\dots,\theta_{d_1})\in\Theta:\sum_{k=1}^{d_1}|\theta_k-\theta_{0,k}|\geqslant\delta\}$, where $\delta>0$.

Let $v(N)=N^{\psi}$, where $\psi\in\mathbb{R}$ and define the two semi-norms $\|u(\cdot)\|_{n,\infty}=\sup_{h+1\leqslant k\leqslant n}|u(N_{k-1})|$ and $\|u(\cdot)\|_n^2=\sum_{k=h+1}^n u^2(N_{k-1})N_{k-1}^{2(1-\psi)-\gamma}\bigl[\sum_{k=h+1}^n N_{k-1}^{2(1-\psi)-\gamma}\bigr]^{-1}$ for some function $u(\cdot)$.

Theorem 2.1. *Assume the following conditions on the non-extinction set* ε_∞:

(1) B1: *for all* $\delta>0$, $\underline{\lim}_{n\to\infty}\inf_{\theta\in B_\delta^c}\|(m_{\theta_0,\nu}(\cdot)-m_{\theta,\nu}(\cdot))v(\cdot)\|_n\overset{\text{a.s.}}{\neq}0$;

B2: $\overline{\lim}_{n\to\infty}\sup_{\theta\in B_\delta^c}\|(m_{\theta_0,\nu}(\cdot)-m_{\theta,\nu}(\cdot))v(\cdot)\|_{n,\infty}<\infty$ *a.s.* B1 *and* B2 *imply that* θ_0 *is uniformly asymptotically identifiable in* $m_{\theta_0,\nu}(\cdot)$ *at the exact rate* $v(\cdot)$ *for the semi-norm* $\{\|\cdot\|_n\}_n$;

(2) B3: $\overline{\lim}_{n\to\infty}\|(m_{\theta_0,\nu_0}(\cdot)-m_{\theta_0,\nu}(\cdot))v(\cdot)\|_n\overset{a.s.}{=}0$, *i.e.* $(m_{\theta_0,\nu_0}(\cdot)-m_{\theta_0,\nu}(\cdot))v(\cdot)$ *is asymptotically negligible*;

(3) B4: (i) $\{\sum_{k=h+1}^n N_{k-1}^{2(1-\psi)-\gamma}\}_n$ *increases a.s. to* ∞, *as* $n\to\infty$;

(ii) *there exists* h_0 *such that* $\sum_{k=h_0+1}^\infty N_{k-1}^{2[2(1-\psi)-\gamma]}N_{k-1}^{\beta+2\psi-1}(\sum_{l=h+1}^k N_{l-1}^{2(1-\psi)-\gamma})^{-2}<\infty$ *a.s.* (*take* $h_0=h$ *when* h *is fixed*);

(4) B5: *for all* $\delta>0$, *for all* N, $\sup_{\theta\in B_\delta^c}(m_{\theta_0,\nu}(N)-m_{\theta,\nu}(N))$ *is attained for some* $\theta_N^{\sup}$ (*respectively,* $\inf_{\theta\in B_\delta^c}(m_{\theta_0,\nu}(N)-m_{\theta,\nu}(N))$ *is attained for some* $\theta_N^{\inf}$).

Then, $\{\hat\theta_{h,n,\gamma,\nu}\}_n$ *is strongly consistent.*

Proof. The proof (detailed in [6]) relies on a sufficient condition concerning the minimum contrast method [9] and on the martingale difference property of $\{N_k-m_{\theta_0,\nu_0}(N_{k-1})N_{k-1}\}_k$ [2]. □

Remark 1. Let $a_0=1$ and for $n\geqslant 1$, $a_n=m^n$ in the supercritical case, $a_n=a_{n-1}m(a_{n-1})$ in the near-critical case and let $W_n=N_na_n^{-1}$. Under A1, A2 and A3, there exists an integrable random variable W such that

$$\lim_{n\to\infty}W_n\overset{\text{a.s.}}{=}W;\quad P(W>0)>0;\quad \varepsilon_\infty=\{W>0\text{ a.s.}\}=\Bigl\{\lim_{n\to\infty}N_n\overset{\text{a.s.}}{=}\infty\Bigr\}. \tag{3}$$

In the near-critical case, $W\overset{a.s.}{=}1$ (see Klebaner [4] for the proof in the supercritical case and Kersting [3] in the near-critical one). Thanks to Toeplitz lemma, relationship (3) implies that we can replace $\{N_{k-1}\}_k$ by $\{a_{k-1}\}_k$ in assumption B4.

Remark 2. When $\gamma=1+\beta$ and h is fixed, B4(ii) is always satisfied (see page 158 of [2]).

3. Strong consistency and rate of convergence in a subclass of model (A1, A2, A3)

We study the properties of $\hat\theta_{h,n,\gamma,\nu}-\theta_0$ properly normalized under the more accurate model ($\widetilde{A}1$, $\widetilde{A}2$, $A3$) which is a particular case of (A1, A2, A3), where $\widetilde{A}1$: for all $N\geqslant 1$, $m_{\theta,\nu}(N)=m+\theta N^{-\alpha}+r_{\theta,\nu}(N)$, where m and α are assumed known and $r_{\theta,\nu}(N)=0(N^{-\bar\alpha})$ with $\bar\alpha>\alpha$;

$\widetilde{A}2$: when $m>1$, for all $N\geqslant 1, \sigma^2(N)=\sigma^2N^\beta-r_+(N)$, and when $m=1$, for N large enough, $\sigma^2(N)=\sigma^2N^\beta-r_+(N), r_+(N)\geqslant 0, r_+(N)=\mathrm{o}(N^\beta)$, $0<\sigma^2<\infty$ and $-1\leqslant\beta<1$.

Let $B_n=\sqrt{\sum_{k=h+1}^n a_{k-1}^{1+\beta+2(1-\alpha)-2\gamma}}$, $D_n=\sum_{k=h+1}^n a_{k-1}^{2(1-\alpha)-\gamma}$ and $\Phi_{h,n,\gamma}=B_nD_n^{-1}$.

Let the assumptions

C1: (i) $\overline{\lim}_{N\to\infty}\sup_\theta|r_{\theta,\nu}(N)|N^{\bar\alpha}<\infty$;

(ii) B5, i.e. for all $\delta>0$, for all N, $\sup_{\theta\in B_\delta^c}(m_{\theta_0,\nu}(N)-m_{\theta,\nu}(N))$ is attained for some $\theta_N^{\sup}$ (respectively $\inf_{\theta\in B_\delta^c}(m_{\theta_0,\nu}(N)-m_{\theta,\nu}(N))$ is attained for some $\theta_N^{\inf}$);

C2: (i) $\{\sum_{k=h+1}^n a_{k-1}^{2(1-\alpha)-\gamma}\}_n$ increases to ∞, as $n\to\infty$;

(ii) there exists h_0 such that $\sum_{k=h_0+1}^\infty a_{k-1}^{2[2(1-\alpha)-\gamma]+[\beta+2\alpha-1]}(\sum_{l=h+1}^k a_{l-1}^{2(1-\alpha)-\gamma})^{-2}<\infty$;

C3: (i) for all N, $\theta\mapsto r_{\theta,\nu}(N)$ is twice continuously differentiable in a neighborhood of θ_0 and there exist $M'<\infty$, $M''<\infty$ such that $\sup_\theta|r'_{\theta,\nu}(N)|N^{\bar\alpha}\leqslant M'$ and $\sup_\theta|r''_{\theta,\nu}(N)|N^{\bar\alpha}\leqslant M''$, for all N;

(ii) for all N, $\sup_\theta r''_{\theta,\nu}(N)$ is attained for some $\theta_N^{\sup}$ (resp. $\inf_\theta r''_{\theta,\nu}(N)$ is attained for some $\theta_N^{\inf}$);

C4: when h is fixed, $\lim_{n\to\infty}B_n=\infty$;

C5: if $\nu\neq\nu_0$, $\lim_{n\to\infty}[\sum_{k=h+1}^n a_{k-1}^{2-\gamma-\bar\alpha-\alpha}]B_n^{-1}=0$;

C6: for all $x \in \mathbb{R}$, $\lim_{H\to\infty}\lim_{n\to\infty}\sup_{H+1\leqslant k\leqslant n} E(R_k^2 1_{\{R_k^2\geqslant B_{k,n}^2 x^2\}}|F_{k-1}) \overset{\text{a.s.}}{=} 0$, where $R_k = [Y_{k,1} - m_{\theta_0,\nu_0}(N_{k-1})] \times m'_{\theta_0,\nu}(N_{k-1}) W_{k-1}^{1-\gamma} a_{k-1}^{\frac{1}{2}(-\beta+2\alpha)}$ and $B_{k,n} = B_n a_{k-1}^{\frac{1}{2}[-\beta-2(1-\alpha)+2\gamma]}$ (Lindeberg condition type).

According to the linear Taylor series expansion applied to the first derivative of the contrast (2):

$$\Phi_{h,n,\gamma}^{-1}(\hat{\theta}_{h,n,\gamma,\nu} - \theta_0) = P_{h,n,\gamma,\nu}[Q_{h,n,\gamma,\nu}]^{-1}, \tag{4}$$

where $P_{h,n,\gamma,\nu} = [\sum_{k=h+1}^{n}\sum_{i=1}^{N_{k-1}}[Y_{k,i} - m_{\theta_0,\nu}(N_{k-1})]m'_{\theta_0,\nu}(N_{k-1})N_{k-1}^{1-\gamma}]B_n^{-1}$, $Q_{h,n,\gamma,\nu} = [m'^2_{\theta_n,\nu}(N_{k-1}) + m_{\theta_n,\nu}(N_{k-1})m''_{\theta_n,\nu}(N_{k-1}) - N_k N_{k-1}^{-1} m''_{\theta_n,\nu}(N_{k-1})]N_{k-1}^{2-\gamma} D_n^{-1}$ and $\theta_n \in\,]\min(\hat{\theta}_{h,n,\gamma,\nu}, \theta_0), \max(\hat{\theta}_{h,n,\gamma,\nu}, \theta_0)[$; the symbol $'$ designates the derivative with respect to θ.

Theorem 3.1. *Under* C1 *and* C2 *on* ε_∞, $\lim_{n\to\infty}\hat{\theta}_{h,n,\gamma,\nu} \overset{a.s.}{=} \theta_0$. *Assume in addition* C3 *to* C6 *on* ε_∞, *then*

$$\lim_{n\to\infty} P_{h,n,\gamma,\nu} \overset{d}{=} P, \quad E\big[\exp(\mathrm{i}tP)\big] = E\left[\exp\left(-\frac{t^2}{2}\sigma^2 W^{1+\beta+2(1-\alpha)-2\gamma}\right)\right],$$
$$\lim_{n\to\infty} Q_{h,n,\gamma,\nu} \overset{a.s.}{=} W^{2(1-\alpha)-\gamma}.$$

The best rate of convergence is attained for $\gamma = 1+\beta$, *and for this value of* γ, *under* $\beta + 2\alpha < 1$, *in the supercritical case,* $\Phi_{h,n,\gamma}^{-1} = 0(a_n^{\frac{1-\beta-2\alpha}{2}})$, *and in the near-critical case,* $\Phi_{h,n,\gamma}^{-1} > a_h^{\frac{1-\beta-2\alpha}{2}}\sqrt{n-h}$. *Under* $\beta + 2\alpha = 1$, $\Phi_{h,n,\gamma}^{-1} = \sqrt{n-h}$.

Proof. Strong consistency is a consequence of Theorem 2.1 and Remark 1. We establish that the asymptotic distribution of $\{P_{h,n,\gamma,\nu}\}_n$ is a mixture of Gaussian laws by using the asymptotic behavior of the process (3) and a central limit theorem for random sums (Rahimov [7]). The proof of the a.s. limit of $\{Q_{h,n,\gamma,\nu}\}_n$ relies on (3) and a strong law of large numbers (Hall and Heyde [2]). Details are given in [6]. □

Remark 3. Following the same reasoning, Theorem 3.1 also holds when estimating the unknown limit parameter m in the supercritical model $m(N) = m + 0(N^{-\alpha})$. The only amendment to make is to replace α (respectively $\bar{\alpha}$) by 0 (respectively α) in conditions C1–C6 and in the statement of the theorem. Since (m, θ) is not asymptotically identifiable, we cannot estimate simultaneously m and θ (see [6]).

References

[1] P. Guttorp, Statistical Inference for Branching Processes, Wiley Ser. Probab. Math. Statist., Academic Press, New York, 1991.

[2] P. Hall, C.C. Heyde, Martingale Limit Theory and its Application, Probab. Math. Statist., Academic Press, New York, 1980.

[3] G. Kersting, Some properties of stochastic difference equations, in: P. Tautu (Ed.), Stochastic Modelling in Biology, World Scientific, Singapore, 1990, pp. 328–339.

[4] F.C. Klebaner, On population-size dependent branching processes, Adv. Appl. Probab. 16 (1984) 30–55.

[5] T.L. Lai, Asymptotic properties of nonlinear least squares estimates in stochastic regression models, Ann. Statist. 22 (1994) 1917–1930.

[6] N. Lalam, C. Jacob, Estimation of the offspring mean in a supercritical or near-critical size-dependent branching process, Technical report, 2003.

[7] I. Rahimov, Random Sums and Branching Stochastic Processes, Lecture Notes in Statist., Springer-Verlag, New York, 1995.

[8] K. Skouras, Strong consistency in nonlinear stochastic regression models, Ann. Statist. 28 (2000) 871–879.

[9] C.F. Wu, Asymptotic theory of nonlinear least squares estimation, Ann. Statist. 9 (1981) 501–513.