

FLUCTUATION LIMIT OF BRANCHING PROCESSES WITH IMMIGRATION AND ESTIMATION OF THE MEANS

M. ISPÁNY* ** AND

G. PAP,* *** University of Debrecen

M. C. A. VAN ZUIJLEN,**** Radboud University Nijmegen

Abstract

We investigate a sequence of Galton–Watson branching processes with immigration, where the offspring mean tends to its critical value 1 and the offspring variance tends to 0. It is shown that the fluctuation limit is an Ornstein–Uhlenbeck-type process. As a consequence, in contrast to the case in which the offspring variance tends to a positive limit, it transpires that the conditional least-squares estimator of the offspring mean is asymptotically normal. The norming factor is $n^{3/2}$, in contrast to both the subcritical case, in which it is $n^{1/2}$, and the nearly critical case with positive limiting offspring variance, in which it is n .

Keywords: Conditional least-squares estimator; fluctuation limit; nearly critical Galton–Watson branching process with immigration; Ornstein–Uhlenbeck-type process

2000 Mathematics Subject Classification: Primary 62F12

Secondary 60J80; 60J60; 62M05

1. Introduction

Let $\{\xi_{k,j}, \varepsilon_k : k, j \in \mathbb{N}\}$ be independent, nonnegative integer-valued random variables such that $\{\xi_{k,j} : k, j \in \mathbb{N}\}$ and $\{\varepsilon_k : k \in \mathbb{N}\}$ are identically distributed. Recursively define

$$X_k = \sum_{j=1}^{X_{k-1}} \xi_{k,j} + \varepsilon_k \quad \text{for } k \in \mathbb{N}, \quad X_0 = 0. \quad (1.1)$$

The sequence $(X_k)_{k \in \mathbb{Z}_+}$ is called a *branching process with immigration*. We can interpret X_k as the size of the k th generation of a population, where $\xi_{k,j}$ is the number of offspring of the j th individual in the $(k-1)$ th generation and ε_k is the number of immigrants contributing to the k th generation. Assume that $m := E \xi_{1,1}$, $\lambda := E \varepsilon_1$, $\sigma^2 := \text{var } \xi_{1,1}$, and $b^2 := \text{var } \varepsilon_1$ are finite. The cases $m < 1$, $m = 1$, and $m > 1$ are referred to respectively as *subcritical*, *critical*, and *supercritical*. For $k \in \mathbb{Z}_+$, let \mathcal{F}_k denote the σ -algebra generated by $\{X_0, X_1, \dots, X_k\}$. Then, by (1.1),

$$E(X_k \mid \mathcal{F}_{k-1}) = mX_{k-1} + \lambda, \quad k \in \mathbb{N}.$$

Received 10 August 2004; revision received 10 December 2004.

* Postal address: Faculty of Informatics, University of Debrecen, Pf. 12, H-4010 Debrecen, Hungary.

** Email address: ispany@inf.unideb.hu

*** Email address: papgy@inf.unideb.hu

**** Postal address: Faculty of Mathematics, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. Email address: m.vanzuijlen@science.ru.nl

Clearly,

$$M_k := X_k - E(X_k \mid \mathcal{F}_{k-1}) = X_k - mX_{k-1} - \lambda, \quad k \in \mathbb{N},$$

defines a martingale difference sequence $(M_k)_{k \in \mathbb{N}}$ with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{Z}_+}$. Moreover, we obtain the recursion

$$X_k = mX_{k-1} + \lambda + M_k \quad \text{for } k \in \mathbb{N}. \tag{1.2}$$

In the critical case $m = 1$, Wei and Winnicki [16] considered the random step functions

$$\mathcal{X}^{(n)}(t) := X_{\lfloor nt \rfloor}, \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N},$$

as random elements in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$, where $\lfloor \cdot \rfloor$ denotes the lower integer part, and proved the weak convergence

$$\frac{1}{n} \mathcal{X}^{(n)} \xrightarrow{D} \mathcal{X}, \quad \text{as } n \rightarrow \infty, \tag{1.3}$$

in $D(\mathbb{R}_+, \mathbb{R}_+)$, where $(\mathcal{X}(t))_{t \in \mathbb{R}_+}$ is a (nonnegative) diffusion process with generator A such that

$$Af(x) = \lambda f'(x) + \frac{1}{2} \sigma^2 x f''(x), \quad f \in C_c^\infty(\mathbb{R}_+),$$

and $\mathcal{X}(0) = 0$. Here $C_c^\infty(\mathbb{R}_+)$ is the space of infinitely differentiable functions on \mathbb{R}_+ that have compact support. The process $(\mathcal{X}(t))_{t \in \mathbb{R}_+}$ can also be characterized as the (unique) solution to the stochastic differential equation

$$d\mathcal{X}(t) = \lambda dt + \sigma \sqrt{\mathcal{X}(t)} dW(t), \quad t \in \mathbb{R}_+,$$

with initial condition $\mathcal{X}(0) = 0$, where $(W(t))_{t \in \mathbb{R}_+}$ is a standard Wiener process.

For each $n \in \mathbb{N}$, let $\{\xi_{k,j}^{(n)}, \varepsilon_k^{(n)} : k, j \in \mathbb{N}\}$ be independent, nonnegative integer-valued random variables such that $\{\xi_{k,j}^{(n)} : k, j \in \mathbb{N}\}$ and $\{\varepsilon_k^{(n)} : k, j \in \mathbb{N}\}$ are identically distributed. In this paper, we consider a sequence of branching processes with immigration $(X_k^{(n)})_{k \in \mathbb{Z}_+}, n \in \mathbb{N}$, given by the recursion

$$X_k^{(n)} = \sum_{j=1}^{X_{k-1}^{(n)}} \xi_{k,j}^{(n)} + \varepsilon_k^{(n)} \quad \text{for } k, n \in \mathbb{N}, \quad X_0^{(n)} = 0. \tag{1.4}$$

Assume that $m_n := E \xi_{1,1}^{(n)}, \lambda_n := E \varepsilon_1^{(n)}, \sigma_n^2 := \text{var } \xi_{1,1}^{(n)}$, and $b_n^2 := \text{var } \varepsilon_1^{(n)}$ are finite for all $n \in \mathbb{N}$. The sequence (1.4) is said to be *nearly critical* if $m_n \rightarrow 1$ as $n \rightarrow \infty$. Now introduce the random step functions

$$\mathcal{X}^{(n)}(t) := X_{\lfloor nt \rfloor}^{(n)}, \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Under the assumptions that

- (i) $m_n = 1 + \alpha n^{-1} + o(n^{-1})$ as $n \rightarrow \infty$, for some $\alpha \in \mathbb{R}$;
- (ii) $\sigma_n^2 \rightarrow \sigma^2 > 0$ as $n \rightarrow \infty$;
- (iii) $E \left(|\xi_{1,1}^{(n)} - m_n|^2 \mathbf{1}_{\{|\xi_{1,1}^{(n)} - m_n| \geq \theta \sqrt{n}\}} \right) \rightarrow 0$ as $n \rightarrow \infty$, for all $\theta > 0$; and
- (iv) $\lambda_n \rightarrow \lambda > 0$ and $b_n^2 \rightarrow b^2 > 0$ as $n \rightarrow \infty$,

Sriram [15] proved that

$$\frac{1}{n} \mathcal{X}^{(n)} \xrightarrow{D} \mathcal{X}_\alpha \quad \text{as } n \rightarrow \infty, \tag{1.5}$$

where $(\mathcal{X}_\alpha(t))_{t \in \mathbb{R}_+}$ is a (nonnegative) diffusion process with generator A_α such that

$$A_\alpha f(x) = (\lambda + \alpha x) f'(x) + \frac{1}{2} \sigma^2 x f''(x), \quad f \in C_c^\infty(\mathbb{R}_+),$$

and $\mathcal{X}_\alpha(0) = 0$. The process $(\mathcal{X}_\alpha(t))_{t \in \mathbb{R}_+}$ is the (unique) solution of the stochastic differential equation

$$d\mathcal{X}_\alpha(t) = (\lambda + \alpha \mathcal{X}_\alpha(t)) dt + \sigma \sqrt{\mathcal{X}_\alpha(t)} dW(t), \quad t \in \mathbb{R}_+,$$

with initial condition $\mathcal{X}_\alpha(0) = 0$. In Theorem 2.1, below, we show that Sriram’s result (1.5) is also valid if $\sigma_n^2 \rightarrow 0$ (in which case condition (iii) is not needed). In this case, the limit process \mathcal{X}_α is a deterministic function $\mathcal{X}_\alpha(t) = \mu_{\mathcal{X}}(t) := \lambda \int_0^t e^{\alpha s} ds, t \in \mathbb{R}_+$, satisfying the (nonrandom) differential equation $d\mu_{\mathcal{X}}(t) = (\lambda + \alpha \mu_{\mathcal{X}}(t)) dt, t \in \mathbb{R}_+$. In fact, this function can be considered to be a degenerate, i.e. deterministic, diffusion process with generator A_α such that $A_\alpha f(x) = (\lambda + \alpha x) f'(x), f \in C_c^\infty(\mathbb{R}_+)$. Note that the convergence of finite-dimensional distributions of a sequence of branching processes with immigration has been investigated by Kawazu and Watanabe [10] and Aliev [1].

Based on Sriram’s result (1.5), we can easily obtain the asymptotic behaviour of the least-squares estimators of m_n and λ_n (see Section 3). (Note that these statistics have also been investigated in the subcritical and supercritical cases; see also Section 3.) We are interested in the asymptotic behaviour of these estimators in a nearly critical case in which the offspring variance σ_n^2 tends to 0. For this purpose, the limit theorem (1.5) of Sriram does not suffice, as will be explained in Remark 3.3. Instead, we have to go one step further in the investigation of the asymptotic behaviour of the sequence $\mathcal{X}^{(n)}$. In Section 2, we prove a fluctuation limit theorem in the case in which $\sigma_n^2 \rightarrow 0$, i.e. we show that the sequence $(\mathcal{X}^{(n)} - E \mathcal{X}^{(n)})/n^{1/2}$ has a limit process $\tilde{\mathcal{X}}$ as $n \rightarrow \infty$. The process $(\tilde{\mathcal{X}}(t))_{t \in \mathbb{R}_+}$ turns out to be an Ornstein–Uhlenbeck-type process driven by a Wiener process in the sense of Definition I.4.9 of [9].

We remark that Li [13] proved a similar result for sequences of continuous-time discrete-state branching processes with immigration. He started the n th process from an appropriate nondegenerate initial distribution in such a way that its mean function became a constant an , with $a > 0$. After centring by subtracting an , and normalizing by dividing by $n^{1/2}$, he also obtained an Ornstein–Uhlenbeck-type limit process, but driven by a spectrally positive Lévy process. The basic difference comes from the fact that Li did not assume finite second moments of the offspring distributions. The method of proof is also different: Li applied Laplace transforms, while we have chosen another approach. To explain our method, let $\mathcal{F}_k^{(n)}$ denote the σ -algebra generated by $\{X_0^{(n)}, X_1^{(n)}, \dots, X_k^{(n)}\}$, for $n \in \mathbb{N}$ and $k \in \mathbb{Z}_+$. Let

$$M_k^{(n)} := X_k^{(n)} - E(X_k^{(n)} | \mathcal{F}_{k-1}^{(n)}) = X_k^{(n)} - m_n X_{k-1}^{(n)} - \lambda_n, \quad k, n \in \mathbb{N},$$

and introduce the random step functions

$$\mathcal{M}^{(n)}(t) := \sum_{k=1}^{\lfloor nt \rfloor} M_k^{(n)}, \quad t \in \mathbb{R}_+, n \in \mathbb{N}.$$

In order to prove convergence of the sequence $(\mathcal{X}^{(n)} - E \mathcal{X}^{(n)})/n^{1/2}$ as $n \rightarrow \infty$, we first show, using the martingale central limit theorem, that $\mathcal{M}^{(n)}/n^{1/2}$ has a limit process $\tilde{\mathcal{M}}$ as $n \rightarrow \infty$,

where $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$ is a Wiener process. Then we show that $(\mathcal{X}^{(n)} - E \mathcal{X}^{(n)})/n^{1/2}$ is a function of $\mathcal{M}^{(n)}/n^{1/2}$, and use a continuous-mapping-type argument to derive its convergence.

Grimvall [4] proved a fluctuation-type limit theorem for a sequence of branching processes without immigration. (See also Lamperti [12].) In this case, the processes $(X_k^{(n)})_{k \in \mathbb{Z}_+}$ cannot start from 0, and the process $X_{\lfloor nt \rfloor}^{(n)}$ can be centred by subtracting the initial value $X_0^{(n)}$, as in [13]. With suitable normalization, the limiting process will be a Wiener process, and its drift and variance depend on the limiting behaviour of the offspring mean and variance, respectively; see [4, Theorem 4.4]. In our case, the (deterministic) time change concerning the limit process $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$ is usually not linear, which is an effect of the immigration part. Grimvall not only gave sufficient conditions for the convergence of a suitable normalized sequence $X_{\lfloor nt \rfloor}^{(n)} - X_0^{(n)}$, but also proved that the Lindeberg-type condition on the offspring distribution is necessary and sufficient for the convergence. This suggests that our Lindeberg-type conditions on the offspring and immigration distributions are close to being optimal.

Based on the result of Section 2, we prove in Section 3 that the least-squares estimators of m_n and λ_n are asymptotically normal, in contrast to the case in which the offspring variance tends to a positive limit. The norming factor for the offspring mean is $n^{3/2}$, in contrast to both the subcritical case, in which it is $n^{1/2}$, and the nearly critical case with positive limiting offspring variance, in which it is n . We remark that the results of the present paper are generalizations of those of [7] and [8], where a Bernoulli offspring distribution was assumed.

2. Fluctuation limit theorem

Consider the sequence of branching processes with immigration given in (1.4). We first investigate the asymptotic behaviour of the sequence $\mathcal{X}^{(n)}/n$ in the case that $\sigma_n^2 \rightarrow 0$. We prove the following analogue of Sriram’s result (1.5).

Theorem 2.1. *Suppose that*

- (i) $m_n = 1 + \alpha n^{-1} + o(n^{-1})$ as $n \rightarrow \infty$, for some $\alpha \in \mathbb{R}$;
- (ii) $\sigma_n^2 \rightarrow 0$ as $n \rightarrow \infty$; and
- (iii) $\lambda_n \rightarrow \lambda$ and $b_n^2 \rightarrow b^2$ as $n \rightarrow \infty$, for some $\lambda \geq 0$ and $b^2 \geq 0$.

Then, weakly in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$,

$$n^{-1} \mathcal{X}^{(n)} \xrightarrow{D} \mu_{\mathcal{X}} \quad \text{as } n \rightarrow \infty,$$

where

$$\mu_{\mathcal{X}}(t) := \lambda \int_0^t e^{\alpha s} ds, \quad t \in \mathbb{R}_+.$$

Remark 2.1. Note that, in the case in which $\sigma_n^2 \rightarrow 0$, no Lindeberg condition (such as condition (iii) of Sriram [15] or condition (iii) or (v) of our Theorem 2.2, below) is needed on the random variables $\{\xi_{k,j}^{(n)} : n \in \mathbb{N}, 1 \leq k, j \leq n\}$ or $\{\varepsilon_j^{(n)} : n \in \mathbb{N}, 1 \leq j \leq n\}$.

Proof of Theorem 2.1. The theorem can be proved by an argument similar to that of [3, Chapter 9, Theorem 1.3], where it was applied to a branching process without immigration; see also [16] in the case of a single branching process with immigration, and [15] in the case of a sequence of branching processes with immigration when $\sigma_n^2 \rightarrow \sigma^2 > 0$.

Observe that $(X_k^{(n)}/n)_{k \in \mathbb{Z}_+}$ is a Markov chain with values in $E_n := \{\ell/n : \ell \in \mathbb{Z}_+\}$. For each $f \in C_c^\infty(\mathbb{R}_+)$, define

$$T_n f(x) := E f\left(n^{-1}\left(\sum_{j=1}^{nx} \xi_{1,j}^{(n)} + \varepsilon_1^{(n)}\right)\right), \quad x \in E_n.$$

Since $\mathcal{X}^{(n)}(0) = 0$, $n \in \mathbb{N}$, from the results of [3, Chapter 1, Theorem 6.5 and Chapter 4, Corollary 8.9] it is sufficient to show that

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |\Delta_n^f(x)| = 0 \quad \text{for all } f \in C_c^\infty(\mathbb{R}_+), \tag{2.1}$$

where $\Delta_n^f(x) := n(T_n f(x) - f(x)) - (\lambda + \alpha x) f'(x)$ for $x \in E_n$ and $f \in C_c^\infty(\mathbb{R}_+)$. Introducing

$$\tilde{\mathcal{J}}_k^{(n)} := \sum_{j=1}^k (\xi_{1,j}^{(n)} - 1) + \varepsilon_1^{(n)}, \quad k \in \mathbb{Z}_+, \quad n \in \mathbb{N},$$

we have $T_n f(x) = E f(x + n^{-1} \tilde{\mathcal{J}}_{nx}^{(n)})$. By Taylor’s formula,

$$T_n f(x) - f(x) = f'(x) n^{-1} E \tilde{\mathcal{J}}_{nx}^{(n)} + n^{-2} E \left((\tilde{\mathcal{J}}_{nx}^{(n)})^2 \int_0^1 (1-v) f''(x + vn^{-1} \tilde{\mathcal{J}}_{nx}^{(n)}) dv \right).$$

Since, for $x \in E_n$,

$$\begin{aligned} E \tilde{\mathcal{J}}_{nx}^{(n)} &= n(m_n - 1)x + \lambda_n, \\ E (\tilde{\mathcal{J}}_{nx}^{(n)})^2 &= n\sigma_n^2 x + b_n^2 + \lambda_n^2 + n^2(m_n - 1)^2 x^2 + 2n(m_n - 1)x\lambda_n, \end{aligned} \tag{2.2}$$

we have $\Delta_n^f(x) = \Delta_{n,1}^f(x) + \Delta_{n,2}^f(x) + \Delta_{n,3}^f(x)$, where

$$\begin{aligned} \Delta_{n,1}^f(x) &:= f'(x)((n(m_n - 1) - \alpha)x + (\lambda_n - \lambda)), \\ \Delta_{n,2}^f(x) &:= n^{-1} E \left((\tilde{\mathcal{J}}_{nx}^{(n)})^2 \int_0^1 (1-v)(f''(x + vn^{-1} \tilde{\mathcal{J}}_{nx}^{(n)}) - f''(x)) dv \right), \\ \Delta_{n,3}^f(x) &:= \frac{1}{2} x f''(x) \sigma_n^2 + \frac{1}{2n} f''(x)(b_n^2 + \lambda_n^2 + n^2(m_n - 1)^2 x^2 + 2n(m_n - 1)x\lambda_n). \end{aligned}$$

To prove (2.1), it suffices to show that $\lim_{n \rightarrow \infty} \Delta_n^f(x_n) = 0$ for every sequence $(x_n)_{n \in \mathbb{N}}$, with $x_n \in E_n$ and $n \in \mathbb{N}$, such that $x_n \rightarrow x \in [0, \infty]$. Assumptions (i)–(iii) clearly imply that $\lim_{n \rightarrow \infty} \Delta_{n,i}^f(x_n) = 0$ for $i = 1, 3$ and for all such sequences $x_n \rightarrow x \in [0, \infty]$. In order to handle $\Delta_{n,2}^f(x_n)$, suppose that the support of f is contained in $[0, c]$. Since

$$x + vn^{-1} \tilde{\mathcal{J}}_{nx}^{(n)} = x + vn^{-1} \left(\sum_{j=1}^{nx} (\xi_{1,j}^{(n)} - 1) + \varepsilon_1^{(n)} \right) \geq x(1-v),$$

the first term of the integrand in $\Delta_{n,2}^f(x)$ is 0 if $v < 1 - c/x$. Consequently,

$$|\Delta_{n,2}^f(x_n)| \leq n^{-1} (\|f''\|_\infty ((c/x_n) \wedge 1)^2 + |f''(x_n)|) E (\tilde{\mathcal{J}}_{nx_n}^{(n)})^2, \tag{2.3}$$

where $\|\cdot\|_\infty$ denotes the supremum norm. Using (2.2), we can easily check that the right-hand side of (2.3) tends to 0 for all sequences $x_n \rightarrow x \in [0, \infty]$. Thus, we conclude that $\lim_{n \rightarrow \infty} \Delta_{n,2}^f(x_n) = 0$ and, hence, obtain (2.1).

The main result of the paper is the following fluctuation limit theorem.

Theorem 2.2. *Suppose that*

- (i) $m_n = 1 + \alpha n^{-1} + o(n^{-1})$ as $n \rightarrow \infty$, for some $\alpha \in \mathbb{R}$;
- (ii) $\sigma_n^2 = \beta n^{-1} + o(n^{-1})$ as $n \rightarrow \infty$, for some $\beta \geq 0$;
- (iii) $n E \left(|\xi_{1,1}^{(n)} - m_n|^2 \mathbf{1}_{\{|\xi_{1,1}^{(n)} - m_n| > \theta n^{1/2}\}} \right) \rightarrow 0$ as $n \rightarrow \infty$, for all $\theta > 0$;
- (iv) $\lambda_n \rightarrow \lambda$ and $b_n^2 \rightarrow b^2$ as $n \rightarrow \infty$, for some $\lambda \geq 0$ and $b^2 \geq 0$; and
- (v) $E \left(|\varepsilon_1^{(n)} - \lambda_n|^2 \mathbf{1}_{\{|\varepsilon_1^{(n)} - \lambda_n| > \theta n^{1/2}\}} \right) \rightarrow 0$ as $n \rightarrow \infty$, for all $\theta > 0$.

Then, weakly in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}^2)$,

$$n^{-1/2}(\mathcal{X}^{(n)} - E \mathcal{X}^{(n)}, \mathcal{M}^{(n)}) \xrightarrow{D} (\tilde{\mathcal{X}}, \tilde{\mathcal{M}}) \quad \text{as } n \rightarrow \infty,$$

where $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$ is a Wiener process $\tilde{\mathcal{M}}(t) = W(T(t))$, $t \in \mathbb{R}_+$, with

$$T(t) := \int_0^t \varrho(s) \, ds, \quad \varrho(t) := b^2 + \beta \lambda \int_0^t e^{\alpha s} \, ds, \quad t \in \mathbb{R}_+;$$

$(W(t))_{t \in \mathbb{R}_+}$ is a standard Wiener process; and

$$\tilde{\mathcal{X}}(t) := \int_0^t e^{\alpha(t-s)} \, d\tilde{\mathcal{M}}(s), \quad t \in \mathbb{R}_+,$$

is an Ornstein–Uhlenbeck-type process driven by $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$.

Remark 2.2. Conditions (iii) and (v) are, in fact, the Lindeberg conditions for the triangular systems $\{\xi_{k,j}^{(n)}/n^{1/2} : n \in \mathbb{N}, 1 \leq k, j \leq n\}$ and $\{\varepsilon_j^{(n)}/n^{1/2} : n \in \mathbb{N}, 1 \leq j \leq n\}$, respectively; see also Grimvall [4], who investigated fluctuation theorems for sequences of branching processes without immigration, and the remarks in the introduction. Clearly, conditions (iii) and (v) hold if the offspring and immigration distributions satisfy Lyapunov-type conditions, i.e. if there exists a $\gamma > 0$ such that $n^{1-\gamma/2} E |\xi_{1,1}^{(n)} - m_n|^{2+\gamma} \rightarrow 0$ and $n^{-\gamma/2} E |\varepsilon_1^{(n)} - \lambda_n|^{2+\gamma} \rightarrow 0$ as $n \rightarrow \infty$.

Example 2.1. If $\xi_{1,1}^{(n)}$ has a Bernoulli distribution with mean $1 - an^{-1}$, where $a \geq 0$, then $m_n = 1 - an^{-1}$ and $\sigma_n^2 = an^{-1} + o(n^{-1})$ as $n \rightarrow \infty$; that is, we have $\alpha = -a$ and $\beta = a$. This special case was investigated in [7] and [8]. If the offspring distributions are geometric distributions with parameter $p_n = 1 - an^{-1}$, where $a \geq 0$, i.e. $P(\xi_{1,1}^{(n)} = \ell) = p_n(1 - p_n)^{\ell-1}$, $\ell = 1, 2, \dots$, then $m_n = p_n^{-1} = 1 + an^{-1} + o(n^{-1})$ and $\sigma_n^2 = (1 - p_n)p_n^{-2} = an^{-1} + o(n^{-1})$ as $n \rightarrow \infty$. In this case, $\alpha = a$ and $\beta = a$. Finally, suppose that the offspring distributions have a common range $\{0, 1, 2\}$, such that $P(\xi_{1,1}^{(n)} = 0) = an^{-1}$, $P(\xi_{1,1}^{(n)} = 1) = 1 - (a + b)n^{-1}$, and $P(\xi_{1,1}^{(n)} = 2) = bn^{-1}$, where $a, b \geq 0$. Then $m_n = 1 + (b - a)n^{-1}$ and $\sigma_n^2 = (a + b)n^{-1} + o(n^{-1})$ as $n \rightarrow \infty$, i.e. $\alpha = b - a$ and $\beta = a + b$.

We note that condition (iii) is trivially satisfied in the first and third cases. In the second case, the centred fourth moment of the geometric distribution with the above parametrization tends to 0 as $n \rightarrow \infty$ and, thus, by Remark 2.2, condition (iii) holds.

Remark 2.3. We remark that $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$ is a continuous zero-mean Gaussian process with independent (but not necessarily stationary) increments. It has stationary increments if and only if $\beta\lambda = 0$, when $\tilde{\mathcal{M}}(t) = W(b^2t)$, $t \in \mathbb{R}_+$, is a Wiener process with variance b^2 . The process $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$ is always a martingale, meaning that we can define stochastic integrals with respect to it. Its covariance function has the form

$$\text{cov}(\tilde{\mathcal{M}}(s), \tilde{\mathcal{M}}(t)) = T(s \wedge t) \quad \text{for } s, t \in \mathbb{R}_+.$$

By comparing the covariance structures, we obtain another representation of the process, of the form

$$\tilde{\mathcal{M}}(t) = \int_0^t \sqrt{\varrho(s)} \, dW(s), \quad t \in \mathbb{R}_+.$$

Consequently, the process $(\tilde{\mathcal{M}}(t))_{t \in \mathbb{R}_+}$ is the unique solution of the stochastic differential equation

$$d\tilde{\mathcal{M}}(t) = \sqrt{\varrho(t)} \, dW(t), \quad t \in \mathbb{R}_+,$$

with initial condition $\tilde{\mathcal{M}}(0) = 0$. The process $(\tilde{\mathcal{X}}(t))_{t \in \mathbb{R}_+}$ is a continuous zero-mean Gaussian martingale with covariance function

$$\text{cov}(\tilde{\mathcal{X}}(s), \tilde{\mathcal{X}}(t)) = \int_0^{s \wedge t} e^{\alpha(s+t-2u)} \varrho(u) \, du \quad \text{for } s, t \in \mathbb{R}_+.$$

We remark that the process $(\tilde{\mathcal{X}}(t))_{t \in \mathbb{R}_+}$ has independent increments if and only if $\alpha = 0$, when $\tilde{\mathcal{X}} = \tilde{\mathcal{M}}$. By comparing the covariance structures, we have the representation

$$\tilde{\mathcal{X}}(t) = e^{\alpha t} \int_0^t e^{-\alpha s} \sqrt{\varrho(s)} \, dW(s) \quad \text{for } t \in \mathbb{R}_+.$$

Hence, for the process $\tilde{\mathcal{Y}}(t) := e^{-\alpha t} \tilde{\mathcal{X}}(t)$, $t \in \mathbb{R}_+$, we have $d\tilde{\mathcal{Y}}(t) = e^{-\alpha t} (\varrho(t))^{1/2} \, dW(t)$. From Itô’s formula, we find that the process $\tilde{\mathcal{X}}(t) = e^{\alpha t} \tilde{\mathcal{Y}}(t)$ is the unique solution of the stochastic differential equation

$$d\tilde{\mathcal{X}}(t) = \alpha \tilde{\mathcal{X}}(t) \, dt + \sqrt{\varrho(t)} \, dW(t), \quad t \in \mathbb{R}_+, \quad \tilde{\mathcal{X}}(0) = 0.$$

In order to prove Theorem 2.2, we need the following lemma.

Lemma 2.1. *Let $(X_k)_{k \in \mathbb{Z}_+}$ be the branching processes with immigration given in (1.1). Then, for all $k \in \mathbb{Z}_+$ and $m \neq 1$, we have*

$$E X_k = \frac{m^k - 1}{m - 1} \lambda, \quad \text{var } X_k = \frac{m^{2k} - 1}{m^2 - 1} b^2 + \frac{(m^k - 1)(m^{k-1} - 1)}{(m - 1)(m^2 - 1)} \lambda \sigma^2,$$

and, if $m = 1$, we have $E X_k = k\lambda$ and $\text{var } X_k = kb^2 + \frac{1}{2}k(k - 1)\lambda\sigma^2$. Moreover,

$$\text{cov}(X_k, X_\ell) = m^{|k-\ell|} \text{var } X_{k \wedge \ell}, \quad E(M_k^2 \mid \mathcal{F}_{k-1}) = \sigma^2 X_{k-1} + b^2$$

for all $k, \ell \in \mathbb{Z}_+$.

Proof. The formulae for $E X_k$, $\text{var } X_k$, and $\text{cov}(X_k, X_\ell)$ are direct consequences of the recursive relationships

$$\begin{aligned} E X_k &= m E X_{k-1} + \lambda && \text{for } k \in \mathbb{N}, \\ \text{var } X_k &= m^2 \text{var } X_{k-1} + \sigma^2 E X_{k-1} + b^2 && \text{for } k \in \mathbb{N}, \\ \text{cov}(X_k, X_\ell) &= m \text{cov}(X_k, X_{\ell-1}) && \text{for } 0 \leq k < \ell, \end{aligned} \tag{2.4}$$

which follow from calculating the conditional expectation with respect to X_{k-1} in the first two cases, with respect to the σ -algebra $\mathcal{F}_{\ell-1}$ in the third case, and then taking the expectation. The final assertion of the theorem is obvious.

Proof of Theorem 2.2. We will make the following steps.

- (A) We prove that $\tilde{\mathcal{M}}^{(n)} := \mathcal{M}^{(n)} / n^{1/2} \xrightarrow{D} \tilde{\mathcal{M}}$, using the martingale central limit theorem.
- (B) We show that $\tilde{\mathcal{X}}^{(n)} := (\mathcal{X}^{(n)} - E \mathcal{X}^{(n)}) / n^{1/2} = \Phi_n(\tilde{\mathcal{M}}^{(n)})$, where $\Phi_n, \Phi: D(\mathbb{R}_+, \mathbb{R}) \rightarrow D(\mathbb{R}_+, \mathbb{R})$ are measurable mappings such that $\Phi_n \rightarrow \Phi$ in an appropriate sense, implying that $(\tilde{\mathcal{X}}^{(n)}, \tilde{\mathcal{M}}^{(n)}) \xrightarrow{D} (\Phi(\tilde{\mathcal{M}}), \tilde{\mathcal{M}})$.
- (C) We prove that $\Phi(\tilde{\mathcal{M}}) = \tilde{\mathcal{X}}$.

Step (A). By the martingale central limit theorem (see, e.g. [9, Theorem VIII.3.33]), it suffices to prove that, for all $t \in \mathbb{R}_+$,

$$\frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} E((M_k^{(n)})^2 \mid \mathcal{F}_{k-1}^{(n)}) \xrightarrow{P} T(t) \quad \text{as } n \rightarrow \infty, \tag{2.5}$$

where ‘ \xrightarrow{P} ’ denotes convergence in probability, and that, for all $\theta > 0$ and $t \in \mathbb{R}_+$,

$$\frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} E \left((M_k^{(n)})^2 \mathbf{1}_{\{|M_k^{(n)}| > \theta \sqrt{n}\}} \mid \mathcal{F}_{k-1}^{(n)} \right) \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty. \tag{2.6}$$

By Lemma 2.1, $E((M_k^{(n)})^2 \mid \mathcal{F}_{k-1}^{(n)}) = \sigma_n^2 X_{k-1}^{(n)} + b_n^2$. Thus, in order to prove (2.5), we have to show that

$$\frac{\sigma_n^2}{n} \sum_{k=1}^{\lfloor nt \rfloor} X_{k-1}^{(n)} \xrightarrow{P} \beta \lambda \int_0^t \left(\int_0^v e^{\alpha u} du \right) dv \quad \text{as } n \rightarrow \infty. \tag{2.7}$$

This statement will clearly hold once we have proved that

$$E \left(\frac{\sigma_n^2}{n} \sum_{k=1}^{\lfloor nt \rfloor} X_{k-1}^{(n)} \right) \rightarrow \beta \lambda \int_0^t \left(\int_0^v e^{\alpha u} du \right) dv \quad \text{as } n \rightarrow \infty, \tag{2.8}$$

$$\text{var} \left(\frac{\sigma_n^2}{n} \sum_{k=1}^{\lfloor nt \rfloor} X_{k-1}^{(n)} \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{2.9}$$

By Lemma 2.1, along a subsequence with $m_n = 1$, we have

$$\frac{\sigma_n^2}{n} \sum_{k=1}^{\lfloor nt \rfloor} E X_{k-1}^{(n)} = n \sigma_n^2 \lambda_n \frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{2n^2} \rightarrow \frac{1}{2} \beta \lambda t^2 \quad \text{as } n \rightarrow \infty. \tag{2.10}$$

Assumption (i) of Theorem 2.2 implies that

$$n(m_n - 1) \rightarrow \alpha \quad \text{as } n \rightarrow \infty. \tag{2.11}$$

For sufficiently large $n \in \mathbb{N}$, we have $m_n = e^{\alpha_n/n}$, where

$$\alpha_n = n \log m_n = \log \left(1 + \frac{\alpha + o(1)}{n} \right)^n \rightarrow \alpha \quad \text{as } n \rightarrow \infty.$$

Thus,

$$m_n^{\lfloor nt \rfloor} = e^{\alpha_n \lfloor nt \rfloor / n} \rightarrow e^{\alpha t} \quad \text{as } n \rightarrow \infty. \tag{2.12}$$

Consequently, along a subsequence with $m_n \neq 1$, we find that

$$\frac{\sigma_n^2}{n} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} X_{k-1}^{(n)} = \frac{n\sigma_n^2 \lambda_n}{n(m_n - 1)} \left(\frac{m_n^{\lfloor nt \rfloor} - 1}{n(m_n - 1)} - \frac{\lfloor nt \rfloor}{n} \right) \rightarrow \frac{\beta \lambda}{\alpha} \left(\frac{e^{\alpha t} - 1}{\alpha} - t \right)$$

if $\alpha \neq 0$; otherwise, the limit is $\frac{1}{2}\beta\lambda t^2$. Taking this convergence and (2.10) into account, we recover (2.8).

In order to prove (2.9), we first note that, by Lemma 2.1,

$$\text{var} \left(\frac{\sigma_n^2}{n} \sum_{k=1}^{\lfloor nt \rfloor} X_{k-1}^{(n)} \right) = (U_n(t)b_n^2 + V_n(t)\lambda_n\sigma_n^2) \frac{\sigma_n^4}{n^2}, \tag{2.13}$$

where

$$U_n(t) := \begin{cases} \sum_{k=1}^{\lfloor nt \rfloor} \frac{m_n^{2k-2} - 1}{m_n^2 - 1} \left(2 \frac{m_n^{\lfloor nt \rfloor - k + 1} - 1}{m_n - 1} - 1 \right) & \text{if } m_n \neq 1, \\ \sum_{k=1}^{\lfloor nt \rfloor} (k-1)(2(\lfloor nt \rfloor - k + 1) - 1) & \text{if } m_n = 1, \end{cases}$$

$$V_n(t) := \begin{cases} \sum_{k=1}^{\lfloor nt \rfloor} \frac{(m_n^{k-1} - 1)(m_n^{k-2} - 1)}{(m_n - 1)(m_n^2 - 1)} \left(2 \frac{m_n^{\lfloor nt \rfloor - k + 1} - 1}{m_n - 1} - 1 \right) & \text{if } m_n \neq 1, \\ \sum_{k=1}^{\lfloor nt \rfloor} \frac{1}{2} (k-1)(k-2)(2(\lfloor nt \rfloor - k + 1) - 1) & \text{if } m_n = 1. \end{cases}$$

If $m_n = 1$ then $\lim_{n \rightarrow \infty} U_n(t)/n^3 = \frac{1}{3}t^3$ for all $t \in \mathbb{R}_+$ and $\lim_{n \rightarrow \infty} V_n(t)/n^4 = \frac{1}{12}t^4$ for all $t \in \mathbb{R}_+$. Thus, taking into account assumptions (ii) and (iv), along a subsequence with $m_n = 1$, we have

$$(U_n(t)b_n^2 + V_n(t)\lambda_n\sigma_n^2) \frac{\sigma_n^4}{n^2} = \frac{n^2\sigma_n^4 b_n^2}{n} \frac{U_n(t)}{n^3} + \frac{n^3\sigma_n^6 \lambda_n}{n} \frac{V_n(t)}{n^4} \rightarrow 0 \quad \text{as } n \rightarrow \infty, \tag{2.14}$$

for all $t \in \mathbb{R}_+$. If $m_n \neq 1$ then, using (2.11) and (2.12) in the case of $\alpha \neq 0$, we have

$$\lim_{n \rightarrow \infty} \frac{U_n(t)}{n^3} = \frac{e^{2\alpha t} - 4e^{\alpha t} + 2\alpha t + 3}{2\alpha^3},$$

$$\lim_{n \rightarrow \infty} \frac{V_n(t)}{n^4} = \frac{e^{2\alpha t} - 4(\alpha t - 1)e^{\alpha t} - 2\alpha t - 5}{2\alpha^4};$$

otherwise, the limits are $\frac{1}{3}t^3$ and $\frac{1}{12}t^4$, respectively. Hence, for all $t \in \mathbb{R}_+$, along a subsequence with $m_n \neq 1$ we again obtain (2.14). From (2.13), we recover (2.9) and, finally, (2.5).

To prove the conditional Lindeberg condition (2.6), we consider the decomposition $M_k^{(n)} = N_{n,k}^{(1)} + N_{n,k}^{(2)}$, where

$$N_{n,k}^{(1)} := \sum_{j=1}^{X_{k-1}^{(n)}} (\xi_{k,j}^{(n)} - m_n), \quad N_{n,k}^{(2)} := \varepsilon_k^{(n)} - \lambda_n.$$

Note that, for any pair of random variables, Y and Z , and for any $\theta > 0$, we have

$$\mathbf{1}_{\{|Y+Z|>\theta\}} \leq \mathbf{1}_{\{|Y|>\theta/2\}} + \mathbf{1}_{\{|Z|>\theta/2\}}. \tag{2.15}$$

Hence, it suffices to show that

$$L_{i,j}^{(n)}(t) := \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} \left((N_{n,k}^{(i)})^2 \mathbf{1}_{\{|N_{n,k}^{(j)}|>\theta\sqrt{n}\}} \mid \mathcal{F}_{k-1}^{(n)} \right) \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty, \tag{2.16}$$

for $i, j = 1, 2$, all $\theta > 0$, and all $t \in \mathbb{R}_+$. To prove (2.16) for $i = j = 1$, we introduce the random step functions

$$\mathcal{J}_n(t) := \sum_{j=1}^{\lfloor nt \rfloor} (\xi_{1,j}^{(n)} - m_n), \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

We note that $\mathbb{E} \mathcal{J}_n(t) = 0$, while condition (ii) implies that $\text{var} \mathcal{J}_n(t) = \lfloor nt \rfloor \sigma_n^2 \rightarrow \beta t$. Denote by $D([0, t], \mathbb{R}_+)$ the Skorokhod space of nonnegative, càdlàg functions on $[0, t]$, and let $C([0, t], \mathbb{R}_+)$ be the closed convex cone of nonnegative, continuous functions on $[0, t]$. For fixed $\theta > 0$ and $t \in \mathbb{R}_+$, we have

$$L_{1,1}^{(n)}(t) = \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} \left(\mathcal{J}_n \left(\frac{\ell}{n} \right)^2 \mathbf{1}_{\{|\mathcal{J}_n(\ell/n)|>\theta\sqrt{n}\}} \right) \Bigg|_{\ell=X_{k-1}^{(n)}} = F_n \left(\frac{1}{n} \mathcal{X}^{(n)} \right),$$

where the measurable mapping $F_n : D([0, t], \mathbb{R}_+) \rightarrow \mathbb{R}$ is defined by

$$F_n(x) := \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} \left(\mathcal{J}_n \left(x \left(\frac{k-1}{n} \right) \right)^2 \mathbf{1}_{\{|\mathcal{J}_n(x((k-1)/n))|>\theta\sqrt{n}\}} \right) \quad \text{for } x \in D([0, t], \mathbb{R}_+).$$

By Theorem 2.1, we have $n^{-1} \mathcal{X}^{(n)} \xrightarrow{D} \mu_{\mathcal{X}}$ as $n \rightarrow \infty$, where $\mu_{\mathcal{X}}$ is a continuous, nonnegative, monotone-increasing function. In view of the continuous-mapping theorem (see [2, Theorem 5.5]), in order to prove (2.16) it suffices to show that $F_n(x_n) \rightarrow 0$ if $x \in C([0, t], \mathbb{R}_+)$ and $x_n \in D([0, t], \mathbb{R}_+)$, with $\|x_n - x\|_{\infty} \rightarrow 0$ as $n \rightarrow \infty$. (We use the fact that, in this case, $x_n \rightarrow x$ in $D([0, t], \mathbb{R}_+)$ if and only if $\|x_n - x\|_{\infty} \rightarrow 0$; see [9, Proposition VI.1.17].) Note that, in fact, we may restrict ourselves to just the function $x = \mu_{\mathcal{X}} \mathbf{1}_{[0,t]}$ since the limit process concentrates on the point $\mu_{\mathcal{X}} \in C(\mathbb{R}_+, \mathbb{R}_+)$, but the proof is the same for a general function x . We can see that $F_n(x_n) = A_n + B_n$, $n \in \mathbb{N}$, where

$$A_n := \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nx_n((k-1)/n) \rfloor} \mathbb{E} \left(|\xi_{1,j}^{(n)} - m_n|^2 \mathbf{1}_{\{|\mathcal{J}_n(x_n((k-1)/n))|>\theta\sqrt{n}\}} \right),$$

$$B_n := \frac{2}{n} \sum_{k=1}^{\lfloor nt \rfloor} \sum_{i=2}^{\lfloor nx_n((k-1)/n) \rfloor} \sum_{j=1}^{i-1} \mathbb{E} \left((\xi_{1,i}^{(n)} - m_n)(\xi_{1,j}^{(n)} - m_n) \mathbf{1}_{\{|\mathcal{J}_n(x_n((k-1)/n))|>\theta\sqrt{n}\}} \right).$$

Define the random step functions $\mathfrak{g}_{n,j}(u) := \mathfrak{g}_n(u) - (\xi_{1,j}^{(n)} - m_n) \mathbf{1}_{\{j \leq nu\}}$ for $u \in \mathbb{R}_+$, $j, n \in \mathbb{N}$. Using (2.15) we obtain, for any $u \in \mathbb{R}_+$,

$$E\left(|\xi_{1,j}^{(n)} - m_n|^2 \mathbf{1}_{\{|\mathfrak{g}_n(u)| > \theta\sqrt{n}\}}\right) \leq E\left(|\xi_{1,j}^{(n)} - m_n|^2 \left(\mathbf{1}_{\{|\xi_{1,j}^{(n)} - m_n| > \theta\sqrt{n}/2\}} + \mathbf{1}_{\{|\mathfrak{g}_{n,j}(u)| > \theta\sqrt{n}/2\}}\right)\right).$$

The random variable $\xi_{1,j}^{(n)}$ and the random step function $(\mathfrak{g}_{n,j}(u))_{u \in \mathbb{R}_+}$ are independent for all $j, n \in \mathbb{N}$ so, using the Markov inequality, we have

$$E\left(|\xi_{1,j}^{(n)} - m_n|^2 \mathbf{1}_{\{|\mathfrak{g}_{n,j}(u)| > \theta\sqrt{n}/2\}}\right) \leq 4\theta^{-2}n^{-1}\sigma_n^4(\lfloor nu \rfloor - 1).$$

Since x_n tends to a continuous function in the uniform topology, there exists $N \in \mathbb{N}$ such that $x_n((k-1)/n) \leq \|x\|_\infty + 1$ for all $k \in \mathbb{N}$, if $n > N$. Thus,

$$A_n \leq (\|x\|_\infty + 1)tn E\left(|\xi_{1,1}^{(n)} - m_n|^2 \mathbf{1}_{\{|\xi_{1,1}^{(n)} - m_n| > \theta\sqrt{n}\}}\right) + 4\theta^{-2}(\|x\|_\infty + 1)^2 t n \sigma_n^4$$

if $n > N$. The first term on the right-hand side tends to 0 by assumption (iii), and the second term is $O(n^{-1})$ by assumption (ii). Thus, we find that $A_n \rightarrow 0$ as $n \rightarrow \infty$. To estimate B_n , introduce the random step functions

$$\mathcal{V}_n(u) = 2 \sum_{i=2}^{\lfloor nu \rfloor} \sum_{j=1}^{i-1} (\xi_{1,i}^{(n)} - m_n)(\xi_{1,j}^{(n)} - m_n), \quad u \in \mathbb{R}_+, n \in \mathbb{N}.$$

Using the Cauchy–Schwarz and Markov inequalities, we obtain

$$E\left(|\mathcal{V}_n(u)| \mathbf{1}_{\{|\mathfrak{g}_n(u)| > \theta\sqrt{n}\}}\right) \leq \theta^{-1}n^{-1/2}(E \mathcal{V}_n^2(u) E \mathfrak{g}_n^2(u))^{1/2}$$

for all $u \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Since the random variables $\xi_{1,i}^{(n)}$ and $\xi_{1,j}^{(n)}$, $n \in \mathbb{N}$, are independent if $i \neq j$, we see that

$$E \mathcal{V}_n^2(u) = 2\lfloor nu \rfloor (\lfloor nu \rfloor - 1)\sigma_n^4 \quad \text{for all } u \in \mathbb{R}_+ \text{ and } n \in \mathbb{N}.$$

Thus, for $n > N$, we have

$$|B_n| \leq Cn^{-3/2}\sigma_n^3 \sum_{k=1}^{\lfloor nt \rfloor} \left[nx_n\left(\frac{k-1}{n}\right) \right] \left(\left[nx_n\left(\frac{k-1}{n}\right) \right] - 1 \right)^{1/2} \leq Ct(\|x\|_\infty + 1)^{3/2}n\sigma_n^3,$$

where $C := 2^{1/2}\theta^{-1}$. Clearly, the right-hand side tends to 0 since $n\sigma_n^3 \rightarrow 0$ as $n \rightarrow \infty$, by assumption (ii). Thus, we find that $B_n \rightarrow 0$ as $n \rightarrow \infty$. Consequently, we conclude that $F_n(x_n) \rightarrow 0$ and, finally, we recover (2.16).

In the case that $i = 1$ and $j = 2$, we note that

$$E\left((N_{n,k}^{(1)})^2 \mathbf{1}_{\{|N_{n,k}^{(2)}| > \theta\sqrt{n}\}} \mid \mathcal{F}_{k-1}^{(n)}\right) = \sigma_n^2 X_{k-1}^{(n)} P(|\varepsilon_1^{(n)} - \lambda_n| > \theta\sqrt{n})$$

and, moreover,

$$P(|\varepsilon_1^{(n)} - \lambda_n| > \theta\sqrt{n}) \leq \theta^{-2}n^{-1} E\left((\varepsilon_1^{(n)} - \lambda_n)^2 \mathbf{1}_{\{|\varepsilon_1^{(n)} - \lambda_n| > \theta\sqrt{n}\}}\right);$$

hence, (2.16) is a consequence of (2.7) and assumptions (ii) and (v). In order to establish (2.16) for $i = 2$ and $j = 1$, we use the estimate

$$\begin{aligned} E \left((N_{n,k}^{(2)})^2 \mathbf{1}_{\{|N_{n,k}^{(1)}| > \theta \sqrt{n}\}} \mid \mathcal{F}_{k-1}^{(n)} \right) &\leq \theta^{-2} n^{-1} E \left((N_{n,k}^{(2)})^2 (N_{n,k}^{(1)})^2 \mid \mathcal{F}_{k-1}^{(n)} \right) \\ &= \theta^{-2} n^{-1} E \left((\varepsilon_k^{(n)} - \lambda_n)^2 \left(\sum_{j=1}^{X_{k-1}^{(n)}} (\xi_{k,j} - m_n) \right)^2 \mid \mathcal{F}_{k-1}^{(n)} \right) \\ &= \theta^{-2} n^{-1} b_n^2 \sigma_n^2 X_{k-1}^{(n)}. \end{aligned}$$

Thus, (2.16) follows from (2.7) and assumptions (ii) and (iv). Finally, we have

$$E \left((N_{n,k}^{(2)})^2 \mathbf{1}_{\{|N_{n,k}^{(2)}| > \theta \sqrt{n}\}} \mid \mathcal{F}_{k-1}^{(n)} \right) = E \left((\varepsilon_1^{(n)} - \lambda_n)^2 \mathbf{1}_{\{|\varepsilon_1^{(n)} - \lambda_n| > \theta \sqrt{n}\}} \right)$$

and, hence, the case $i = j = 2$ follows from assumption (v). Having thus proved (2.16), we conclude step (A).

Step (B). By the difference equation (1.2) and the recursion (2.4), we obtain the recursion

$$X_k^{(n)} - E X_k^{(n)} = m_n (X_{k-1}^{(n)} - E X_{k-1}^{(n)}) + M_k^{(n)},$$

with solution $X_k^{(n)} - E X_k^{(n)} = \sum_{j=1}^k m_n^{k-j} M_j^{(n)}$. Hence,

$$\tilde{X}^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nt \rfloor} m_n^{\lfloor nt \rfloor - j} M_j^{(n)} = \sum_{j=1}^{\lfloor nt \rfloor} m_n^{\lfloor nt \rfloor - j} \left(\tilde{\mathcal{M}}^{(n)} \left(\frac{j}{n} \right) - \tilde{\mathcal{M}}^{(n)} \left(\frac{j-1}{n} \right) \right).$$

Again writing $m_n = e^{\alpha_n/n}$, where $\alpha_n \rightarrow \alpha$, we have

$$\tilde{X}^{(n)}(t) = \int_0^{\lfloor nt \rfloor/n} e^{\alpha_n(\lfloor nt \rfloor/n - s)} d\tilde{\mathcal{M}}^{(n)}(s),$$

which suggests that $\tilde{X}^{(n)} \xrightarrow{D} \int_0^t e^{\alpha(t-s)} d\tilde{\mathcal{M}}(s)$. We prove this weak convergence in the following way: we rearrange the sum, then use the continuous-mapping theorem, and, finally, rearrange the result using Itô’s formula. Thus, we write

$$\tilde{X}^{(n)}(t) = \tilde{\mathcal{M}}^{(n)} \left(\frac{\lfloor nt \rfloor}{n} \right) + \alpha_n \int_0^{\lfloor nt \rfloor/n} e^{\alpha_n(\lfloor nt \rfloor/n - s)} \tilde{\mathcal{M}}^{(n)}(s) ds.$$

This implies that $(\tilde{X}^{(n)}, \tilde{\mathcal{M}}^{(n)}) = \Psi_n(\tilde{\mathcal{M}}^{(n)})$, with the mapping $\Psi_n : D(\mathbb{R}_+, \mathbb{R}) \rightarrow D(\mathbb{R}_+, \mathbb{R}^2)$ defined by

$$\Psi_n(x)(t) = \left(x \left(\frac{\lfloor nt \rfloor}{n} \right) + \alpha_n \int_0^{\lfloor nt \rfloor/n} e^{\alpha_n(\lfloor nt \rfloor/n - s)} x(s) ds, x(t) \right) \quad \text{for } x \in D(\mathbb{R}_+, \mathbb{R}).$$

We want to show that $\Psi_n(\tilde{\mathcal{M}}^{(n)}) \xrightarrow{D} \Psi(\tilde{\mathcal{M}})$, where the mapping $\Psi : D(\mathbb{R}_+, \mathbb{R}) \rightarrow D(\mathbb{R}_+, \mathbb{R}^2)$ is defined by

$$\Psi(x)(t) = \left(x(t) + \alpha \int_0^t e^{\alpha(t-s)} x(s) ds, x(t) \right) \quad \text{for } x \in D(\mathbb{R}_+, \mathbb{R}).$$

Since almost all trajectories of the limit process are continuous, in view of the continuous-mapping theorem, it suffices to check that $\Psi_n(x_n) \rightarrow \Psi(x)$ if $x \in C([0, t], \mathbb{R})$ and $x_n \in D([0, t], \mathbb{R})$, with $x_n \rightarrow x$ as $n \rightarrow \infty$. (Note that the mappings Φ_n and Φ can be defined as the first coordinates of the mappings Ψ_n and Ψ , respectively.)

Step (C). Itô's formula yields

$$\tilde{X}(t) = \int_0^t e^{\alpha(t-s)} d\tilde{M}(s) = \tilde{M}(t) + \alpha \int_0^t e^{\alpha(t-s)} \tilde{M}(s) ds,$$

whence $\Psi(\tilde{M}) = (\tilde{X}, \tilde{M})$. This completes the proof of Theorem 2.2.

3. Asymptotics of the least-squares estimators

Consider the branching process with immigration given in (1.1). If the immigration mean λ is known, then the conditional least-squares estimator \hat{m}_n , based on the regression equation (1.2), can be obtained by minimizing the sum of squares

$$\sum_{k=1}^n (X_k - mX_{k-1} - \lambda)^2 \tag{3.1}$$

with respect to m . It takes the form

$$\hat{m}_n = \frac{\sum_{k=1}^n X_{k-1}(X_k - \lambda)}{\sum_{k=1}^n X_{k-1}^2}.$$

If the immigration mean λ is unknown, then the joint conditional least-squares estimator $(\tilde{m}_n, \tilde{\lambda}_n)$ of the vector (m, λ) can be obtained by minimizing the sum of squares (3.1) with respect to m and λ . The components of this estimator take the form

$$\tilde{m}_n = \frac{\sum_{k=1}^n X_{k-1}(X_k - \bar{X})}{\sum_{k=1}^n (X_{k-1} - \bar{X}_*)^2}, \quad \tilde{\lambda}_n = \bar{X} - \tilde{m}_n \bar{X}_*,$$

where

$$\bar{X} := \frac{1}{n} \sum_{k=1}^n X_k, \quad \bar{X}_* := \frac{1}{n} \sum_{k=1}^n X_{k-1}.$$

In the subcritical case ($m < 1$), under the assumptions that $E \xi_{1,1}^3 < \infty$ and $E \varepsilon_1^3 < \infty$, the estimators \hat{m}_n and $(\tilde{m}_n, \tilde{\lambda}_n)$ are asymptotically normal:

$$\begin{aligned} n^{1/2}(\hat{m}_n - m) &\xrightarrow{D} \mathcal{N}(0, c^2) \quad \text{as } n \rightarrow \infty, \\ \begin{pmatrix} n^{1/2}(\tilde{m}_n - m) \\ n^{1/2}(\tilde{\lambda}_n - \lambda) \end{pmatrix} &\xrightarrow{D} \mathcal{N}(0, \Sigma) \quad \text{as } n \rightarrow \infty, \end{aligned}$$

where the variance c^2 and the covariance matrix Σ can be respectively expressed in terms of the moments, up to the third order, of the offspring and immigration distributions (see Klimko and Nelson [11]); closely related estimators were proposed and studied by Heyde and Seneta [5], [6] and Quine [14].

In the critical case ($m = 1$), the estimators \widehat{m}_n and $(\widetilde{m}_n, \widetilde{\lambda}_n)$ are not asymptotically normal, but

$$n(\widehat{m}_n - 1) \xrightarrow{D} \frac{\frac{1}{2}\mathcal{X}(1)^2 - (\lambda + \frac{1}{2}\sigma^2) \int_0^1 \mathcal{X}(t) dt}{\int_0^1 \mathcal{X}(t)^2 dt} \quad \text{as } n \rightarrow \infty, \tag{3.2}$$

and

$$n(\widetilde{m}_n - 1) \xrightarrow{D} \frac{\frac{1}{2}\mathcal{X}(1)^2 - (\mathcal{X}(1) + \frac{1}{2}\sigma^2) \int_0^1 \mathcal{X}(t) dt}{\int_0^1 \mathcal{X}(t)^2 dt - (\int_0^1 \mathcal{X}(t) dt)^2} \quad \text{as } n \rightarrow \infty; \tag{3.3}$$

see [17]. The proof is based on the convergence result (1.3). Wei and Winnicki [17] also proved that $\widetilde{\lambda}_n$ is not a consistent estimator of λ .

Now let us consider the sequence of branching processes with immigration given in (1.4). In the case in which λ_n is known, the estimator \widehat{m}_n for m_n is given by

$$\widehat{m}_n = \frac{\sum_{k=1}^n X_{k-1}^{(n)} (X_k^{(n)} - \lambda_n)}{\sum_{k=1}^n (X_{k-1}^{(n)})^2}.$$

If both m_n and λ_n are unknown, then the components of the joint estimator $(\widetilde{m}_n, \widetilde{\lambda}_n)$ are given by

$$\widetilde{m}_n = \frac{\sum_{k=1}^n X_{k-1}^{(n)} (X_k^{(n)} - \overline{X}^{(n)})}{\sum_{k=1}^n (X_{k-1}^{(n)} - \overline{X}_*^{(n)})^2}, \quad \widetilde{\lambda}_n = \overline{X}^{(n)} - \widetilde{m}_n \overline{X}_*^{(n)},$$

where, now,

$$\overline{X}^{(n)} := \frac{1}{n} \sum_{k=1}^n X_k^{(n)}, \quad \overline{X}_*^{(n)} := \frac{1}{n} \sum_{k=1}^n X_{k-1}^{(n)}.$$

Using the convergence result (1.5), due to Sriram [15], we can easily show that (3.2) and (3.3) hold with \mathcal{X} replaced by \mathcal{X}_α in the case that $\sigma_n^2 \rightarrow \sigma^2 > 0$.

By applying the continuous-mapping theorem and using Slutsky’s argument, in the nearly critical model of Theorem 2.2 we can derive the asymptotic behaviour of the estimators \widehat{m}_n and $(\widetilde{m}_n, \widetilde{\lambda}_n)$ in exactly the same way as was done in the case of a Bernoulli offspring distribution in [7] and [8].

Theorem 3.1. *Suppose that the assumptions of Theorem 2.2 hold, for some $\lambda > 0$. Then*

$$n^{3/2}(\widehat{m}_n - m_n) \xrightarrow{D} \frac{\int_0^1 \mu_{\mathcal{X}}(t) d\widetilde{\mathcal{M}}(t)}{\int_0^1 \mu_{\mathcal{X}}(t)^2 dt} \stackrel{D}{=} \mathcal{N}(0, c^2),$$

where ‘ $\stackrel{D}{=}$ ’ denotes equality in distribution, $\mu_{\mathcal{X}}(t) = \lambda \int_0^t e^{au} du$, $t \in \mathbb{R}_+$, and

$$c^2 := \frac{\int_0^1 \mu_{\mathcal{X}}(t)^2 \varrho(t) dt}{(\int_0^1 \mu_{\mathcal{X}}(t)^2 dt)^2},$$

with $\varrho(t) = b^2 + \beta \mu_{\mathcal{X}}(t)$, $t \in \mathbb{R}_+$.

Moreover,

$$\begin{pmatrix} n^{3/2}(\widetilde{m}_n - m_n) \\ n^{1/2}(\widetilde{\lambda}_n - \lambda_n) \end{pmatrix} \xrightarrow{D} \begin{pmatrix} \frac{\int_0^1 \mu_{\mathcal{X}}(t) d\widetilde{\mathcal{M}}(t) - \overline{\mu}_{\mathcal{X}} \widetilde{\mathcal{M}}(1)}{\int_0^1 (\mu_{\mathcal{X}}(t) - \overline{\mu}_{\mathcal{X}})^2 dt} \\ \frac{\widetilde{\mathcal{M}}(1) \int_0^1 \mu_{\mathcal{X}}(t)^2 dt - \overline{\mu}_{\mathcal{X}} \int_0^1 \mu_{\mathcal{X}}(t) d\widetilde{\mathcal{M}}(t)}{\int_0^1 (\mu_{\mathcal{X}}(t) - \overline{\mu}_{\mathcal{X}})^2 dt} \end{pmatrix} \stackrel{D}{=} \mathcal{N}(0, \Sigma),$$

where $\bar{\mu}_X := \int_0^1 \mu_X(t) dt$ and

$$\Sigma := \frac{1}{\left(\int_0^1 (\mu_X(t) - \bar{\mu}_X)^2 dt\right)^2} (\sigma_{i,j})_{1 \leq i, j \leq 2},$$

with

$$\begin{aligned} \sigma_{1,1} &:= \int_0^1 (\mu_X(t) - \bar{\mu}_X)^2 \varrho(t) dt, \\ \sigma_{1,2} = \sigma_{2,1} &:= \int_0^1 (\mu_X(t) - \bar{\mu}_X) \left(\int_0^1 \mu_X(t)^2 dt - \bar{\mu}_X \mu_X(t) \right) \varrho(t) dt, \\ \sigma_{2,2} &:= \int_0^1 \left(\int_0^1 \mu_X(t)^2 dt - \bar{\mu}_X \mu_X(t) \right)^2 \varrho(t) dt. \end{aligned}$$

Remark 3.1. We remark that, in this case, $\tilde{\lambda}_n$ is again a consistent estimator, in contrast to the case in which $\sigma_n^2 \rightarrow \sigma^2 > 0$.

Remark 3.2. If $b^2 = \beta = 0$ then the limiting normal distributions are degenerate, i.e. $c^2 = 0$ and $\Sigma = \mathbf{0}$. Thus, in this case, we find that $n^{3/2}(\hat{m}_n - m_n) \xrightarrow{P} 0$, $n^{3/2}(\tilde{m}_n - m_n) \xrightarrow{P} 0$, and $n^{1/2}(\tilde{\lambda}_n - \lambda_n) \xrightarrow{P} 0$ as $n \rightarrow \infty$, which means that the norming factors $n^{3/2}$ and $n^{1/2}$ are not appropriate.

Remark 3.3. The result of Sriram’s [15] convergence theorem, i.e. $X^{(n)}/n \xrightarrow{D} \mu_X$, implies only that $n(\hat{m}_n - m_n) \xrightarrow{P} 0$, while, from the convergence $\tilde{X}^{(n)} = (X^{(n)} - E X^{(n)})/n^{1/2} \xrightarrow{D} \tilde{X}$ of Theorem 2.2, we can derive that $n^{3/2}(\hat{m}_n - m_n) \xrightarrow{D} \mathcal{N}(0, c^2)$. A heuristic argument for this goes as follows. On the one hand, we have

$$n(\hat{m}_n - 1) = \frac{\int_0^1 X^{(n)}(t) dX^{(n)}(t) - n\lambda_n \int_0^1 X^{(n)}(t) dt}{\int_0^1 X^{(n)}(t)^2 dt}.$$

From Sriram’s result, we find that

$$n(\hat{m}_n - 1) \xrightarrow{D} \frac{\int_0^1 \mu_X(t) d\mu_X(t) - \lambda \int_0^1 \mu_X(t) dt}{\int_0^1 \mu_X(t)^2 dt}$$

and, from $d\mu_X(t) = (\lambda + \alpha\mu_X(t)) dt$, we conclude that $n(\hat{m}_n - 1) \xrightarrow{D} \alpha$ and, thus,

$$n(\hat{m}_n - m_n) = n(\hat{m}_n - 1) - n(m_n - 1) \xrightarrow{P} 0.$$

On the other hand, we have

$$n^{3/2}(\hat{m}_n - m_n) = \frac{\int_0^1 (n^{-1/2} \tilde{X}^{(n)}(t) + n^{-1} E X^{(n)}(t)) d\tilde{M}^{(n)}(t)}{\int_0^1 (n^{-1/2} \tilde{X}^{(n)}(t) + n^{-1} E X^{(n)}(t))^2 dt}.$$

Theorem 2.2 and the fact that $n^{-1} E X^{(n)}(t) \rightarrow \mu_X(t)$ imply that

$$n^{3/2}(\hat{m}_n - m_n) \xrightarrow{D} \frac{\int_0^1 \mu_X(t) d\tilde{M}(t)}{\int_0^1 \mu_X(t)^2 dt},$$

as stated. The above considerations show that the ‘main term’ of each integrand becomes the nonrandom function μ_X , while the random fluctuation term \tilde{X} disappears as $n \rightarrow \infty$; this results in the asymptotic normality of the estimator \hat{m}_n .

Acknowledgements

This research was supported by the Hungarian Scientific Research Fund grants OTKA-F032060/2000 and OTKA-T032361/2000. The work of the first author was also supported by the János Bolyai Scholarship of the Hungarian Academy of Sciences.

References

- [1] ALIEV, S. A. (1985). A limit theorem for Galton–Watson branching processes with immigration. *Ukrain. Mat. Zh.* **37**, 656–659.
- [2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.
- [3] ETHIER, S. N. AND KURTZ, T. G. (1968). *Markov Processes*. John Wiley, New York.
- [4] GRIMVALL, A. (1974). On the convergence of sequences of branching processes. *Ann. Prob.* **2**, 1027–1045.
- [5] HEYDE, C. C. AND SENETA, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Prob.* **9**, 235–256.
- [6] HEYDE, C. C. AND SENETA, E. (1974). Notes on ‘Estimation theory for growth and immigration rates in a multiplicative process’. *J. Appl. Prob.* **11**, 572–577.
- [7] ISPÁNY, M., PAP, G. AND VAN ZUIJLEN, M. C. A. (2003). Asymptotic behaviour of estimators of the parameters of nearly unstable INAR(1) models. In *Foundations of Statistical Inference*, eds Y. Haitovsky, H. R. Lerche and Y. Ritov, Physica, Heidelberg, pp. 193–204.
- [8] ISPÁNY, M., PAP, G. AND VAN ZUIJLEN, M. C. A. (2003). Asymptotic inference for nearly unstable INAR(1) models. *J. Appl. Prob.* **40**, 750–765.
- [9] JACOD, J. AND SHIRYAYEV, A. N. (2003). *Limit Theorems for Stochastic Processes*. Springer, Berlin.
- [10] KAWAZU, K. AND WATANABE, S. (1971). Branching processes with immigration and related limit theorems. *Theory Prob. Appl.* **16**, 36–54.
- [11] KLIMKO, L. A. AND NELSON, P. I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* **6**, 629–642.
- [12] LAMPERTI, J. (1967). The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitsth.* **7**, 271–288.
- [13] LI, Z.-H. (2000). Ornstein–Uhlenbeck type processes and branching processes with immigration. *J. Appl. Prob.* **37**, 627–634.
- [14] QUINE, M. P. (1976). Asymptotic results for estimators in a subcritical branching process with immigration. *Ann. Prob.* **4**, 319–325.
- [15] SRIRAM, T. N. (1994). Invalidity of bootstrap for critical branching process with immigration. *Ann. Statist.* **22**, 1013–1023.
- [16] WEI, C. Z. AND WINNICKI, J. (1989). Some asymptotic results for the branching process with immigration. *Stoch. Process. Appl.* **31**, 261–282.
- [17] WEI, C. Z. AND WINNICKI, J. (1990). Estimation of the means in the branching process with immigration. *Ann. Statist.* **18**, 1757–1773.