DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN The Netherlands

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

Márton Ispány, Gyula Pap and Martien C. A. van Zuijlen

Report No. 0131 (November 2001)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN
Toernooiveld
6525 ED Nijmegen
The Netherlands

Fluctuation limit of branching processes with immigration and estimation of the means

Márton Ispány and Gyula Pap
Institute of Mathematics and Informatics, University of Debrecen
Pf. 12, H–4010 Debrecen, Hungary

Martien C. A. van Zuijlen
Department of Mathematics, University of Nijmegen
Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Abstract

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

Keywords. Subcritical and nearly critical Galton–Watson branching processes with immigration, conditional least squares estimators, Ornstein–Uhlenbeck type processes, fluctuation limit.

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.

This research has been supported by the Hungarian Scientific Research Fund under Grant No. OTKA–T032361/2000 and Grant No. OTKA–F032060/2000. M. Ispány is also supported by the János Bolyai Scholarship of the Hungarian Academy of Sciences.

Define recursively

$$\begin{cases} X_k = \sum_{j=1}^{X_{k-1}} \xi_{k,j} + \varepsilon_k \quad \text{for} \;\; k \in \mathbb{N}, \\ X_0 = 0. \end{cases} \tag{1.1}$$

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 offsprings of the j^{th} individual in the $(k-1)^{\text{st}}$ generation and ε_k is the number of immigrants contributing to the k^{th} generation. Assume that

$$m := \mathsf{E}\xi_{1,1} < \infty, \qquad \lambda := \mathsf{E}\varepsilon_1 < \infty, \qquad \sigma^2 := \mathsf{Var}\xi_{1,1} < \infty, \qquad b^2 := \mathsf{Var}\varepsilon_1 < \infty.$$

The cases $m < 1$, $m = 1$, $m > 1$ are referred to respectively as **subcritical**, **critical** and **supercritical**.

For $k \in \mathbb{Z}_+$, let $\mathcal{F}_k$ denote the σ–algebra generated by $\{X_0, X_1, \ldots, X_k\}$. Then by (1.1),

$$\mathsf{E}(X_k \mid \mathcal{F}_{k-1}) = mX_{k-1} + \lambda, \qquad k \in \mathbb{N}. \tag{1.2}$$

Clearly,

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

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 regression equation

$$X_k = mX_{k-1} + \lambda + M_k, \qquad \text{for} \;\; k \in \mathbb{N}. \tag{1.3}$$

In the critical case, $m = 1$, Wei and Winnicki [17] proved that for the random step functions

$$\mathcal{X}^{(n)}(t) := X_{[nt]} \qquad \text{for} \;\; t \in \mathbb{R}_+, \;\; n \in \mathbb{N},$$

which can be considered as random elements taking their values in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$, we have

$$\frac{1}{n}\mathcal{X}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{X} \qquad \text{as} \;\; n \to \infty, \tag{1.4}$$

that is, weakly in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$, where $(\mathcal{X}(t))_{t\in\mathbb{R}_+}$ is a (non-negative) diffusion process with generator

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

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

$$\begin{cases} \mathrm{d}\mathcal{X}(t) = \lambda\,\mathrm{d}t + \sigma\sqrt{\mathcal{X}(t)}\,\mathrm{d}W(t), \qquad t \in \mathbb{R}_+, \\ \mathcal{X}(0) = 0, \end{cases}$$

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

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

$$\begin{cases} X_k^{(n)} = \sum_{j=1}^{X_{k-1}^{(n)}} \xi_{k,j}^{(n)} + \varepsilon_k^{(n)} \quad \text{for } k,\, n\in\mathbb{N}, \\ X_0^{(n)} = 0. \end{cases} \tag{1.5}$$

Assume that $m_n := \mathsf{E}\xi_{1,1}^{(n)} < \infty$, $\lambda_n := \mathsf{E}\varepsilon_1^{(n)} < \infty$, $\sigma_n^2 := \mathsf{Var}\xi_{1,1}^{(n)} < \infty$, $b_n^2 := \mathsf{Var}\varepsilon_1^{(n)} < \infty$ for all $n\in\mathbb{N}$. The sequence (1.5) is called **nearly critical** if $m_n \to 1$ as $n\to\infty$. Introduce the random step functions

$$\mathcal{X}^{(n)}(t) := X_{[nt]}^{(n)} \qquad \text{for } t\in\mathbb{R}_+,\ n\in\mathbb{N}.$$

Sriram [15] proved that under the assumptions

(i) $m_n = 1 + \alpha n^{-1} + o(n^{-1})$ as $n\to\infty$ with some $\alpha\in\mathbb{R}$,

(ii) $\sigma_n^2 \to \sigma^2 > 0$ as $n\to\infty$,

(iii) $\mathsf{E}\left(|\xi_{1,1}^{(n)} - m_n|^2 \mathbb{1}_{\{|\xi_{1,1}^{(n)} - m_n| \geqslant \theta\sqrt{n}\}}\right) \to 0$ as $n\to\infty$ for all $\theta > 0$,

(iv) $\lambda_n \to \lambda > 0$ and $b_n^2 \to b^2 > 0$ as $n\to\infty$,

we have

$$\frac{1}{n}\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}_\alpha \qquad \text{as } n\to\infty, \tag{1.6}$$

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

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

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

$$\begin{cases} \mathrm{d}\mathcal{X}_\alpha(t) = (\lambda + \alpha\mathcal{X}_\alpha(t))\,\mathrm{d}t + \sigma\sqrt{\mathcal{X}_\alpha(t)}\,\mathrm{d}W(t), \qquad t\in\mathbb{R}_+, \\ \mathcal{X}_\alpha(0) = 0, \end{cases}$$

In Theorem 2.1 we show that Sriram's result (1.6) is also valid if $\sigma_n^2 \to 0$ (and then condition (iii) is not needed). In this case the limit process $\mathcal{X}_\alpha$ is a deterministic function, namely, $\mathcal{X}_\alpha(t) = \mu_{\mathcal{X}}(t) = \lambda\int_0^t \mathrm{e}^{\alpha s}\,\mathrm{d}s$, $t\in\mathbb{R}_+$, satisfying the (nonrandom) differential equation $\mathrm{d}\mu_{\mathcal{X}}(t) = (\lambda + \alpha\mu_{\mathcal{X}}(t))\,\mathrm{d}t$, $t\geqslant 0$. In fact, this function can be considered as a degenerated, i.e., deterministic diffusion process with generator

$A_\alpha f(x) = (\lambda + \alpha x) f'(x)$, $f \in C_c^\infty(\mathbb{R}_+)$. Remark that 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 (1.6), one can easily obtain the asymptotic behavior of the least squares estimators of m_n and λ_n (see Section 3). (Remark that these statistics have also been investigated in the subcritical and supercritical cases, see Section 3.) We are interested in the asymptotic behaviour of these estimators in a nearly critical case where the offspring variance σ_n^2 tends to zero. For this purpose the limit theorem (1.6) of Sriram does not suffice, as will be explained in Remark 3.4. We have to go on 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 case where $\sigma_n^2 \to 0$, namely, we show that the sequence $(\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)})/\sqrt{n}$ has a limit process $\widetilde{\mathcal{X}}$ as $n \to \infty$. The process $\big(\widetilde{\mathcal{X}}(t)\big)_{t \in \mathbb{R}_+}$ turns out to be an Ornstein–Uhlenbeck type process driven by a time changed Wiener process. We remark that Li [13] proved a similar result for sequences of continuous time discrete state branching processes with immigration. Li [13] applied Laplace transforms, while we have chosen another approach. To explain our method, let $\mathcal{F}_k^{(n)}$ denote the σ–algebra generated by $\big\{X_0^{(n)}, X_1^{(n)}, \ldots, X_k^{(n)}\big\}$ for $n \in \mathbb{N}$ and $k \in \mathbb{Z}_+$. Let

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

Introduce the random step functions

$$\mathcal{M}^{(n)}(t) := \sum_{k=1}^{[nt]} M_k^{(n)} \qquad \text{for } t \in \mathbb{R}_+, \ n \in \mathbb{N}.$$

In order to prove convergence of the sequence $(\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)})/\sqrt{n}$ as $n \to \infty$, first we show, by the help of the martingale central limit theorem, that $\mathcal{M}^{(n)}/\sqrt{n}$ has a limit process $\widetilde{\mathcal{M}}$ as $n \to \infty$, where $\big(\widetilde{\mathcal{M}}(t)\big)_{t \in \mathbb{R}_+}$ is a time-changed Wiener process. Then we show that $(\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)})/\sqrt{n}$ is a function of $\mathcal{M}^{(n)}/\sqrt{n}$, and we use continuous mapping type argument to derive convergence of the sequence $(\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)})/\sqrt{n}$.

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 $(X_k^{(n)})_{k \in \mathbb{Z}_+}$ can not start from zero, and the process $X_{[nt]}^{(n)}$ can be centered by substracting the initial value $X_0^{(n)}$. With suitable normalization, the limiting process will be a zero mean Wiener process, and its variance depends on the limiting behaviour of the offspring variance. In our case the (deterministic) time change mentioned concerning the limit process $\big(\widetilde{\mathcal{M}}(t)\big)_{t \in \mathbb{R}_+}$ is usually not linear, which is the effect of the immigration part. Grimvall [4] not only gave sufficient conditions for the convergence of a suitable normalized sequence $X_{[nt]}^{(n)} - X_0^{(n)}$, 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 be 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 where the offspring variance tends to a positive limit. The norming factor for the offspring mean is $n^{3/2}$, in contrast to the subcritical case where it is $n^{1/2}$, and in contrast to the nearly critical case with positive limiting offspring variance, where it is n. Remark that the results of the present paper are generalizations of those in Ispány et al. [7], [8], where a Bernoulli offspring distribution has been taken.

2 Fluctuation limit theorem

Consider a sequence of branching processes with immigration given in (1.5). First we investigate the asymptotic behaviour of the sequence $\mathcal{X}^{(n)}/n$ in case $\sigma_n^2 \to 0$. We prove the following analogue of Sriram's result (1.6).

2.1 Theorem. *Suppose that*

(i) $m_n = 1 + \alpha n^{-1} + o(n^{-1})$ *as* $n \to \infty$ *with some* $\alpha \in \mathbb{R}$,

(ii) $\sigma_n^2 \to 0$ *as* $n \to \infty$,

(iii) $\lambda_n \to \lambda$ *and* $b_n^2 \to b^2$ *as* $n \to \infty$ *with some* $\lambda \geqslant 0$ *and* $b^2 \geqslant 0$.

Then

$$\frac{1}{n}\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mu_{\mathcal{X}} \qquad \text{as } n \to \infty,$$

that is, weakly in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$, *where*

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

2.2 Remark. If $\lambda = 0$ then the limit function is degenerated, that is, $\mu_{\mathcal{X}}(t) = 0$ for all $t \in \mathbb{R}_+$.

2.3 Remark. Note that in case $\sigma_n^2 \to 0$ no Lindeberg condition (like condition (iii) of Sriram or condition (iii) or (v) of our Theorem 2.4) is needed for the triangular systems $\{\xi_{1,j}^{(n)} : n \in \mathbb{N},\, 1 \leqslant j \leqslant n\}$ or $\{\varepsilon_j^{(n)}/\sqrt{n} : n \in \mathbb{N},\, 1 \leqslant j \leqslant n\}$.

Proof of Theorem 2.1. The theorem can be proved by an argument similar to that in Ethier and Kurtz [3, Chapter 9, Theorem 1.3], where it has been applied for a branching process without immigration. See also Wei and Winnicki [17] in case of a single branching process with immigration, and Sriram [15] in case of a sequence of branching processes with immigration when $\sigma_n^2 \to \sigma^2 > 0$.

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

$$T_n f(x) := \mathsf{E} f\biggl(n^{-1}\Bigl(\sum_{j=1}^{nx} \xi_{1,j}^{(n)} + \varepsilon_1^{(n)}\Bigr)\biggr), \qquad x \in E_n.$$

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

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

where

$$\Delta_n^f(x) := n\bigl(T_n f(x) - f(x)\bigr) - (\lambda + \alpha x) f'(x), \qquad x \in E_n, \quad f \in C_c^\infty(\mathbb{R}_+).$$

Introducing

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

we have $T_n f(x) = \mathsf{E} f\bigl(x + n^{-1}\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr)$. By Taylor's formula,

$$T_n f(x) - f(x) = f'(x) n^{-1} \mathsf{E}\bigl(\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr) + n^{-2} \mathsf{E}\left(\bigl(\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr)^2 \int_0^1 (1-v) f''\bigl(x + v n^{-1}\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr)\,\mathrm{d}v\right).$$

Since for $x \in E_n$

$$\mathsf{E}\widetilde{\mathcal{S}}_{nx}^{(n)} = n(m_n - 1)x + \lambda_n, \tag{2.2}$$

$$\mathsf{E}\bigl(\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr)^2 = n\sigma_n^2 x + b_n^2 + \lambda_n^2 + n^2(m_n-1)^2x^2 + 2n(m_n-1)x\lambda_n, \tag{2.3}$$

we have

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

where

$$\Delta_{n,1}^f(x) := f'(x)\Bigl(\bigl(n(m_n-1) - \alpha\bigr)x + (\lambda_n - \lambda)\Bigr),$$

$$\Delta_{n,2}^f(x) := n^{-1}\mathsf{E}\left(\bigl(\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr)^2 \int_0^1 (1-v)\Bigl(f''\bigl(x + v n^{-1}\widetilde{\mathcal{S}}_{nx}^{(n)}\bigr) - f''(x)\Bigr)\,\mathrm{d}v\right),$$

$$\Delta_{n,3}^f(x) := \frac{1}{2} x f''(x)\sigma_n^2 + \frac{1}{2n} f''(x)\Bigl(b_n^2 + \lambda_n^2 + n^2(m_n-1)^2x^2 + 2n(m_n-1)x\lambda_n\Bigr).$$

To prove (2.1), it is enough to show $\lim_{n\to\infty} \Delta_n^f(x_n) = 0$ for every sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \in E_n$, $n \in \mathbb{N}$, such that $x_n \to x \in [0, +\infty]$. Assumptions (i)–(iii) clearly imply $\lim_{n\to\infty} \Delta_{n,i}^f(x_n) = 0$ for $i = 1$ and $i = 3$ and for all such sequences

$x_n \to x \in [0, +\infty]$. In order to deal with $\Delta^f_{n,2}(x_n)$, suppose that the support of f is contained in $[0, c]$. Since

$$x + vn^{-1}\widetilde{\mathcal{S}}^{(n)}_{nx} = x + vn^{-1}\left(\sum_{j=1}^{nx}(\xi^{(n)}_{1,j} - 1) + \varepsilon^{(n)}_1\right) \geqslant x(1 - v),$$

the integrand in $\Delta^f_{n,2}(x)$ is zero if $v \leqslant 1 - c/x$. Consequently,

$$|\Delta^f_{n,2}(x_n)| \leqslant n^{-1}\|f''\|_\infty\big((c/x_n) \wedge 1\big)^2 \mathsf{E}\big(\widetilde{\mathcal{S}}^{(n)}_{nx_n}\big)^2, \qquad (2.4)$$

where $\|\cdot\|_\infty$ denotes the supremum norm. Using (2.3), one can easily check that the right hand side of (2.4) tends to 0 for all sequences $x_n \to x \in [0, +\infty]$. Thus we conclude $\lim_{n\to\infty} \Delta^f_{n,2}(x_n) = 0$, hence finally we obtain (2.1). □

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

2.4 Theorem. *Suppose that*

(i) $m_n = 1 + \alpha n^{-1} + o(n^{-1})$ *as* $n \to \infty$ *with some* $\alpha \in \mathbb{R}$,

(ii) $\sigma_n^2 = \beta n^{-1} + o(n^{-1})$ *as* $n \to \infty$ *with some* $\beta \geqslant 0$,

(iii) $n\mathsf{E}\left(|\xi^{(n)}_{1,1} - m_n|^2 \mathbb{1}_{\{|\xi^{(n)}_{1,1} - m_n| \geqslant \theta\}}\right) \to 0$ *as* $n \to \infty$ *for all* $\theta > 0$,

(iv) $\lambda_n \to \lambda$ *and* $b_n^2 \to b^2$ *as* $n \to \infty$ *with some* $\lambda \geqslant 0$ *and* $b^2 \geqslant 0$,

(v) $\mathsf{E}\left(|\varepsilon^{(n)}_1 - \lambda_n|^2 \mathbb{1}_{\{|\varepsilon^{(n)}_1 - \lambda_n| \geqslant \theta\sqrt{n}\}}\right) \to 0$ *as* $n \to \infty$ *for all* $\theta > 0$.

Then

$$\left(\frac{\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)}}{\sqrt{n}}, \frac{\mathcal{M}^{(n)}}{\sqrt{n}}\right) \xrightarrow{\mathcal{D}} (\widetilde{\mathcal{X}}, \widetilde{\mathcal{M}}) \qquad \text{as } n \to \infty,$$

that is, weakly in the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}^2)$, *where* $\big(\widetilde{\mathcal{M}}(t)\big)_{t\in\mathbb{R}_+}$ *is a time–changed Wiener process, namely,*

$$\widetilde{\mathcal{M}}(t) = W(T(t)), \quad t \in \mathbb{R}_+$$

with

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

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

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

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

2.5 Remark. If $b^2 = 0$ and $\beta\lambda = 0$ then the limit processes are degenerated, that is, $\widetilde{\mathcal{X}}(t) = \widetilde{\mathcal{M}}(t) = 0$ for all $t \in \mathbb{R}_+$.

2.6 Remark. Conditions (iii) and (v) are, in fact, the Lindeberg conditions for the triangular systems $\{\xi_{1,j}^{(n)} : n \in \mathbb{N}, 1 \leqslant j \leqslant n\}$ and $\{\varepsilon_j^{(n)}/\sqrt{n} : n \in \mathbb{N}, 1 \leqslant j \leqslant n\}$, respectively. See also Grimvall [4] who investigated fluctuation theorems for sequences of branching processes without immigration, and the remarks in the Introduction. Clearly, if there exists $\gamma > 0$ such that $n\mathsf{E}|\xi_{1,1}^{(n)} - m_n|^{2+\gamma} \to 0$ and $n^{-\gamma/2}\mathsf{E}|\varepsilon_1^{(n)} - \lambda_n|^{2+\gamma} \to 0$ as $n \to \infty$ then conditions (iii) and (v) are satisfied.

2.7 Remark. We remark that $\big(\widetilde{\mathcal{M}}(t)\big)_{t\in\mathbb{R}_+}$ 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 $\widetilde{\mathcal{M}}(t) = W(b^2t)$, $t \in \mathbb{R}_+$, is a Wiener process. The process $\big(\widetilde{\mathcal{M}}(t)\big)_{t\in\mathbb{R}_+}$ is always a martingale, so that we can define stochastic integrals with respect to it. Its covariance function has the form

$$\mathsf{Cov}(\widetilde{\mathcal{M}}(s), \widetilde{\mathcal{M}}(t)) = T(s \wedge t) \qquad \text{for } s, t \in \mathbb{R}_+.$$

Comparing the covariance structures we obtain another representation of the process in the form

$$\widetilde{\mathcal{M}}(t) = \int_0^t \sqrt{\varrho(s)}\,\mathrm{d}W(s) \qquad \text{for } t \in \mathbb{R}_+,$$

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

$$\begin{cases} \mathrm{d}\widetilde{\mathcal{M}}(t) = \sqrt{\varrho(t)}\,\mathrm{d}W(t), & t \geqslant 0, \\ \widetilde{\mathcal{M}}(0) = 0. \end{cases}$$

The process $\big(\widetilde{\mathcal{X}}(t)\big)_{t\in\mathbb{R}_+}$ is a continuous zero mean Gaussian martingale with covariance function

$$\mathsf{Cov}(\widetilde{\mathcal{X}}(s), \widetilde{\mathcal{X}}(t)) = \int_0^{s\wedge t} \mathrm{e}^{\alpha(s+t-2u)}\varrho(u)\,\mathrm{d}u \qquad \text{for } s, t \in \mathbb{R}_+.$$

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

$$\widetilde{\mathcal{X}}(t) = \mathrm{e}^{\alpha t}\int_0^t \mathrm{e}^{-\alpha s}\sqrt{\varrho(s)}\,\mathrm{d}W(s) \qquad \text{for } t \in \mathbb{R}_+.$$

This implies that for the process $\widetilde{\mathcal{Y}}(t) := \mathrm{e}^{-\alpha t}\widetilde{\mathcal{X}}(t)$, $t \geqslant 0$, we have

$$\mathrm{d}\widetilde{\mathcal{Y}}(t) = \mathrm{e}^{-\alpha t}\sqrt{\varrho(t)}\,\mathrm{d}W(t).$$

By Itô's formula, we obtain that the process $\widetilde{\mathcal{X}}(t) = \mathrm{e}^{\alpha t}\widetilde{\mathcal{Y}}(t)$ is the unique solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}\widetilde{\mathcal{X}}(t) = \alpha\widetilde{\mathcal{X}}(t)\,\mathrm{d}t + \sqrt{\varrho(t)}\,\mathrm{d}W(t), \quad t \geqslant 0, \\ \widetilde{\mathcal{X}}(0) = 0. \end{cases}$$

In order to prove Theorem 2.4 we need formulas for $\mathsf{E}X_k^{(n)}$, $\mathsf{Cov}\big(X_k^{(n)}, X_\ell^{(n)}\big)$ and $\mathsf{E}\Big(\big(M_k^{(n)}\big)^2 \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big)$.

2.8 Lemma. *Let* $(X_k)_{k\in\mathbb{Z}_+}$ *be a branching processes with immigration given in* (1.1)*. Then for all* $k \in \mathbb{Z}_+$*,*

$$\mathsf{E}X_k = \begin{cases} \dfrac{m^k - 1}{m - 1}\lambda & \text{if } m \neq 1, \\ k\lambda & \text{if } m = 1, \end{cases}$$

$$\mathsf{Var}X_k = \begin{cases} \dfrac{m^{2k} - 1}{m^2 - 1}b^2 + \dfrac{(m^k - 1)(m^{k-1} - 1)}{(m - 1)(m^2 - 1)}\lambda\sigma^2 & \text{if } m \neq 1, \\ kb^2 + \dfrac{k(k-1)}{2}\lambda\sigma^2 & \text{if } m = 1. \end{cases}$$

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

$$\mathsf{Cov}(X_k,\, X_\ell) = m^{|k-\ell|}\mathsf{Var}X_{k\wedge\ell}.$$

Furthermore, for all $k \in \mathbb{Z}_+$*,*

$$\mathsf{E}\big(M_k^2 \,\big|\, \mathcal{F}_{k-1}\big) = \sigma^2 X_{k-1} + b^2.$$

Proof. By (1.2), we obtain the recursion

$$\mathsf{E}X_k = m\mathsf{E}X_{k-1} + \lambda \qquad \text{for } k \in \mathbb{N}. \tag{2.5}$$

Moreover,

$$\mathsf{Var}X_k = m^2\mathsf{Var}X_{k-1} + \sigma^2\mathsf{E}X_{k-1} + b^2 \qquad \text{for } k \in \mathbb{N}. \tag{2.6}$$

Indeed, by (1.1),

$$\begin{aligned} \mathsf{E}\big((X_k - \mathsf{E}X_k)^2 \,\big|\, X_{k-1}\big) &= \mathsf{E}\big((X_k - m\mathsf{E}X_{k-1} - \lambda)^2 \mid X_{k-1}\big) \\ &= \mathsf{E}\Bigg(\bigg(\sum_{j=1}^{X_{k-1}} \xi_{k,j} + \varepsilon_k - m\mathsf{E}X_{k-1} - \lambda\bigg)^2 \,\Bigg|\, X_{k-1}\Bigg) \\ &= \mathsf{E}\Bigg(\bigg(\sum_{j=1}^{X_{k-1}} (\xi_{k,j} - m) + m(X_{k-1} - \mathsf{E}X_{k-1}) + (\varepsilon_k - \lambda)\bigg)^2 \,\Bigg|\, X_{k-1}\Bigg). \end{aligned}$$

The random variables $\xi_{k,j} - m$ and $\varepsilon_k - \lambda$ are independent of X_{k-1} and have zero mean and variances σ^2 and b^2, respectively. Consequently,

$$\mathsf{E}\big((X_k - \mathsf{E}X_k)^2 \,\big|\, X_{k-1}\big) = \sigma^2 X_{k-1} + m^2 (X_{k-1} - \mathsf{E}X_{k-1})^2 + b^2,$$

which implies (2.6). From (2.5) and (2.6) we obtain the vector recursion

$$\begin{pmatrix} \mathsf{E}X_k \\ \mathsf{Var}X_k \end{pmatrix} = \begin{pmatrix} m & 0 \\ \sigma^2 & m^2 \end{pmatrix} \begin{pmatrix} \mathsf{E}X_{k-1} \\ \mathsf{Var}X_{k-1} \end{pmatrix} + \begin{pmatrix} \lambda \\ b^2 \end{pmatrix} \qquad \text{for } n \in \mathbb{N}.$$

Obviously, $\mathsf{E}X_0 = \mathsf{Var}X_0 = 0$, hence

$$\begin{pmatrix} \mathsf{E}X_k \\ \mathsf{Var}X_k \end{pmatrix} = \sum_{j=0}^{k-1} \begin{pmatrix} m & 0 \\ \sigma^2 & m^2 \end{pmatrix}^j \begin{pmatrix} \lambda \\ b^2 \end{pmatrix} \qquad \text{for } n \in \mathbb{N}.$$

Clearly

$$\begin{pmatrix} m & 0 \\ \sigma^2 & m^2 \end{pmatrix}^j = \begin{pmatrix} m^j & 0 \\ \sigma^2 \sum\limits_{i=j-1}^{2j-2} m^i & m^{2j} \end{pmatrix} \qquad \text{for } j \in \mathbb{N}.$$

Hence we conclude that

$$\begin{aligned} \mathsf{E}X_k &= \lambda \sum_{j=0}^{k-1} m^j, \\ \mathsf{Var}X_k &= b^2 \sum_{j=0}^{k-1} m^{2j} + \lambda\sigma^2 \sum_{j=1}^{k-1} \sum_{i=j-1}^{2j-2} m^i, \end{aligned}$$

which imply the formulas for $\mathsf{E}X_k$ and $\mathsf{Var}X_k$.

The formula for the covariances $\mathsf{Cov}(X_k, X_\ell)$ follows from the recursion

$$\mathsf{Cov}(X_k, X_\ell) = m\mathsf{Cov}(X_k, X_{\ell-1}) \qquad \text{for } 0 \leqslant k < \ell. \tag{2.7}$$

Indeed, by (2.5),

$$\begin{aligned} &\mathsf{E}\big((X_k - \mathsf{E}X_k)(X_\ell - \mathsf{E}X_\ell) \,\big|\, \mathcal{F}_{\ell-1}\big) = (X_k - \mathsf{E}X_k)\mathsf{E}(X_\ell - m\mathsf{E}X_{\ell-1} - \lambda \,|\, \mathcal{F}_{\ell-1}) \\ &= (X_k - \mathsf{E}X_k)\mathsf{E}\Bigg(\sum_{j=1}^{X_{\ell-1}} \xi_{\ell,j} + \varepsilon_\ell - m\mathsf{E}X_{\ell-1} - \lambda \,\Bigg|\, \mathcal{F}_{\ell-1} \Bigg) \\ &= (X_k - \mathsf{E}X_k)\mathsf{E}\Bigg(\sum_{j=1}^{X_{\ell-1}} (\xi_{\ell,j} - m) + m(X_{\ell-1} - \mathsf{E}X_{\ell-1}) + (\varepsilon_\ell - \lambda) \,\Bigg|\, \mathcal{F}_{\ell-1} \Bigg) \\ &= m(X_k - \mathsf{E}X_k)(X_{\ell-1} - \mathsf{E}X_{\ell-1}), \end{aligned}$$

which implies (2.7).

Finally,

$$\begin{aligned}\mathsf{E}\big(M_k^2 \,|\, \mathcal{F}_{k-1}\big) &= \mathsf{E}\big((X_k - mX_{k-1} - \lambda)^2 \,\big|\, \mathcal{F}_{k-1}\big)\\ &= \mathsf{E}\Bigg(\bigg(\sum_{j=1}^{X_{k-1}} (\xi_{k,j} - m) + (\varepsilon_k - \lambda)\bigg)^2 \,\Bigg|\, \mathcal{F}_{k-1}\Bigg) = \sigma^2 X_{k-1} + b^2,\end{aligned}$$

and we finished the proof of the lemma. □

We remark that $\mathsf{E}X_k$, $\mathsf{Var}X_k$ and $\mathsf{Cov}(X_k,\, X_\ell)$ continuously depend on m.

Proof of Theorem 2.4. We will make the following steps:

- **(A)** we prove that $\widetilde{\mathcal{M}}^{(n)} := \mathcal{M}^{(n)}/\sqrt{n} \stackrel{\mathcal{D}}{\longrightarrow} \widetilde{\mathcal{M}}$ by the help of the martingale central limit theorem;
- **(B)** we show that $\widetilde{\mathcal{X}}^{(n)} := \big(\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)}\big)/\sqrt{n} = \Phi_n\big(\widetilde{\mathcal{M}}^{(n)}\big)$ with some measurable mappings $\Phi_n : D(\mathbb{R}_+,\mathbb{R}) \to D(\mathbb{R}_+,\mathbb{R})$ such that $\Phi_n \to \Phi$ in an appropriate sense, where $\Phi : D(\mathbb{R}_+,\mathbb{R}) \to D(\mathbb{R}_+,\mathbb{R})$ is a measurable mapping, implying that $\big(\widetilde{\mathcal{X}}^{(n)}, \widetilde{\mathcal{M}}^{(n)}\big) \stackrel{\mathcal{D}}{\longrightarrow} \big(\Phi(\widetilde{\mathcal{M}}), \widetilde{\mathcal{M}}\big)$;
- **(C)** we derive that $\Phi(\widetilde{\mathcal{M}}) = \widetilde{\mathcal{X}}$.

(A). By the martingale central limit theorem (see, e.g. Jacod and Shiryaev [9, Theorem VIII. 3.33]), it suffices to prove that for all $t \geqslant 0$,

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(M_k^{(n)}\big)^2 \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \stackrel{\mathsf{P}}{\longrightarrow} T(t) \qquad \text{as } n \to \infty, \tag{2.8}$$

$$\forall\, \theta > 0 \qquad \frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(M_k^{(n)}\big)^2 \mathbb{1}_{\{|M_k^{(n)}| > \theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \stackrel{\mathsf{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty, \tag{2.9}$$

By Lemma 2.8,

$$\mathsf{E}\Big(\big(M_k^{(n)}\big)^2 \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) = \sigma_n^2 X_{k-1}^{(n)} + b_n^2.$$

Thus, in order to prove (2.8), we have to show that

$$\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} X_{k-1}^{(n)} \stackrel{\mathsf{P}}{\longrightarrow} \beta\lambda \int_0^t \bigg(\int_0^v \mathrm{e}^{\alpha u}\,\mathrm{d}u\bigg)\mathrm{d}v \qquad \text{as } n \to \infty. \tag{2.10}$$

This statement will clearly follow once we prove

$$\mathsf{E}\bigg(\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} X_{k-1}^{(n)}\bigg) \to \beta\lambda \int_0^t \bigg(\int_0^v \mathrm{e}^{\alpha u}\,\mathrm{d}u\bigg)\mathrm{d}v \qquad \text{as } n \to \infty, \tag{2.11}$$

$$\mathsf{Var}\bigg(\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} X_{k-1}^{(n)}\bigg) \to 0 \qquad \text{as } n \to \infty. \tag{2.12}$$

If $m_n = 1$ then by Lemma 2.8,

$$\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} \mathsf{E} X_{k-1}^{(n)} = \frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]}(k-1)\lambda_n = n\sigma_n^2\lambda_n\frac{[nt]([nt]-1)}{2n^2}.$$

Hence, along a subsequence with $m_n = 1$, we have

$$\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} \mathsf{E} X_{k-1}^{(n)} \to \frac{1}{2}\beta\lambda t^2 \qquad \text{as } n\to\infty. \tag{2.13}$$

If $m_n \neq 1$ then again by Lemma 2.8,

$$\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} \mathsf{E} X_{k-1}^{(n)} = \frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]}\frac{m_n^{k-1}-1}{m_n-1}\lambda_n = \frac{n\sigma_n^2\lambda_n}{n(m_n-1)}\left(\frac{m_n^{[nt]}-1}{n(m_n-1)} - \frac{[nt]}{n}\right).$$

On one hand, assumption (i) implies that

$$n(m_n-1) \to \alpha \qquad \text{as } n\to\infty. \tag{2.14}$$

On the other hand, for sufficiently large $n\in\mathbb{N}$, $m_n = \mathrm{e}^{\alpha_n/n}$, where $\alpha_n\to\alpha$ as $n\to\infty$, since

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

Thus

$$m_n^{[nt]} = \mathrm{e}^{\alpha_n[nt]/n} \to \mathrm{e}^{\alpha t} \qquad \text{as } n\to\infty. \tag{2.15}$$

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

$$\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} \mathsf{E} X_{k-1}^{(n)} \to \begin{cases} \dfrac{\beta\lambda}{\alpha}\left(\dfrac{\mathrm{e}^{\alpha t}-1}{\alpha} - t\right) & \text{if } \alpha\neq 0,\\[2mm] \dfrac{1}{2}\beta\lambda t^2, & \text{if } \alpha = 0.\end{cases}$$

Taking this convergence and (2.13) into account, we conclude (2.11).

In order to prove (2.12) first we note that by Lemma 2.8,

$$\begin{aligned}\mathsf{Var}\left(\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} X_{k-1}^{(n)}\right) &= \frac{\sigma_n^4}{n^2}\sum_{k=1}^{[nt]}\sum_{\ell=1}^{[nt]} m_n^{k-\ell}\mathsf{Var}\big(X_{k\wedge\ell-1}^{(n)}\big)\\ &= \frac{\sigma_n^4}{n^2}\left(\sum_{k=1}^{[nt]}\mathsf{Var}\big(X_{k-1}^{(n)}\big) + 2\sum_{k=1}^{[nt]-1}\sum_{j=1}^{[nt]-k} m_n^j\mathsf{Var}\big(X_{k-1}^{(n)}\big)\right)\\ &= \frac{\sigma_n^4}{n^2}\sum_{k=1}^{[nt]}\left(2\sum_{j=0}^{[nt]-k} m_n^j - 1\right)\mathsf{Var}\big(X_{k-1}^{(n)}\big).\end{aligned}$$

Thus

$$\operatorname{Var}\biggl(\frac{\sigma_n^2}{n}\sum_{k=1}^{[nt]} X_{k-1}^{(n)}\biggr) = \bigl(U_n(t)b_n^2 + V_n(t)\lambda_n\sigma_n^2\bigr)\frac{\sigma_n^4}{n^2}, \tag{2.16}$$

where

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

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

If $m_n = 1$ then

$$U_n(t) = \sum_{k=1}^{[nt]}(k-1)\bigl(2[nt]-2k+1\bigr) = \frac{[nt]\bigl([nt]-1\bigr)\bigl(2[nt]-1\bigr)}{6},$$

hence for all $t \in \mathbb{R}_+$,

$$\frac{U_n(t)}{n^3} \to \frac{t^3}{3} \qquad \text{as } n \to \infty.$$

Moreover,

$$V_n(t) = \frac{1}{2}\sum_{k=1}^{[nt]}(k-1)(k-2)\bigl(2[nt]-2k+1\bigr) = \frac{[nt]\bigl([nt]-1\bigr)^2\bigl([nt]-2\bigr)}{12},$$

hence for all $t \in \mathbb{R}_+$,

$$\frac{V_n(t)}{n^4} \to \frac{t^4}{12} \qquad \text{as } n \to \infty.$$

Thus, taking into account assumptions (ii) and (iv), for all $t \in \mathbb{R}_+$, along a subsequence with $m_n = 1$,

$$\bigl(U_n(t)b_n^2 + V_n(t)\lambda_n\sigma_n^2\bigr)\frac{\sigma_n^4}{n^2} = \frac{n^2\sigma_n^4 b_n^2}{n}\frac{U_n(t)}{n^3} + \frac{n^3\sigma_n^6\lambda_n}{n}\frac{V_n(t)}{n^4} \to 0 \qquad \text{as } n \to \infty. \tag{2.17}$$

If $m_n \neq 1$ then

$$U_n(t) = \frac{1}{(m_n-1)^2(m_n+1)} \sum_{k=1}^{[nt]} (m_n^{2k-2} - 1)\big(2m^{[nt]-k+1} - (m_n+1)\big)$$
$$= \frac{1}{(m_n-1)^2(m_n+1)} \sum_{k=1}^{[nt]} \big(2m_n^{[nt]+k-1} - 2m_n^{[nt]-k+1} - (m_n+1)m_n^{2k-2} + (m_n+1)\big)$$
$$= \frac{1}{(m_n-1)^2(m_n+1)} \left(2\frac{m_n^{2[nt]} - m_n^{[nt]}}{m_n-1} - 2\frac{m_n^{[nt]+1} - m_n}{m_n-1} - \frac{m_n^{2[nt]} - 1}{m_n-1} + (m_n+1)[nt]\right).$$

Using (2.14) and (2.15)

$$\frac{U_n(t)}{n^3} \to \begin{cases} \frac{1}{2\alpha^2}\left(2\frac{\mathrm{e}^{2\alpha t} - \mathrm{e}^{\alpha t}}{\alpha} - 2\frac{\mathrm{e}^{\alpha t} - 1}{\alpha} - \frac{\mathrm{e}^{2\alpha t} - 1}{\alpha} + 2t\right) = \frac{\mathrm{e}^{2\alpha t} - 4\mathrm{e}^{\alpha t} + 2\alpha t + 3}{2\alpha^3} & \text{if } \alpha \neq 0, \\ \frac{t^3}{3} & \text{if } \alpha = 0. \end{cases}$$

In a similar way,

$$\frac{V_n(t)}{n^4} \to \begin{cases} \frac{\mathrm{e}^{2\alpha t} - 4(\alpha t - 1)\mathrm{e}^{\alpha t} - 2\alpha t - 5}{2\alpha^4} & \text{if } \alpha \neq 0, \\ \frac{t^4}{12} & \text{if } \alpha = 0. \end{cases}$$

Hence, for all $t \in \mathbb{R}_+$, along a subsequence with $m_n \neq 1$, we obtain again (2.17). By (2.16), we conclude (2.12), and finally (2.8). (Note that $\lim_{n\to\infty} U_n(t)/n^3$ and $\lim_{n\to\infty} V_n(t)/n^4$ depend continuously on α.)

To prove the conditional Lindeberg condition (2.9) we consider the decomposition

$$M_k^{(n)} = \sum_{j=1}^{X_{k-1}^{(n)}} \xi_{k,j}^{(n)} + \varepsilon_k^{(n)} - m_n X_{k-1}^{(n)} - \lambda_n = N_k^{(n)} + \delta_k^{(n)},$$

where

$$N_k^{(n)} := \sum_{j=1}^{X_{k-1}^{(n)}} (\xi_{k,j}^{(n)} - m_n), \qquad \delta_k^{(n)} := \varepsilon_k^{(n)} - \lambda_n.$$

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

$$\mathbb{1}_{\{|Y+Z|>\theta\}} \leqslant \mathbb{1}_{\{|Y|>\theta/2\}} + \mathbb{1}_{\{|Z|>\theta/2\}}.$$

Hence, it suffices to show

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(N_k^{(n)}\big)^2 \mathbb{1}_{\{|N_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \stackrel{\mathsf{P}}{\longrightarrow} 0 \qquad \text{as} \quad n\to\infty, \tag{2.18}$$

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(N_k^{(n)}\big)^2 \mathbb{1}_{\{|\delta_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \stackrel{\mathsf{P}}{\longrightarrow} 0 \qquad \text{as} \quad n\to\infty, \tag{2.19}$$

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(\delta_k^{(n)}\big)^2 \mathbb{1}_{\{|N_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \stackrel{\mathsf{P}}{\longrightarrow} 0 \qquad \text{as} \quad n\to\infty, \tag{2.20}$$

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(\delta_k^{(n)}\big)^2 \mathbb{1}_{\{|\delta_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \stackrel{\mathsf{P}}{\longrightarrow} 0 \qquad \text{as} \quad n\to\infty \tag{2.21}$$

for all $\theta>0$ and all $t>0$.

To prove (2.18), introduce the random step functions

$$\mathcal{S}_n(t) := \sum_{j=1}^{[nt]} (\xi_{1,j}^{(n)} - m_n) \qquad \text{for} \quad t\in\mathbb{R}_+, \quad n\in\mathbb{N}.$$

We note that conditions (i) and (ii) imply that $\mathsf{E}\mathcal{S}_n(t)=0$ and $\mathsf{Var}(\mathcal{S}_n(t)) = [nt]\sigma_n^2 \to \beta t$. Together with the Lindeberg condition (iii), this garantees that

$$\mathcal{S}_n \stackrel{\mathcal{D}}{\longrightarrow} W_\beta \qquad \text{as} \quad n\to\infty,$$

where $(W_\beta(t))_{t\in\mathbb{R}_+}$ is a Wiener process with $\mathsf{E}W_\beta(t)=0$ and $\mathsf{Var}W_\beta(t)=\beta t$, $t\in\mathbb{R}_+$. Moreover,

$$\begin{aligned}
&\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\big(N_k^{(n)}\big)^2 \mathbb{1}_{\{|N_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big)\\
&= \frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\bigg(\Big(\sum_{j=1}^{\ell}(\xi_{k,j}^{(n)} - m_n)\Big)^2 \mathbb{1}_{\{|\sum_{j=1}^{\ell}(\xi_{k,j}^{(n)}-m_n)|>\theta\sqrt{n}\}}\bigg)\bigg|_{\ell=X_{k-1}^{(n)}}\\
&= \frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\mathcal{S}_n\left(\tfrac{\ell}{n}\right)^2 \mathbb{1}_{\{|\mathcal{S}_n(\frac{\ell}{n})|>\theta\sqrt{n}\}}\Big)\Big|_{\ell=X_{k-1}^{(n)}} = F_n\Big(\frac{1}{n}\mathcal{X}^{(n)}\Big),
\end{aligned}$$

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

$$F_n(x) := \frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}\Big(\mathcal{S}_n\left(x\left(\tfrac{k-1}{n}\right)\right)^2 \mathbb{1}_{\{|\mathcal{S}_n(x(\frac{k-1}{n}))|>\theta\sqrt{n}\}}\Big) \qquad \text{for} \quad x\in D([0,t],\mathbb{R}).$$

By Theorem 2.1, we have

$$\frac{1}{n}\mathcal{X}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \mu_{\mathcal{X}} \qquad \text{as } n \to \infty,$$

where $\mu_{\mathcal{X}}$ is a continuous function. In view of the continuous mapping theorem (see Billingsley [2, Theorem 5.5]), in order to prove (2.18) it suffices to show that

$$F_n(x_n) \to 0 \ \text{ if } \ x \in C([0,t],\mathbb{R}), \ x_n \in D([0,t],\mathbb{R}) \ \text{ with } \ \|x_n - x\|_\infty \to 0 \ \text{ as } \ n \to \infty.$$

(We used that in this case $x_n \to x$ in $D([0,t],\mathbb{R})$ if and only if $\|x_n - x\|_\infty \to 0$, see, e.g. Jacod and Shiryaev [9, VI,1.17].) We have

$$F_n(x_n) = \mathsf{E} G_n(\mathcal{S}_n),$$

where the measurable mapping $G_n : D([0,L],\mathbb{R}) \to \mathbb{R}$ is defined by

$$G_n(y) := \frac{1}{n}\sum_{k=1}^{[nt]} y\left(x_n\left(\tfrac{k-1}{n}\right)\right)^2 \mathbb{1}_{\{|y(x_n(\frac{k-1}{n}))|>\theta\sqrt{n}\}} \qquad \text{for } y \in D([0,L],\mathbb{R}),$$

and

$$L := \sup_{n\in\mathbb{N}} \sup_{s\in[0,t]} |x_n(s)| < \infty.$$

Obviously

$$G_n(y) \leqslant H_n(y) \qquad \text{for all } y \in D([0,L],\mathbb{R}),$$

where the measurable mapping $H_n : D([0,L],\mathbb{R}) \to \mathbb{R}$ is defined by

$$H_n(y) := \frac{1}{n}\sum_{k=1}^{[nt]} y\left(x_n\left(\tfrac{k-1}{n}\right)\right)^2 \qquad \text{for } y \in D([0,L],\mathbb{R}).$$

In view of the dominated convergence theorem (see, e.g. Ethier and Kurtz [3, Appendices, Theorem 1.2]), in order to show $\mathsf{E} G_n(\mathcal{S}_n) \to 0$ it suffices to prove that

(a) $G_n(\mathcal{S}_n) \stackrel{\mathcal{D}}{\longrightarrow} 0$ as $n \to \infty$,

(b) $H_n(\mathcal{S}_n) \stackrel{\mathcal{D}}{\longrightarrow} H(W_\beta)$ as $n \to \infty$, where the measurable mapping $H : D([0,L],\mathbb{R}) \to \mathbb{R}$ is defined by

$$H(y) := \int_0^t y(x(s))^2 \, \mathrm{d}s \qquad \text{for } y \in D([0,L],\mathbb{R}),$$

(c) $\mathsf{E} H_n(\mathcal{S}_n) \to \mathsf{E} H(W_\beta)$ as $n \to \infty$.

To prove (a) and (b), we use again the continuous mapping theorem. Since $\mathcal{S}_n \xrightarrow{\mathcal{D}} W_\beta$ and almost all trajectories of W_β are continuous, it suffices to show that

$$G_n(y_n) \to 0 \quad \text{if} \quad y \in C([0,L],\mathbb{R}), \ \ y_n \in D([0,L],\mathbb{R}) \ \ \text{with} \ \ \|y_n - y\|_\infty \to 0 \ \ \text{as} \ \ n \to \infty.$$

We have

$$G_n(y_n) = \frac{1}{n} \sum_{k=1}^{[nt]} y_n\left(x_n\left(\tfrac{k-1}{n}\right)\right)^2 \mathbb{1}_{\{|y_n(x_n(\frac{k-1}{n}))| > \theta\sqrt{n}\}}\Big) = 0$$

for sufficiently large $n \in \mathbb{N}$, since

$$\sup_{n\in\mathbb{N}} \|y_n \circ x_n\|_\infty < \infty.$$

Indeed, $\|y_n \circ x_n\|_\infty \leqslant \|y_n\|_\infty \leqslant \|y_n - y\|_\infty + \|y\|_\infty$, where $\|y_n - y\|_\infty \to 0$ as $n \to \infty$, and $\|y\|_\infty < \infty$. Hence we conclude (a). To prove (b), it is enough to check that

$$H_n(y_n) \to H(y) \quad \text{if} \quad y \in C([0,L],\mathbb{R}), \ \ y_n \in D([0,L],\mathbb{R}) \ \ \text{with} \ \ \|y_n - y\|_\infty \to 0 \ \ \text{as} \ \ n \to \infty.$$

Moreover, $\mathsf{E}\mathcal{S}_n(s)^2 = [ns]\sigma_n^2$ implies

$$\mathsf{E}H_n(\mathcal{S}_n) = \frac{\sigma_n^2}{n} \sum_{k=1}^{[nt]} \left[n x_n\left(\tfrac{k-1}{n}\right)\right] \to \beta \int_0^t x(s)\,\mathrm{d}s.$$

By Fubini's theorem,

$$\mathsf{E}H(W_\beta) = \int_0^t \mathsf{E}W_\beta(x(s))^2\,\mathrm{d}s = \beta \int_0^t x(s)\,\mathrm{d}s,$$

hence we obtain (c). Consequently, we conclude $\mathsf{E}G_n(\mathcal{S}_n) \to 0$, which implies $F_n(x_n) \to 0$, and finally, we obtain (2.18).

To prove (2.19) we note that

$$\mathsf{E}\Big(\big(N_k^{(n)}\big)^2 \mathbb{1}_{\{|\delta_k^{(n)}| > \theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) = \mathsf{E}\Bigg(\Bigg(\sum_{j=1}^{X_{k-1}^{(n)}} (\xi_{k,j} - m_n)\Bigg)^2 \mathbb{1}_{\{|\delta_k^{(n)}| > \theta\sqrt{n}\}} \,\Bigg|\, \mathcal{F}_{k-1}^{(n)}\Bigg)$$
$$= \sigma_n^2 X_{k-1}^{(n)} \mathsf{E}\mathbb{1}_{\{|\delta_k^{(n)}| > \theta\sqrt{n}\}} = \sigma_n^2 X_{k-1}^{(n)} \mathsf{P}\big(|\varepsilon_1^{(n)} - \lambda_n| > \theta\sqrt{n}\big).$$

Moreover

$$\mathsf{P}\big(|\varepsilon_1^{(n)} - \lambda_n| > \theta\sqrt{n}\big) \leqslant \theta^{-2} n^{-1} \mathsf{E}\Big(\big(\varepsilon_1^{(n)} - \lambda_n\big)^2 \mathbb{1}_{\{|\varepsilon_1^{(n)} - \lambda_n| > \theta\sqrt{n}\}}\Big),$$

hence (2.19) is a consequence of (2.10) and the assumptions (ii) and (v).

In order to show (2.20) we use the estimate

$$\begin{aligned}
&\mathsf{E}\Big(\big(\delta_k^{(n)}\big)^2 \mathbb{1}_{\{|N_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) \leqslant \theta^{-2}n^{-1}\mathsf{E}\Big(\big(\delta_k^{(n)}\big)^2 \big(N_k^{(n)}\big)^2 \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big)\\
&= \theta^{-2}n^{-1}\mathsf{E}\Bigg(\big(\varepsilon_k^{(n)}-\lambda_n\big)^2\Bigg(\sum_{j=1}^{X_{k-1}^{(n)}}(\xi_{k,j}-m_n)\Bigg)^2 \,\Bigg|\, \mathcal{F}_{k-1}^{(n)}\Bigg) = \theta^{-2}n^{-1}b_n^2\sigma_n^2 X_{k-1}^{(n)}.
\end{aligned}$$

Thus (2.20) follows from (2.10) and the assumptions (ii) and (iv).

We have

$$\mathsf{E}\Big(\big(\delta_k^{(n)}\big)^2 \mathbb{1}_{\{|\delta_k^{(n)}|>\theta\sqrt{n}\}} \,\Big|\, \mathcal{F}_{k-1}^{(n)}\Big) = \mathsf{E}\Big(\big(\varepsilon_1^{(n)}-\lambda_n\big)^2 \mathbb{1}_{\{|\varepsilon_1^{(n)}-\lambda_n|>\theta\sqrt{n}\}}\Big),$$

hence (2.21) follows from the assumption (v). We finished the proof of (2.9), hence the proof of **(A)** is complete.

(B). By the regression equation (1.3) and by the recursion (2.5), we obtain the regression equation

$$X_k^{(n)} - \mathsf{E}X_k^{(n)} = m_n\big(X_{k-1}^{(n)} - \mathsf{E}X_{k-1}^{(n)}\big) + M_k^{(n)}.$$

It has the solution

$$X_k^{(n)} - \mathsf{E}X_k^{(n)} = \sum_{j=1}^{k} m_n^{k-j} M_j^{(n)}.$$

Hence

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

Writing again $m_n = \mathrm{e}^{\alpha_n/n}$ where $\alpha_n \to \alpha$, we have

$$\widetilde{\mathcal{X}}^{(n)}(t) = \int_0^{[nt]/n} \mathrm{e}^{\alpha_n([nt]/n-s)}\,\mathrm{d}\widetilde{\mathcal{M}}^{(n)}(s),$$

which suggests $\widetilde{\mathcal{X}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \int_0^t \mathrm{e}^{\alpha(t-s)}\,\mathrm{d}\widetilde{\mathcal{M}}(s)$. Instead of proving the convergence of stochastic integrals, we choose a simpler way, namely, we rearrange the sum, then we use the continuous mapping theorem and finally we rearrange the result by Itô's formula. Thus we write

$$\begin{aligned}
\widetilde{\mathcal{X}}^{(n)}(t) &= \widetilde{\mathcal{M}}^{(n)}\left(\frac{[nt]}{n}\right) - \sum_{j=1}^{[nt]-1}\left(\mathrm{e}^{\alpha_n([nt]-j-1)/n} - \mathrm{e}^{\alpha_n([nt]-j)/n}\right)\widetilde{\mathcal{M}}^{(n)}\left(\frac{j}{n}\right)\\
&= \widetilde{\mathcal{M}}^{(n)}\left(\frac{[nt]}{n}\right) + \alpha_n\sum_{j=1}^{[nt]-1}\int_{j/n}^{(j+1)/n} \mathrm{e}^{\alpha_n([nt]/n-s)}\,\mathrm{d}s\,\widetilde{\mathcal{M}}^{(n)}\left(\frac{j}{n}\right)\\
&= \widetilde{\mathcal{M}}^{(n)}\left(\frac{[nt]}{n}\right) + \alpha_n\int_0^{[nt]/n} \mathrm{e}^{\alpha_n([nt]/n-s)}\widetilde{\mathcal{M}}^{(n)}(s)\,\mathrm{d}s.
\end{aligned}$$

This implies $\left(\widetilde{\mathcal{X}}^{(n)}, \widetilde{\mathcal{M}}^{(n)}\right) = \Psi_n\big(\widetilde{\mathcal{M}}^{(n)}\big)$ with the mapping $\Psi_n : D(\mathbb{R}_+, \mathbb{R}) \to D(\mathbb{R}_+, \mathbb{R}^2)$ defined by

$$\Psi_n(x)(t) = \left(x\left(\frac{[nt]}{n}\right) + \alpha_n \int_0^{[nt]/n} \mathrm{e}^{\alpha_n([nt]/n - s)} x(s) \,\mathrm{d}s,\, x(t) \right) \qquad \text{for } x \in D(\mathbb{R}_+, \mathbb{R}).$$

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

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

Since almost all trajectories of the limit process are continuous, in view of the continuous mapping theorem, it suffices to check that

$$\Psi_n(x_n) \to \Psi(x) \text{ if } x \in C([0,t], \mathbb{R}),\; x_n \in D([0,t], \mathbb{R}) \text{ with } x_n \to x \text{ as } n \to \infty.$$

(C). Itô's formula yields

$$\widetilde{\mathcal{X}}(t) = \int_0^t \mathrm{e}^{\alpha(t-s)} \,\mathrm{d}\widetilde{\mathcal{M}}(s) = \widetilde{\mathcal{M}}(t) + \alpha \int_0^t \mathrm{e}^{\alpha(t-s)} \widetilde{\mathcal{M}}(s) \,\mathrm{d}s,$$

hence $\Psi(\widetilde{\mathcal{M}}) = (\widetilde{\mathcal{X}}, \widetilde{\mathcal{M}})$, and we finished the proof. □

3 Asymptotics of the least squares estimators

Consider a branching process with immigration given in (1.1). If the immigration mean λ is known then the conditional least squares estimator $\widehat{m}_n$ based on the regression equation (1.3) 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, and it has the form

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

If the immigration mean λ is unknown then the joint conditional least squares estimator $\big(\widetilde{m}_n, \widetilde{\lambda}_n\big)$ of the vector (m, λ) can be obtained by minimizing the sum of squares (3.1) with respect to m and λ, and it has the form

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

where

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

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

$$n^{1/2}(\widehat{m}_n - m) \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}(0, c^2) \qquad \text{as } n \to \infty,$$

$$\begin{pmatrix} n^{1/2}(\widetilde{m}_n - m) \\ n^{1/2}(\widetilde{\lambda}_n - \lambda) \end{pmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}(0, \Sigma) \qquad \text{as } n \to \infty,$$

where the variance c^2 and the covariance matrix Σ can be expressed by the moments up to the third order of the offspring and immigration distribution (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 $\big(\widetilde{m}_n, \widetilde{\lambda}_n\big)$ are not asymptotically normal, but

$$n(\widehat{m}_n - 1) \stackrel{\mathcal{D}}{\longrightarrow} \frac{\frac{1}{2}\mathcal{X}(1)^2 - \big(\lambda + \frac{1}{2}\sigma^2\big)\int_0^1 \mathcal{X}(t)\,\mathrm{d}t}{\int_0^1 \mathcal{X}(t)^2\,\mathrm{d}t} \qquad \text{as } n \to \infty, \tag{3.2}$$

and

$$n(\widetilde{m}_n - 1) \stackrel{\mathcal{D}}{\longrightarrow} \frac{\frac{1}{2}\mathcal{X}(1)^2 - \big(\mathcal{X}(1) + \frac{1}{2}\sigma^2\big)\int_0^1 \mathcal{X}(t)\,\mathrm{d}t}{\int_0^1 \mathcal{X}(t)^2\,\mathrm{d}t - \big(\int_0^1 \mathcal{X}(t)\,\mathrm{d}t\big)^2} \qquad \text{as } n \to \infty. \tag{3.3}$$

(See Wei and Winnicki [16], [18].) The proof is based on the convergence result (1.4). Wei and Winnicki [16], [18] also proved that $\widetilde{\lambda}_n$ is not a consistent estimator of λ.

Now let us consider a sequence of branching processes with immigration given in (1.5). Based on convergence result (1.6) due to Sriram [15], one can easily obtain that (3.2) and (3.3) hold with $\mathcal{X}$ replaced by $\mathcal{X}_\alpha$.

Applying the continuous mapping theorem and using Slutsky's argument one can derive the asymptotic behaviour of the estimators $\widehat{m}_n$ and $\big(\widetilde{m}_n, \widetilde{\lambda}_n\big)$ in the nearly critical model of Theorem 2.4 exactly in the same way as it has been obtained in the case of a Bernoulli offspring distribution in Ispány et al. [7], [8].

3.1 Theorem. *Suppose that the assumptions of Theorem 2.4 hold with some $\lambda > 0$. Then*

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

where $\mu_{\mathcal{X}}(t) = \lambda \int_0^t \mathrm{e}^{\alpha u}\,\mathrm{d}u, \quad t \in \mathbb{R}_+,$ *and*

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

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

Moreover,

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

where $\overline{\mu}_{\mathcal{X}} := \int_0^1 \mu_{\mathcal{X}}(t)\,\mathrm{d}t$ *and*

$$\Sigma := \frac{1}{\left(\int_0^1 (\mu_{\mathcal{X}}(t) - \overline{\mu}_{\mathcal{X}})^2\,\mathrm{d}t\right)^2} (\sigma_{i,j})_{1 \leqslant i,j \leqslant 2}$$

with

$$\begin{aligned}
\sigma_{1,1} &:= \int_0^1 (\mu_{\mathcal{X}}(t) - \overline{\mu}_{\mathcal{X}})^2 \varrho(t)\,\mathrm{d}t, \\
\sigma_{1,2} &:= \int_0^1 (\mu_{\mathcal{X}}(t) - \overline{\mu}_{\mathcal{X}}) \left(\int_0^1 \mu_{\mathcal{X}}(t)^2\,\mathrm{d}t - \overline{\mu}_{\mathcal{X}}\mu_{\mathcal{X}}(t)\right) \varrho(t)\,\mathrm{d}t, \\
\sigma_{2,2} &:= \int_0^1 \left(\int_0^1 \mu_{\mathcal{X}}(t)^2\,\mathrm{d}t - \overline{\mu}_{\mathcal{X}}\mu_{\mathcal{X}}(t)\right)^2 \varrho(t)\,\mathrm{d}t.
\end{aligned}$$

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

3.3 Remark. If $b^2 = \beta = 0$ then the limiting normal distributions are degenerated, that is, $c^2 = 0$ and $\Sigma = 0$. Thus in this case we obtain that $n^{3/2}(\widehat{m}_n - m_n) \stackrel{\mathsf{P}}{\longrightarrow} 0$ and $n^{3/2}(\widetilde{m}_n - m_n) \stackrel{\mathsf{P}}{\longrightarrow} 0$, $n^{1/2}(\widetilde{\lambda}_n - \lambda_n) \stackrel{\mathsf{P}}{\longrightarrow} 0$ as $n \to \infty$, which means that the norming factors $n^{3/2}$ and $n^{1/2}$ are not appropriate.

3.4 Remark. We explain heuristically that Sriram's convergence theorem $\mathcal{X}^{(n)}/n \stackrel{\mathcal{D}}{\longrightarrow} \mu_{\mathcal{X}}$ implies only that $n(\widehat{m}_n - m_n) \stackrel{\mathsf{P}}{\longrightarrow} 0$, while from the convergence $(\mathcal{X}^{(n)} - \mathsf{E}\mathcal{X}^{(n)})/\sqrt{n} \stackrel{\mathcal{D}}{\longrightarrow} \widetilde{\mathcal{X}}$ of Theorem 2.4 we can derive $n^{3/2}(\widehat{m}_n - m_n) \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}(0, c^2)$. On one hand, we have

$$n(\widehat{m}_n - 1) = n\frac{\sum_{k=1}^n X_{k-1}^{(n)}\left(X_k^{(n)} - X_{k-1}^{(n)} - \lambda_n\right)}{\sum_{k=1}^n \left(X_{k-1}^{(n)}\right)^2} = \frac{\int_0^1 \mathcal{X}^{(n)}(t)\,\mathrm{d}\mathcal{X}^{(n)}(t) - n\lambda_n \int_0^1 \mathcal{X}^{(n)}(t)\,\mathrm{d}t}{\int_0^1 \mathcal{X}^{(n)}(t)^2\,\mathrm{d}t}.$$

By Sriram's result we obtain that

$$n(\widehat{m}_n - 1) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mu_{\mathcal{X}}(t)\,\mathrm{d}\mu_{\mathcal{X}}(t) - \lambda_n \int_0^1 \mu_{\mathcal{X}}(t)\,\mathrm{d}t}{\int_0^1 \mu_{\mathcal{X}}(t)^2\,\mathrm{d}t}.$$

By $\mathrm{d}\mu_{\mathcal{X}}(t) = (\lambda + \alpha\mu_{\mathcal{X}}(t))\,\mathrm{d}t$, we conclude $n(\widehat{m}_n - 1) \xrightarrow{\mathcal{D}} \alpha$, thus

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

On the other hand,

$$\begin{aligned} n^{3/2}(\widehat{m}_n - m_n) &= n^{3/2}\frac{\sum_{k=1}^n X_{k-1}^{(n)}\big(X_k^{(n)} - m_n X_{k-1}^{(n)} - \lambda_n\big)}{\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2} = n^{3/2}\frac{\sum_{k=1}^n X_{k-1}^{(n)} M_k^{(n)}}{\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2} \\ &= n^{1/2}\frac{\int_0^1 \mathcal{X}^{(n)}(t)\,\mathrm{d}\mathcal{M}^{(n)}(t)}{\int_0^1 \mathcal{X}^{(n)}(t)^2\,\mathrm{d}t} = \frac{\int_0^1 \Big(n^{-1/2}\widetilde{\mathcal{X}}^{(n)}(t) + n^{-1}\mathsf{E}\mathcal{X}^{(n)}(t)\Big)\,\mathrm{d}\widetilde{\mathcal{M}}^{(n)}(t)}{\int_0^1 \Big(n^{-1/2}\widetilde{\mathcal{X}}^{(n)}(t) + n^{-1}\mathsf{E}\mathcal{X}^{(n)}(t)\Big)^2\,\mathrm{d}t}. \end{aligned}$$

Theorem 2.4 and $n^{-1}\mathsf{E}\mathcal{X}^{(n)}(t) \to \mu_{\mathcal{X}}(t)$ imply

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

as stated. The above consideration shows that the 'main term' of the integrands becomes the nonrandom function $\mu_{\mathcal{X}}$, while the random fluctuation term $\widetilde{\mathcal{X}}$ disappears as $n \to \infty$, and this causes the asymptotic normality of the estimator $\widehat{m}_n$.

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 & Sons, New York.

[3] Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes.* John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore.

[4] Grimvall, A. (1974). On the convergence of sequences of branching processes. *Ann. Probab.* **2**, 1027–1045.

[5] Heyde, C. C. and Seneta, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.* **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. Probab.* **11**, 572–577.

[7] Ispány, M., Pap, G. and Zuijlen, M. v. (2001). Asymptotic inference for nearly unstable INAR(1) models. Report No. 0109, University of Nijmegen, The Netherlands. Submitted to J. Appl. Prob.

[8] Ispány, M., Pap, G. and Zuijlen, M. v. (2001). Asymptotic behaviour of estimators of the parameters of nearly unstable INAR(1) models. Report No. 0120, University of Nijmegen, The Netherlands. To appear in: Foundations of Statistical Inference: Applications in the Medical and Social Sciences, in Industry and the Interface with Computer Sciences. Proceedings, Shoresh, 2000. Contributions to Statistics, Springer–Verlag.

[9] Jacod, J. and Shiryayev, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer–Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo.

[10] Kawazu, K. and Watanabe, S. (1971). Branching processes with immigration and related limit theorems. *Theor. 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. Wharscheinlichkeitstheorie verw. Geb.* **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. Probab.* **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. (1987). A unified estimation theory for the branching process with immigration. Technical Report, University of Maryland.

[17] Wei, C. Z. and Winnicki, J. (1989). Some asymptotic results for the branching process with immigration. *Stoch. Proc. Appl.* **31**, 261–282.

[18] Wei, C. Z. and Winnicki, J. (1990). Estimation of the means in the branching process with immigration. *Ann. Statist.* **18**, 1757–1773.