Probab. Th. Rel. Fields 88, 77–106 (1991)

Probability Theory and Related Fields

Estimation of the variances in the branching process with immigration

J. Winnicki *

Department of Statistics, Columbia University, New York, NY 10027, USA

Received July 10, 1989; in revised form August 2, 1990

Summary. Estimation theory for the variances of the offspring and immigration distributions in a simple branching process with immigration is developed, analogous to the estimation theory for the means given by Wei and Winnicki (1990). Conditional and weighted conditional least squares estimators are considered and their asymptotic properties for the full range of parameters are studied. Nonexistence of consistent estimators in the critical case is established, which complements analogous result of Wei and Winnicki for the supercritical case.

1. Introduction

We will consider the simple branching process with immigration, $\{X_n\}$, defined by

$$X_n = \sum_{j=1}^{X_{n-1}} \xi_{n,j} + \eta_n,$$

where $\{\xi_{n,j}\}$, $n=1, 2, \ldots, j=1, 2, \ldots$, and $\{\eta_n\}$, $n=1, 2, \ldots$, are two independent families of $\mathbf{N}_0$-valued, i.i.d. random variables. The initial value X_0 has an arbitrary distribution on $\mathbf{N}_0$. The distribution of $\xi_{n,j}$ is called the offspring distribution and the distribution of η_n is called the immigration distribution. These are assumed to be nondegenerate. Let $E(\xi_{n,j})=m$, $\operatorname{var}(\xi_{n,j})=\sigma^2$, $E(\eta_n)=\lambda$, $\operatorname{var}(\eta_n)=b^2$. We refer the reader to Athreya and Ney (1972) for basic properties of the process $\{X_n\}$.

Problems of statistical estimation of parameters of the process $\{X_n\}$ were considered by a number of authors. The well-known trichotomy $m<1$, $m=1$, $m>1$ forced earlier authors to restrict attention to one of the cases $m<1$ or $m>1$. Recently, Wei and Winnicki (1990) considered the problem of estimating the means m and λ when the range of m is unknown. The estimators considered by Wei and Winnicki are the conditional least squares and the weighted conditional least squares estimators. They examined the asymptotic properties of

* Research supported by the National Science Foundation under Grant NSF-DMS-8801496

these estimators showing that the weighted conditional least squares estimators are superior to the ordinary conditional least squares estimators.

In this paper we consider an analogous problem for the variances σ^2 and b^2. The problem is interesting in its own right, but it is also an important complement to the work of Wei and Winnicki in that the limiting distributions of the estimators for m and λ depend on σ^2 and b^2. Hence, to use their estimators $\tilde{m}_n$ and $\tilde{\lambda}_n$ in practice would require estimating the variances σ^2 and b^2.

The problem of estimating the variance σ^2 has been first considered by Heyde (1974). Under the assumption that the process is supercritical ($m>1$), he proved that the estimator $\frac{1}{n}\sum_{k=1}^{n}[(X_{k+1}-\bar{m}_n X_k)^2/X_k]$, where $\bar{m}_n = X_{n+1}/X_n$, is strongly consistent and asymptotically normal.

Yanev and Tchoukova-Dantcheva (1980) considered the problem of estimating σ^2 and b^2 in the subcritical case ($m<1$). They proposed the estimators

$$\hat{\sigma}_n^2 = \frac{\sum_{k=1}^{n} U_k^2 (X_k - \bar{X}_n)}{\sum_{k=1}^{n} (X_k - \bar{X}_n)^2}, \tag{1.1}$$

$$\hat{b}_n^2 = \frac{\sum_{k=1}^{n} U_k^2 \sum_{k=1}^{n} (X_k^2 - \bar{X}_n X_k)}{n \sum_{k=1}^{n} (X_k - \bar{X}_n)^2}, \tag{1.2}$$

where $\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$, $U_k = X_k - mX_{k-1} - \lambda$. If m and λ are not known, they proposed to use $\hat{U}_k = X_k - \hat{m}_n X_{k-1} - \hat{\lambda}_n$ instead of U_k in (1.1) and (1.2), where $\hat{m}_n$ and $\hat{\lambda} n$ are the estimators of the offspring and immigration distribution means given by Heyde and Seneta (1972). Under certain moment assumptions, Yanev and Tchoukova-Dantcheva stated the asymptotic properties of the estimators $\hat{\sigma}_n^2$ and $\hat{b}_n^2$, including consistency and asymptotic normality.

Notice that the estimators $\hat{\sigma}_n^2$ and $\hat{b}_n^2$ are the conditional least squares estimators [in the sense of Klimko and Nelson (1978) or Wei and Winnicki (1990)]. To see it, we first suppose that m and λ are known. Then write

$$U_k^2 = \sigma^2 X_{k-1} + b^2 + V_k \tag{1.3}$$

and treat the above equation as a stochastic regression equation with the unknown coefficients σ^2 and b^2 and a martingale difference "error" term V_k. The least squares estimators based on Eq. (1.3) are given by (1.1) and (1.2).

Now observe that in the critical and supercritical cases the "error" term V_k in the stochastic regression equation (1.3) is strongly heteroscedastic, as can be seen by computing $\operatorname{var}(V_k) = E(R(X_{k-1})) = O(E(X_{k-1}^2))$, where

$$R(X) = 2\sigma^4 X^2 + (a^4 + 4\sigma^2 b^2 - 3\sigma^4) X + c^4 - b^4,$$
$$a^4 = E((\xi_{n,i} - m)^4), \qquad c^4 = E((\eta_n - \lambda)^4).$$

This suggests that a weighted conditional least squares approach [Nelson (1980); Wei and Winnicki (1990)] may be useful.

We then consider the transformed equation

$$\frac{U_k^2}{X_{k+1}+1}=\sigma^2+(b^2-\sigma^2)\frac{1}{X_{k-1}+1}+\frac{V_k}{X_{k-1}+1}. \tag{1.4}$$

In the stochastic regression equation (1.4) the "error" term $V_k/(X_{k-1}+1)$ satisfies

$$E\left(\frac{V_k}{X_{k-1}+1}\,\middle|\,\mathscr{F}_{k-1}\right)=0$$

and assuming $E(\xi_{n,i}^4)<\infty$ and $E(\eta_n^4)<\infty$,

$$E\left(\left(\frac{V_k}{X_{k-1}+1}\right)^2\,\middle|\,\mathscr{F}_{k-1}\right)=\frac{R(X_{k-1})}{(X_{k-1}+1)^2}<a^4-\sigma^4+c^4-b^4+4\sigma^2b^2<\infty.$$

Furthermore, if $X_k\to\infty$,

$$E\left(\left(\frac{V_k}{X_{k-1}+1}\right)^2\,\middle|\,\mathscr{F}_{k-1}\right)\to 2\sigma^4.$$

Hence, the variances of the "error" terms would approximately be homogeneous. The weighted conditional least squares estimators resulting from Eq. (1.4) are

$$\tilde{\sigma}_n^2=\frac{\sum_{k=1}^n\frac{1}{(X_{k-1}+1)^2}\sum_{k=1}^n\frac{U_k^2}{X_{k-1}+1}-\sum_{k=1}^n\frac{1}{X_{k-1}+1}\sum_{k=1}^n\frac{U_k^2}{(X_{k-1}+1)^2}}{n\sum_{k=1}^n\frac{1}{(X_{k-1}+1)^2}-\left(\sum_{k=1}^n\frac{1}{X_{k-1}+1}\right)^2}, \tag{1.5}$$

$$\tilde{b}_n^2=\frac{\sum_{k=1}^n\frac{X_{k-1}}{X_{k-1}+1}\sum\frac{U_k^2}{(X_{k-1}+1)^2}-\sum_{k=1}^n\frac{X_{k-1}}{(X_{k-1}+1)^2}\sum_{k=1}^n\frac{U_k^2}{X_{k-1}+1}}{n\sum_{k=1}^n\frac{1}{(X_{k-1}+1)^2}-\left(\sum_{k=1}^n\frac{1}{X_{k-1}+1}\right)^2}. \tag{1.6}$$

Here $\tilde{b}_n^2=\tilde{\sigma}_n^2+\tilde{q}_n$, where $\tilde{q}_n$ is the least squares estimator of $q=\sigma^2-b^2$.

Estimators $\tilde{\sigma}_n^2$ and $\tilde{b}_n^2$ can be used if m and λ are known. Otherwise we would use $\tilde{\tilde{\sigma}}_n^2$ and $\tilde{\tilde{b}}_n^2$ which are obtained from $\tilde{\sigma}_n^2$ and $\tilde{b}_n^2$ by replacing U_k^2 with $\tilde{U}_k^2=X_k-\tilde{m}_nX_{k-1}-\tilde{\lambda}_n$ in (1.5) and (1.6).

The first part of this paper is devoted to examining asymptotic properties of the above estimators. Our methods are extensions of those of Wei and Winnicki (1990). In Sect. 2 a number of preliminary asymptotic results is given, mostly concerning the critical case. In particular, using the theory of weak convergence in function spaces, we prove some nonstandard limit theorems for martingale functionals of the branching process. In Sect. 3 we give limit theorems for the statistics of the branching process. We show that the conditional least squares estimator $\hat{\sigma}_n^2$ is consistent only if $m\leq 1$, while $\hat{b}_n^2$ is consistent only if $m<1$. On the other hand, $\tilde{\sigma}_n^2$ is consistent for all m and $\tilde{b}_n^2$ is consistent only if $m<1$ or $m=1$ and $2\lambda \leq \sigma^2$. Rates of convergence in the form of limiting distribution results are also considered.

The second part of the paper is devoted to the question of existence of consistent estimators for b^2 in the case $m=1$, $2\lambda > \sigma^2$. It is known that there is no consistent estimator for b^2 if $m>1$. In fact, Wei and Winnicki (1990) showed that in the case of the supercritical branching process with immigration the only parameters that have consistent estimators are the mean m and the variance σ^2 of the offspring distribution. In view of the exponential growth of X_n and the stationary rate of immigration, this result may be not surprising. In particular, no parameters of the immigration distribution can be estimated consistently in the supercritical case. The situation is more complicated in the critical case, where we know that λ has a consistent estimator. In Theorem 4.5 we prove a general result that the only parameters of the critical, transient branching process with immigration which may have consistent estimators are the first four moments of the offspring distribution and the mean of the immigration distribution. It is worth pointing out that in the traditional statistical setup of independent, identically distributed observations the issues of estimability of parameters do not arise if only all the parameters are identifiable. In statistical inference for stochastic processes these issues are crucial and the example of the branching process with immigration discussed in this paper shows that a careful analysis may be required.

2. Preliminary results

Our analysis of asymptotic behavior of the statistics of the branching process with immigration will be facilitated by establishing asymptotic properties of several key functionals of the process and then showing that our statistics can be expressed in terms of these functionals.

In this section we will carry out the first part of the above plan. We will concentrate on the critical case, which turns out to be the most complicated one.

We will need the following extension of a result of Strasser (1986). Using Strasser's notation, we consider a sequence of filtrations $(\mathscr{F}_{nk}, k \geq 0\}$, $n=1, 2, \ldots$, and double sequences $\{X_{nk}, k \geq 0\}$ and $\{Y_{nk}, k \geq 0\}$ of $\mathscr{F}_{nk}$-adapted, integrable random variables. Let $\tau_n(t)$, $0 \leq t < \infty$, be stopping times for the filtration $\{\mathscr{F}_{nk}\}$, $n=1, 2, \ldots$, such that τ_n is a.s. right continuous, $\mathbf{N}_0$-valued, nondecreasing and taking all values between zero and $\tau_n(t)$ for all $t>0$. We will assume that the random variables $\{X_{nk}\}$ and $\{Y_{nk}\}$ satisfy the condition

$$\sum_{k=1}^{\tau_n(t)} E(|Z_{nk}| I[|Z_{nk}| > \varepsilon] \,|\, \mathscr{F}_{n,k-1}) \xrightarrow{P} 0 \tag{2.1}$$

for all $\varepsilon > 0$, $t > 0$, with $\{Z_{nk}\} = \{X_{nk}\}$ as well as $\{Z_{nk}\} = \{Y_{nk}\}$. Define

$$S_{nk} = \sum_{j=0}^{k} X_{nj}, \qquad S_n(t) = S_{n,\tau_n(t)},$$

$$T_{nk} = \sum_{j=0}^{k} Y_{nj}, \qquad T_n(t) = T_{n,\tau_n(t)}.$$

Proposition 2.1. *Assume that* $\{Y_{nk}, k \geqq 0\}$ *is a martingale difference sequence with respect to* $\{\mathscr{F}_{nk}\}$ *for each* $n=1, 2, \ldots$. *Assume further that the common distributions of* $(S_n(t), T_n(t))$ *converge to the common distribution of a pair of continuous processes* $(S(t), T(t))$, *where* $T(t)$ *is a local martingale. Then for any Hölder continuous function* $\varphi\colon \mathbf{R} \to \mathbf{R}$ *with exponent* $\alpha > 0$, *the distributions of the processes*

$$\sum_{k=1}^{\tau_n(t)} \varphi(S_{n,k-1})\, Y_{nk}$$

converge to the distribution of the process

$$\int_0^t \varphi(S)\, dT.$$

The proof of Proposition 2.1 is similar to the proof of Theorem (1.7) in Strasser (1986).

As a corollary to Proposition 2.1 we can prove Proposition 2.2 below, which generalizes Theorem (1.12) of Strasser (1986).

Consider a third $\mathscr{F}_{nk}$-adapted double sequence $\{Z_{nk}, k \geqq 0\}$ satisfying condition (2.1) and let $I_{nk} = \sum_{j=0}^{k} Z_{nk}$, $I_n(t) = I_{n, \tau_n(t)}$.

Proposition 2.2. *Assume that* $\{Y_{nk}, k \geqq 1\}$ *is a martingale difference sequence with respect to* $\{\mathscr{F}_{nk}\}$ *for each* $n=1, 2, \ldots$ *and that the processes* $S_n(t)$ *are stochastically uniformly bounded (i.e.* $\lim_{a\to\infty} \limsup_{n\to\infty} P(\sup_t |S_n(t)| > a) = 0$). *Assume further that the common distributions of* $(I_n(t), T_n(t))$ *converge to the distribution of a pair of continuous processes* $(I(t), T(t))$, *where* $T(t)$ *is a local martingale and* $I(t) = \int_0^t \psi(s)\, dA(s)$ *for a T-adapted continuous process of bounded variation* $A(t)$ *and a continuous function* $\psi(t)$. *If for a Hölder continuous function* $\varphi\colon \mathbf{R} \to \mathbf{R}$ *with exponent* $\alpha \geqq \frac{1}{2}$

$$S_n(t) = I_n(t) + \sum_{k=1}^{\tau_n(t)} \varphi(S_{n,k-1})\, Y_{nk}, \qquad n = 1, 2, \ldots,$$

then the distributions of $S_n(t)$ *converge to the distribution of the process* $S(t)$ *satisfying the equation*

$$S(t) = \int_0^t \psi(s)\, dA(s) + \int_0^t \varphi(S)\, dT. \tag{2.2}$$

For the branching process with immigrating, $\{X_n\}$, consider $Y_n(t) = X_{[nt]}/n$, $t \geqq 0$. Then $Y_n \in D^+[0, \infty)$, where $D^+[0, \infty)$ is the space of nonnegative, right-continuous functions having left limits. Wei and Winnicki (1989) proved a weak convergence result for $Y_n(t)$. In Proposition 2.3 below we formulate an equivalent form of that result. Its proof is an example of an application of Proposition 2.2 and is omitted.

Proposition 2.3. *If $m=1$, $\sigma^2<\infty$ and $b^2<\infty$, then the distributions of Y_n converge weakly to the distribution of the solution of the equation*

$$dY(t)=\lambda\,dt+\sigma\sqrt{Y(t)}\,dW(t),$$
$$Y(0)=0,$$

where $W(t)$ is a standard Brownian motion.

Let $V_k=U_k^2-\sigma^2 X_{k-1}-b^2$ and define

$$T_n(t)=\sum_{k=1}^{[nt]}\frac{U_k}{\sigma\sqrt{n}\sqrt{X_{k-1}+1}},$$
$$V_n(t)=\sum_{k=1}^{[nt]}\frac{V_k}{\sigma^2\sqrt{2n(X_{k-1}+1)}}.$$

We will need the following lemma.

Lemma 2.4. *If $m=1$, $E(\xi_{n,i}^4)<\infty$ and $E(\eta_n^4)<\infty$, then*

$$(T_n(t),\ V_n(t))\to(W(t),\ B(t))\quad(\text{weakly in } D^2[0,\infty)),$$

where W and B are independent Brownian motions.

Proof. We will first show that

$$T_n(t)\to W(t)\quad(\text{weakly in } D[0,\infty)). \tag{2.3}$$

Let $Y_{nk}=U_k/(\sigma\sqrt{n}\sqrt{X_{k-1}+1})$. We have

$$\sum_{k=1}^{[nt]}E(Y_{nk}^2\,|\,\mathscr{F}_{k-1})=\sum_{k=1}^{[nt]}\frac{\sigma^2 X_{k-1}+b^2}{\sigma^2(X_{k-1}+1)\,n}\xrightarrow{P}t, \tag{2.4}$$

where we used the easily verified fact that

$$\frac{1}{n}\sum_{k=0}^{n-1}(X_k+1)^{-\alpha}\xrightarrow{P}0, \tag{2.5}$$

for any $\alpha>0$. Furthermore, we claim that

$$\sum_{k=1}^{[nt]}E[Y_{nk}^2 I[|Y_{nk}|>\varepsilon]\,|\,\mathscr{F}_{k-1}] \tag{2.6}$$
$$\le\frac{1}{n}\sum_{k=1}^{[nt]}E\Big[U_k^2/(\sigma^2(X_{k-1}+1))\,I[|U_k|/\sqrt{X_{k-1}+1}>\varepsilon k]\,|\,\mathscr{F}_{k-1}\Big]\xrightarrow{P}0.$$

To prove (2.6) we observe that

$$U_k^2/\sigma^2(X_{k-1}+1)\xrightarrow{d}[N(0,1)]^2,$$

where $N(0, 1)$ denotes a unit normal random variable [Heyde and Seneta (1971)] and

$$E[U_k^2/(\sigma^2(X_{k-1}+1))] = E[(\sigma^2 X_{k-1}+b^2)/(\sigma^2(X_{k-1}+1))] \to 1$$

(since $X_k \xrightarrow{P} \infty$).

It follows that $U_k^2/(\sigma^2(X_{k-1}+1))$ is uniformly integrable and so

$$E\left[U_k^2/(\sigma^2(X_{k-1}+1))\, I[|U_k|/(\sigma\sqrt{X_{k-1}+1})>\varepsilon\sqrt{k}]\right] \to 0.$$

This proves (2.5) and by Theorem 3.2 of Helland (1982), (2.3) follows. The same method of proof can be used to show that

$$V_n(t) \to B(t) \quad (\text{weakly in } D[0, \infty)).$$

To apply Theorem 3.3 of Helland (1982) it remains to show that

$$(2.7)\quad \sum_{k=1}^{[nt]} E\left(\frac{U_k}{\sigma\sqrt{n}\sqrt{(X_{k-1}+1)}} \cdot \frac{V_k}{\sigma^2\sqrt{2n}(X_{k-1}+1)} \,\middle|\, \mathscr{F}_{k-1}\right) \xrightarrow{P} 0 \quad \text{for all } t>0.$$

Now,

$$E\left(\frac{U_k}{\sigma\sqrt{n}\sqrt{(X_{k-1}+1)}} \cdot \frac{V_k}{\sigma^2\sqrt{2n}(X_{k-1}+1)} \,\middle|\, \mathscr{F}_{k-1}\right) = \frac{E[(X_k - X_{k-1}-\lambda)^3 \mid \mathscr{F}_{k-1}]}{n\sigma^3(X_{k-1}+1)^{3/2}}$$
$$= \frac{1}{n}\, \frac{X_{k-1}E(\xi_{n,i}-m)^3 + E(\eta_n-\lambda)^3}{\sigma^3(X_{k-1}+1)^{3/2}},$$

so (2.7) follows by (2.5).

Remark 2.5. Notice that Lemma 2.4 remains valid in the supercritical case ($m>1$). The same proof works except that in order to obtain convergence of $\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}+1)^{-\alpha}$ for $\alpha>0$ we use the fact that

$$(2.8)\qquad X_n/m^n \to Z \quad \text{a.s. where } 0<Z<\infty,$$

[Seneta (1970)].

Remark 2.6. It follows from Proposition 2.3 and Lemma 2.4 that Y and B are independent processes and so

$$(2.9)\qquad (Y_n(t), V_n(t)) \to (Y(t), B(t)) \quad (\text{weakly in } D^2[0, \infty)).$$

Lemma 2.7. *If $m=1$, $\sigma^2<\infty$ and $b^2<\infty$, then*

$$(2.10)\qquad \frac{1}{n}\sum_{k=1}^{n} U_k \xrightarrow{d} Y(1)-\lambda,$$

and

$$\frac{1}{n^2}\sum_{k=1}^{n} X_{k-1}U_k \xrightarrow{d} \sigma\int_0^1 Y^{3/2}(t)\,dW(t). \tag{2.11}$$

Proof. Clearly, (2.10) holds by Proposition 2.2. Moreover, by Proposition 2.3, $(Y_n(t), T_n(t)) \to (Y(t), T(t))$ weakly in $D^2[0, \infty)$. By Proposition 2.1,

$$\frac{1}{n^2}\sum_{k=1}^{n} X_{k-1}U_k = \sigma\sum_{k=1}^{n} [S_{n,k-1}]^{3/2}\, Y_{nk} - \frac{1}{n^2}\sum_{k=1}^{n} U_k \xrightarrow{d} \sigma\int_0^1 Y^{3/2}(t)\,dW(t).$$

Similarly, with the aid of Proposition 2.1 and Remark 2.6 we obtain the following lemma.

Lemma 2.8. *If* $m=1$, $E(\xi_{n,i}^4)<\infty$ *and* $E(\eta_n^4)<\infty$, *then*

$$\frac{1}{\sqrt{2\sigma^2}\,n^{3/2}}\sum_{k=1}^{n} V_k \xrightarrow{d} \int_0^1 Y(t)\,dB(t) \tag{2.12}$$

and

$$\frac{1}{\sqrt{2\sigma^2}\,n^{5/2}}\sum_{k=1}^{n} X_{k-1}V_k \xrightarrow{d} \int_0^1 Y^2(t)\,dB(t). \tag{2.13}$$

Remark 2.9. The distribution of the limit in (2.12) is a mixture of a unit normal random variable and $\left(\int_0^1 Y^2(t)\,dt\right)^{1/2}$, while the limit in (2.13) is a mixture of a unit normal and $\left(\int_0^1 Y^4(t)\,dt\right)^{1/2}$.

In the study of weighted conditional least squares estimators we will also need to consider $\sum_{k=1}^{n}\frac{U_k}{X_{k-1}+1}$, $\sum_{k=1}^{n}\frac{U_k}{(X_{k-1}+1)^2}$, $\sum_{k=1}^{n}\frac{V_k}{X_{k-1}+1}$ and $\sum_{k=1}^{n}\frac{V_k}{(X_{k-1}+1)^2}$.

Lemma 2.10. *Assume that* $m=1$, $\sigma^2<\infty$ *and* $b^2<\infty$. *Then*

(a)

$$\sum_{k=1}^{n}\frac{U_k}{X_{k-1}+1} = o\left(\left(\sum_{k=1}^{n}\frac{1}{X_{k-1}+1}\right)^{\alpha}\right) \quad \text{a.s. for any } \alpha>\frac{1}{2}. \tag{2.14}$$

(b) *If* $\tau \leq 1$, *then*

$$\sum_{k=1}^{n}\frac{U_k}{(X_{k-1}+1)^2} = o\left(\left(\sum_{k=1}^{n}\frac{1}{(X_{k-1}+1)^3}\right)^{\alpha}\right) \quad \text{a.s. for any } \alpha>\frac{1}{2}. \tag{2.15}$$

(c) *If* $\tau > 1$, *then*

(2.16) $$\sum_{k=1}^{n} \frac{U_k}{(X_{k-1}+1)^2} \quad \text{converges almost surely to a finite limit.}$$

Proof. Relations (2.14) and (2.15) follow from the strong law for martingales [Theorem 2.18 in Hall and Heyde (1980)]. We will only prove (2.14), since the proof of (2.15) is analogous.

Note that $U_k/(X_{k-1}+1)$ is a martingale difference with respect to $\{\mathscr{F}_k\}$ and

$$E\left(\left(\frac{U_k}{X_{k-1}-1}\right)^2 \middle| \mathscr{F}_{k-1}\right) = \frac{\sigma^2 X_{k-1}+b^2}{(X_{k-1}+1)^2} = O\left(\frac{1}{X_{k-1}+1}\right) \quad \text{a.s.}$$

Also notice that

(2.17) $$\sum_{k=1}^{n} \frac{1}{X_{k-1}+1} \to \infty \quad \text{a.s.}$$

[Wei and Winnicki (1989)]. Since $\alpha > \frac{1}{2}$,

$$\sum_{k=1}^{\infty} \left[\sum_{j=1}^{k} \frac{1}{X_{j-1}+1}\right]^{-2\alpha} E\left(\left(\frac{U_k}{X_{k-1}+1}\right)^2 \middle| \mathscr{F}_{k-1}\right) < \infty \quad \text{a.s.}$$

This completes the proof of (2.14).

To prove (2.16) we will apply the local martingale convergence theorem [Theorem 2.17 in Hall and Heyde (1980)] to the martingale $\sum_{k=1}^{n} U_k/(X_{k-1}+1)^2$.
It suffices to show that

$$\sum_{k=1}^{\infty} E\left(\left(\frac{U_k}{(X_{k-1}+1)^2}\right)^2 \middle| \mathscr{F}_{k-1}\right) = \sum_{k=1}^{\infty} \frac{\sigma^2 X_{k-1}+b^2}{(X_{k-1}+1)^4} < \infty.$$

But this is an immediate consequence of Lemma 2.13 of Wei and Winnicki (1989).

Turning to $\sum_{k=1}^{n} V_k/(X_{k-1}+1)$ and $\sum_{k=1}^{n} V_k/(X_{k-1}+1)^2$, it is clear from Lemma 2.4 that

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{V_k}{X_{k-1}+1} \xrightarrow{d} N(0, 2\sigma^4).$$

Let $\tau = 2\lambda/\sigma^2$. The following lemma gives the order of magnitude of $\sum_{k=1}^{n} \frac{V_n}{(X_{k-1}+1)^2}$.

Lemma 2.11. *Assume that* $m=1$, $E(\xi_{n,i}^4) < \infty$ *and* $E(\eta_n^4) < \infty$.

(a) *If* $\tau < 1$, *then*

$$\sum_{k=1}^{n} \frac{V_k}{(X_{k-1}+1)^2} = o(\sqrt{n}) \quad a.s. \tag{2.18}$$

(b) *If* $\tau \leqq 1$, *then*

$$\sum_{k=1}^{n} \frac{V_k}{(X_{k-1}+1)^2} = o\left(\sum_{k=1}^{n} \frac{1}{(X_{k-1}+1)^2}\right) \quad a.s. \tag{2.19}$$

(c) *If* $\tau > 1$, *then*

$$\sum_{k=1}^{n} \frac{V_k}{(X_{k-1}+1)^2} \quad \text{converges almost surely to a finite limit.} \tag{2.20}$$

Proof. The proofs of (2.19) and (2.20) are analogous to the proofs of (2.15) and (2.16). Relation (2.18) will be established by another application of the strong law for martingales. It is enough to show that

$$\sum_{k=1}^{\infty} \frac{1}{k} E\left(\left(\frac{V_k}{(X_{k-1}+1)^2}\right)^2 \Bigg| \mathscr{F}_{k-1}\right) = \sum_{k=1}^{\infty} \frac{1}{k} \frac{R(X_{k-1})}{(X_{k-1}+1)^4} < \infty \quad \text{a.s.,}$$

which is implied by

$$\sum_{k=1}^{\infty} \frac{1}{k} E \frac{1}{X_{k-1}+1} < \infty. \tag{2.21}$$

In the case $X_0 = 0$, (2.21) follows by a result of Pakes (1975), cf. Theorem 10, who showed that

$$\sum_{j=1}^{\infty} \frac{1}{j} P(X_n = j \mid X_0 = 0) = O(n^{-\tau})$$

and

$$P(X_n = 0 \mid X_0 = 0) = O(n^{-\tau}).$$

The general case follows by noticing that

$$E\left(\frac{1}{X_{k-1}+1}\right) \leqq E\left(\frac{1}{X_{k-1}+1} \Bigg| X_0 = 0\right), \quad k = 1, 2, \ldots.$$

The next lemma summarizes the asymptotic properties of $\sum_{k=1}^{n} \frac{1}{X_{k-1}+1}$ needed in the context of the present paper. Various parts of this lemma are proved in Wei and Winnicki (1989) (see also Pakes (1975), Theorem 10).

Lemma 2.12. *Assume that* $m = 1$, $b^2 < \infty$ *and* $E(\xi_{n,i}^2 \log^+ \xi_{n,i}) < \infty$.

(a) *If* $\tau \geqq 1$, *then*

$$\sum_{k=1}^{n} \frac{1}{X_{k-1}+1} = o(n^{\alpha}) \quad a.s. \text{ for any } \alpha > 0. \tag{2.22}$$

(b) *If* $\tau < 1$, *then*

$$\sum_{k=1}^{n} \frac{1}{X_{k-1}+1} = O_P(n^{1-\tau}) \tag{2.23}$$

and

$$\sum_{k=1}^{n} \frac{1}{X_{k-1}+1} = O\left(\sum_{k=1}^{n} \frac{1}{(X_{k-1}+1)^2}\right) \quad a.s. \tag{2.24}$$

In developing the estimation theory for the variances σ^2 and b^2 without assuming that m and λ are known, we will need to estimate them. We will use the conditional weighted least squares estimators $\tilde{m}_n$ and $\tilde{\lambda}_n$. Consequently, to prove asymptotic properties of the estimators for σ^2 and b^2 we will need the rates of convergence of $\tilde{m}_n$ and $\tilde{\lambda}_n$.

Lemma 2.13. *If* $m=1$, $\sigma^2 < \infty$ *and* $b^2 < \infty$, *then*

(a)

$$\tilde{m}_n - m = O_P(n^{-1}) \tag{2.25}$$

(b)

$$\tilde{\lambda}_n - \lambda = o_P\left(\left(\sum_{k=1}^{n} \frac{1}{X_{k-1}+1}\right)^{-\alpha}\right) \quad \text{for any } \alpha < \frac{1}{2}. \tag{2.26}$$

Proof. Relation (2.25) is a consequence of the asymptotic distribution result for $\tilde{m}_n$ of Wei and Winnicki (1990) (see also Wei and Winnicki (1989), Corollary 2.3).

To prove (2.26) we notice that in the proof of Theorem 2.5 of Wei and Winnicki (1990) it is essentially shown that

$$\tilde{\lambda}_n - \lambda = O_P\left(\sum_{k=1}^{n} \frac{U_k}{X_{k-1}+1} \Big/ \sum_{k=1}^{n} \frac{1}{X_{k-1}+1}\right) + O_P\left(\left(\sum_{k=1}^{n} \frac{1}{X_{k-1}+1}\right)^{-1}\right) \tag{2.27}$$

and so (2.26) is a consequence of (2.17) and Lemma 2.10.

3. Limit theorems for estimators of the variances in the branching process

We will first consider the subcritical case. Under the assumption that $\sigma^2 < \infty$ and $b^2 < \infty$, by Remark 2.9 of Wei and Winnicki (1989), we may assume that $\{X_n\}$ is stationary. We denote by X a random variable with the stationary distribution of the process.

Proposition 3.1. *If* $m<1$, $\sigma^2<\infty$ *and* $b^2<\infty$ *then* $\hat{\sigma}_n^2 \to \sigma^2$ *and* $\hat{b}_n^2 \to b^2$ *a.s. If, in addition,* $E(\xi_{n,i}^4)<\infty$ *and* $E(\eta_n^4)<\infty$, *then*

$$\sqrt{n}(\hat{\sigma}_n^2-\sigma^2, \hat{b}_n^2-b^2)' \xrightarrow{d} N(0, \Phi^{-1}\Sigma\Phi'^{-1}),$$

where

$$\Sigma=\begin{pmatrix} E(R(X)X^2) & E(R(X)X) \\ E(R(X)X) & E(R(X)) \end{pmatrix}$$

and

$$\Phi=\begin{pmatrix} E(X^2) & E(X) \\ E(X) & 1 \end{pmatrix}.$$

Proof. The strong consistency result is most easily established by an application of the ergodic theorem. We have

$$\frac{1}{n}\sum_{k=1}^{n} X_{k-1}^{\alpha} U_k^2 \to E[X^{\alpha}(\sigma^2 X+b^2)] \quad \text{a.s.} \tag{3.1}$$

and

$$\frac{1}{n}\sum_{k=1}^{n} X_{k-1}^{\alpha} \to E(X^{\alpha}) \quad \text{a.s.} \tag{3.2}$$

($\alpha=0, 1, 2$), which yields $\hat{\sigma}_n^2 \to \sigma^2$ a.s. and $\hat{b}_n^2 \to b^2$ a.s.

The asymptotic distribution of $\tilde{\sigma}_n^2$ and $\tilde{b}_n^2$ is established by an application of the martingale central limit theorem. We can write

$$(\hat{\sigma}_n^2-\sigma^2, \hat{b}_n^2-b^2)' = \begin{pmatrix} \sum_{k=1}^{n} X_{k-1}^2 & \sum_{k=1}^{n} X_{k-1} \\ \sum_{k=1}^{n} X_{k-1} & n \end{pmatrix}^{-1} \begin{pmatrix} \sum_{k-1}^{n} X_{k-1} V_k \\ \sum_{k=1}^{n} V_k \end{pmatrix}. \tag{3.3}$$

Proceeding as in the proof of Theorem 3.1 in Wei and Winnicki (1990), we obtain

$$\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n} X_{k-1} V_k, \sum_{k=1}^{n} V_k\right)' \xrightarrow{d} N(0, \Sigma).$$

This together with (3.2) and (3.3) completes the proof.

Remark 3.2. Using the methods of Quine (1976) we can express the moments of the stationary distribution in terms of the moments of the offspring and immigration distributions. Set

$$\mu=\frac{\lambda}{1-m}$$

$$\nu=\frac{1}{1-m^2}[b^2+\mu\sigma^2]$$

$$\gamma=\frac{1}{1-m^3}[E(\eta_n-\lambda)^3+\mu E(\xi_{n,i}-m)^3+3m\sigma^2\nu]$$

$$\delta=\frac{1}{1-m^4}\{c^4-3b^4+\mu[a^4-3\sigma^4]+4m\nu E(\xi n, i-m)^3+6m^2\sigma^2\gamma+3\sigma^2\nu\}.$$

Then

$$\begin{aligned} EX &= \mu \\ EX^2 &= \gamma + \mu^2 \\ EX^3 &= \gamma + 3\mu\nu + \mu^3 \\ EX^4 &= \delta + 3\nu^2 + 4\mu\gamma + 6\mu^2\nu + \mu^4. \end{aligned}$$

Without assuming the knowledge of m and λ we consider the estimators

$$\hat{\hat{\sigma}}_n^2 = \frac{n \sum_{k=1}^{n} X_{k-1} \hat{U}_k^2 - \sum_{k=1}^{n} X_{k-1} \sum_{k=1}^{n} \hat{U}_k^2}{n \sum_{k=1}^{n} X_{k-1}^2 - \left(\sum_{k=1}^{n} X_{k-1} \right)^2} \tag{3.4}$$

and

$$\hat{\hat{b}}_n^2 = \frac{\sum_{k=1}^{n} X_{k-1}^2 \sum_{k=1}^{n} \hat{U}_k^2 - \sum_{k=1}^{n} X_{k-1} \sum_{k=1}^{n} X_{k-1} \hat{U}_k^2}{n \sum_{k=1}^{n} X_{k-1}^2 - \left(\sum_{k=1}^{n} X_{k-1} \right)^2}, \tag{3.5}$$

where $\hat{U}_k = X_k - \hat{m}_n X_{k-1} - \hat{\lambda}_n$ and $\hat{m}_n$ and $\hat{\lambda}_n$ are consistent estimators for m and λ [see Wei and Winnicki (1990)]. If $\hat{m}_n$ and $\hat{\lambda}_n$ are the estimators of Heyde and Seneta (1972), then (3.4) and (3.5) are the estimators considered by Yanev and Tchoukova-Dantcheva (1980). In the subcritical case the asymptotic results are the same for both interpretations of $\hat{m}_n$ and $\hat{\lambda}_n$.

Theorem 3.3. *If* $m < 1$, $\sigma^2 < \infty$ *and* $b^2 < \infty$, *then* $\hat{\hat{\sigma}}_n^2 \to \sigma^2$ *a.s., and* $\hat{\hat{b}}_n^2 \to b^2$ *a.s. If, in addition,* $E(\xi_{n,i}^4) < \infty$ *and* $E(\eta_n^4) < \infty$, *then*

$$\sqrt{n}(\hat{\hat{\sigma}}_n^2 - \sigma^2, \hat{\hat{b}}_n^2 - b^2)' \xrightarrow{d} N(0, \Phi^{-1} \Sigma \Phi'^{-1}).$$

Proof. We have

$$(\hat{\hat{\sigma}}_n^2 - \hat{\sigma}_n^2, \hat{\hat{b}}_n^2 - \hat{b}_n^2)' = \begin{pmatrix} \sum_{k=1}^{n} X_{k-1}^2 & \sum_{k=1}^{n} X_{k-1} \\ \sum_{k=1}^{n} X_{k-1} & n \end{pmatrix}^{-1} \begin{pmatrix} \sum_{k=1}^{n} X_{k-1} [\hat{U}_k^2 - U_k^2] \\ \sum_{k=1}^{n} [\hat{U}_k^2 - U_k^2] \end{pmatrix}.$$

It is enough to show that

$$\frac{1}{n} \sum_{k=1}^{n} X_{k-1}^{\alpha} [\hat{U}_k^2 - U_k^2] \to 0 \qquad \text{a.s.} \tag{3.6}$$

and

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_{k-1}^{\alpha} [\hat{U}_k^2 - U_k^2] \xrightarrow{P} 0 \tag{3.7}$$

for $\alpha=0$ and $\alpha=1$.

But

$$\begin{aligned}\sum_{k=1}^{n} X_{k-1}^{\alpha} [\hat{U}_k^2 - U_k^2] &= 2(\hat{m}_n - m) \sum_{k=1}^{n} X_{k-1}^{\alpha+1} U_k + 2(\hat{\lambda}_n - \lambda) \sum_{k=1}^{n} X_{k-1}^{\alpha} U_k \\ &+ (\hat{m}_n - m)[(\hat{m}_n - m) + (\hat{\lambda}_n - \lambda)] \sum_{k=1}^{n} X_{k-1}^{\alpha+1} \\ &+ (\hat{\lambda}_n - \lambda)[(\hat{m}_n - m) + (\hat{\lambda}_n - \lambda)] \sum_{k=1}^{n} X_{k-1}^{\alpha}.\end{aligned}$$

Thus (3.6) follows from (3.1), (3.2) and strong consistency of $\hat{m}_n$ and $\hat{\lambda}_n$, while for (3.7) we only have to note that $\sqrt{n}(\hat{m}_n - m, \hat{\lambda}_n - \lambda) \xrightarrow{d} N$, where N has a multivariate normal distribution [Wei and Winnicki (1990); Klimko and Nelson (1978)].

The above methods can also be used to derive asymptotic properties of the weighted conditional least squares estimators, which are summarized in the following theorem. The proof is omitted.

Theorem 3.4. *If $m<1$, $\sigma^2<\infty$ and $b^2<\infty$, then*

$$\tilde{\sigma}_n^2 \to \sigma^2, \quad \tilde{b}_n^2 \to b^2 \quad a.s. \tag{3.8}$$

If, in addition $E(\xi_{n,i}^4)<\infty$ and $E(\eta_n^4)<\infty$, then

$$\sqrt{n}(\tilde{\sigma}_n^2 - \sigma^2, \tilde{b}_n^2 - b^2)' \xrightarrow{d} N(0, \Psi^{-1} \Xi \Psi'^{-1}), \tag{3.9}$$

where

$$\Psi = \begin{pmatrix} E\dfrac{X}{X+1} & E\dfrac{1}{X+1} \\ E\dfrac{X}{(X+1)^2} & E\dfrac{1}{(X+1)^2} \end{pmatrix}$$

and

$$\Xi = \begin{pmatrix} E\dfrac{R(X)}{(X+1)^2} & E\dfrac{R(X)}{(X+1)^3} \\ E\dfrac{R(X)}{(X+1)^3} & E\dfrac{R(X)}{(X+1)^4} \end{pmatrix}.$$

The same statements hold with $\tilde{\sigma}_n^2$ and $\tilde{b}_n^2$ replaced by $\tilde{\tilde{\sigma}}_n^2$ and $\tilde{\tilde{b}}_n^2$.

We will now turn to the critical case. Here we will assume $\hat{m}_n = \tilde{m}_n$ and $\hat{\lambda}_n = \tilde{\lambda}_n$.

Theorem 3.5. *If* $m=1$, $E(\xi_{n,i}^4)<\infty$ *and* $E(\eta_n^4)<\infty$, *then*

(3.10)
$$\sqrt{\frac{n}{2\sigma^4}}\,(\hat{\sigma}_n^2-\sigma^2)\xrightarrow{d} L$$

and

(3.11)
$$\sqrt{\frac{n}{2\sigma^4}}\,(\hat{\hat{\sigma}}_n^2-\sigma^2)\xrightarrow{d} L,$$

where

$$L=\frac{\int_0^1 Y^2(t)\,dB(t)-\int_0^1 Y(t)\,dt\int_0^1 Y(t)\,dB(t)}{\int_0^1 Y^2(t)\,dt-\left(\int_0^1 Y(t)\,dt\right)^2}.$$

Proof. Notice that

(3.12)
$$\int_0^1 Y^2(t)\,dt-\left(\int_0^1 Y(t)\,dt\right)^2>0 \quad \text{a.s.}$$

We write

$$\sqrt{n}(\hat{\hat{\sigma}}_n^2-\hat{\sigma}_n^2)=\sqrt{n}(\hat{m}_n-m)\left[\frac{2}{n^3}\sum_{k=1}^n X_{k-1}U_k-(\hat{m}_n-m)\frac{1}{n^3}\sum_{k=1}^n X_{k-1}^2\right.$$
$$\left.-\frac{1}{n^2}\sum_{k=1}^n X_{k-1}\frac{2}{n^2}\sum_{k=1}^n U_k-(\hat{m}_n-m)\frac{1}{n}\sum_{k=1}^n X_{k-1}\right]$$
$$\cdot\left[\frac{1}{n^3}\sum_{k=1}^n X_{k-1}^2-\left(\frac{1}{n^2}\sum_{k=1}^2 X_{k-1}\right)^2\right]^{-1}.$$

It is now easily seen by Lemmas 2.7 and 2.13 above and Corollary 2.3 of Wei and Winnicki (1989) that

(3.13)
$$\sqrt{n}(\hat{\hat{\sigma}}_n^2-\hat{\sigma}_n^2)\xrightarrow{P}0$$

and it is enough to establish (3.10). But

$$\sqrt{n}(\hat{\sigma}_n^2-\sigma^2)=\left[\frac{1}{n^{5/2}}\sum_{k=1}^n X_{k-1}V_k-\frac{1}{n^2}\sum_{k=1}^n X_{k-1}\frac{1}{n^{3/2}}\sum_{k=1}^n V_k\right]$$
$$\cdot\left[\frac{1}{n^3}\sum_{k=1}^n X_{k-1}^2-\left(\frac{1}{n^2}\sum_{k=1}^n X_{k-1}\right)^2\right]^{-1},$$

so an application of Lemma 2.8 completes the proof.

Remark 3.6. The limiting random variable L can be seen to have the distribution of a mixture of a unit normal random variable and

$$\left\{\int_0^1\left[Y^2(s)-\int_0^1 Y(t)\,dt\,Y(s)\right]^2\left[\int_0^1 Y^2(t)\,dt-\left(\int_0^1 Y(t)\,dt\right)^2\right]^{-2}ds\right\}^{\frac{1}{2}}.$$

Remark 3.7. Theorem 3.5 implies that

$$\hat{\sigma}_n^2 \xrightarrow{P} \sigma^2 \quad \text{and} \quad \hat{\hat{\sigma}}_n^2 \xrightarrow{P} \sigma^2. \tag{3.14}$$

This was established under the assumption of finiteness of fourth moments. Using truncation arguments it can be shown that $\sum_{k=1}^{n} X_{k-1}^{\alpha} V_k = O(n^{2+\alpha})$ a.s. $(\alpha=0, 1)$, holds under the weaker assumption that $\sigma^2<\infty$ and $b^2<\infty$. Hence, (3.14) holds if only $\sigma^2<\infty$ and $b^2<\infty$.

Theorem 3.8. *If* $m=1$, $E(\xi_{n,i}^4)<\infty$ *and* $E(\eta_n^4)<\infty$. *Then*

$$\frac{1}{\sqrt{2\sigma^4 n}}(\hat{b}_n^2-b^2) \xrightarrow{d} \frac{\int_0^1 Y^2(t)\,dt \int_0^1 Y(t)\,dB(t)-\int_0^1 Y(t)\,dt\int_0^1 Y^2(t)\,dB(t)}{\int_0^1 Y^2(t)\,dt-\left(\int_0^1 Y(t)\,dt\right)^2}. \tag{3.15}$$

Proof. Write

$$\frac{1}{\sqrt{n}}(\hat{b}_n^2-b^2)=\left[\frac{1}{n^3}\sum_{k=1}^{n} X_{k-1}^2\frac{1}{n^{3/2}}\sum_{k=1}^{n} V_k-\frac{1}{n^2}\sum_{k=1}^{n} X_{k-1}\frac{1}{n^{5/2}}\sum_{k=1}^{n} X_{k-1}V_k\right]$$
$$\cdot\left[\frac{1}{n^3}\sum_{k=1}^{n} X_{k-1}^2-\left(\frac{1}{n^2}\sum_{k=1}^{n} X_{k-1}\right)^2\right]^{-1}$$

and apply Lemma 2.8 as in the proof of Theorem 3.5.

Remark 3.9. The limiting random variable in (3.15) has the distribution of a mixture $\eta\cdot N(0, 1)$ where

$$\eta=\left\{\int_0^1 Y^2(s)\left[\int_0^1 Y^2(t)\,dt-\int_0^1 Y(t)\,dt\,Y(s)\right]^2\left[\int_0^1 Y^2(t)\,dt-\left(\int_0^1 Y(t)\,dt\right)^2\right]^{-2}ds\right\}^{\frac{1}{2}}.$$

It is quite clear that for any $c\geqq 0$, $P(Y(s)=c$ for all $s\in[0, 1])=0$. It follows that $P(\eta=0)=0$. Hence,

$$|\hat{b}_n^2-b^2| \xrightarrow{P} \infty, \tag{3.16}$$

i.e. $\hat{b}_n^2$ is not a consistent estimator. It can also be shown, reasoning as in the proof of (3.13), that $\hat{\hat{b}}_n^2-\hat{b}_n^2=O_P(1)$, and it follows from (3.16) that $\hat{\hat{b}}_n^2$ is not a consistent estimator.

Somewhat surprisingly, the limit law in the analogue of Theorem 3.5 for the weighted conditional least squares estimators is standard normal.

Theorem 3.10. *If* $m=1$, $E(\xi_{n,i}^4)<\infty$ *and* $E(\eta_n^4)<\infty$, *then*

(3.17) $$\sqrt{\frac{n}{2\sigma^4}}(\tilde{\sigma}_n^2-\sigma^2)\xrightarrow{d}N(0,1)$$

and

(3.18) $$\sqrt{\frac{n}{2\sigma^4}}(\tilde{\tilde{\sigma}}_n^2-\sigma^2)\xrightarrow{d}N(0,1)$$

Proof. We write

(3.19) $$\tilde{\sigma}_n^2-\sigma^2=\frac{A_n-B_n}{1-C_n},$$

where

$$A_n=\frac{1}{n}\sum_{k=1}^{n}\frac{V_k}{X_{k-1}+1},$$

$$B_n=\sum_{k=1}^{n}\frac{1}{X_{k-1}+1}\sum_{k=1}^{n}\frac{V_k}{(X_{k-1}+1)^2}\Bigg/\left(n\sum_{k=1}^{n}\frac{1}{(X_{k-1}+1)^2}\right)$$

and

$$C_n=\left(\sum_{k=1}^{n}\frac{1}{X_{k-1}+1}\right)^2\Bigg/\left(n\sum_{k=1}^{n}\frac{1}{(X_{k-1}+1)^2}\right).$$

To establish (3.17) notice that by Lemma 2.4,

(3.20) $$\sqrt{n}A_n\xrightarrow{d}N(0,2\sigma^4),$$

while, using Lemma 2.11 and Lemma 2.12, it is elementary to show that $\sqrt{n}B_n\xrightarrow{P}0$ and $C_n\xrightarrow{P}0$. To establish (3.18) it is enough to show that

(3.21) $$\sqrt{n}(\tilde{\tilde{\sigma}}_n^2-\tilde{\sigma}^2)\xrightarrow{P}0.$$

We have

(3.22) $$\tilde{\tilde{\sigma}}_n^2-\tilde{\sigma}^2=\frac{D_n-E_n}{1-C_n},$$

where

$$D_n=\frac{1}{n}\sum_{k=1}^{n}\frac{\tilde{U}_k^2-U_k^2}{X_{k-1}+1}$$

and

$$E_n = \sum_{k=1}^{n} \frac{1}{X_{k-1}+1} \sum_{k=1}^{n} \frac{\tilde{U}_k^2 - U_k^2}{(X_{k-1}+1)^2} \bigg/ \left(n \sum_{k=1}^{n} \frac{1}{(X_{k-1}+1)^2} \right).$$

We only have to show that

(3.23) $$\sqrt{n} D_n \xrightarrow{P} 0$$

and

(3.24) $$\sqrt{n} E_n \xrightarrow{P} 0.$$

For (3.23), we write

(3.25) $$\sqrt{n} D_n = \frac{1}{\sqrt{n}} \left\{ 2(\tilde{m}_n - m) \sum_{k=1}^{n} \frac{X_{k-1}}{X_{k-1}+1} U_k + 2(\tilde{\lambda}_n - \lambda) \sum_{k=1}^{n} \frac{U_k}{X_{k-1}+1} \right.$$
$$+ (\tilde{m}_n - m)[(\tilde{m}_n - m) + (\tilde{\lambda}_n - \lambda)] \sum_{k=1}^{n} \frac{X_{k-1}}{X_{k-1}+1}$$
$$\left. + (\tilde{\lambda}_n - \lambda)[(\tilde{m}_n - m) + (\tilde{\lambda}_n - \lambda)] \sum_{k=1}^{n} \frac{1}{X_{k-1}+1} \right\}.$$

Now apply Lemmas 2.7, 2.10, 2.12 and 2.13. The proof of (3.24) is similar.

Remark 3.11. Relations (3.17) and (3.18) imply that $\tilde{\sigma}_n^2$ and $\tilde{\tilde{\sigma}}_n^2$ are (weakly) consistent estimators. Assuming only that

(3.26) $$E(\xi_{n,i}^2 \log^+ \xi_{n,i}) < \infty \quad \text{and} \quad b^2 < \infty,$$

the relation $C_n \xrightarrow{P} 0$ remains valid. Furthermore, using truncation arguments, it can be shown that $A_n \xrightarrow{P} 0$ and $\sum_{k=1}^{n} V_k/(X_{k-1}+1) = o_P(n^\alpha)$ for any $\alpha > 0$. This shows that assumption (3.26) is sufficient to prove consistency of $\tilde{\sigma}_n^2$. Similarly, $\tilde{\tilde{\sigma}}_n^2$ is consistent under (3.26) only, since the proof of (3.21) does not require higher moment assumptions.

Theorem 3.12. *Assume that* $m=1$, $2\lambda \leq \sigma^2$, $E(\xi_{n,i}^4) < \infty$ *and* $E(\eta_n^4) < \infty$. *Then*

(3.27) $$\tilde{b}_n^2 \xrightarrow{P} b^2$$

and

(3.28) $$\tilde{\tilde{b}}_n^2 \xrightarrow{P} b^2$$

Proof.

$$\tilde{b}_n^2 - b^2 = \frac{F_n - G_n}{1 - C_n},$$

where

$$F_n = \sum_{k=1}^{n} \frac{X_{k-1}}{X_{k-1}+1} \sum_{k=1}^{n} \frac{V_k}{(X_{k-1}+1)^2} \Bigg/ \left(n \sum_{k=1}^{n} \frac{1}{(X_{k-1}+1)^2} \right)$$

and

$$G_n = \sum_{k=1}^{n} \frac{X_{k-1}}{(X_{k-1}+1)^2} \sum_{k=1}^{n} \frac{V_k}{X_{k-1}+1} \Bigg/ \left(n \sum_{k=1}^{n} \frac{1}{(X_{k-1}+1)^2} \right).$$

To obtain (3.27), apply Lemmas 2.11 and 2.12 as in the proof of Theorem 3.10. The proof of (3.28) relies on the relation $\tilde{\tilde{b}}_n^2 - \tilde{b}_n^2 \xrightarrow{P} 0$, which is analogous to (3.21). Details are omitted.

Remark 3.13. If $m=1$ and $2\lambda > \sigma^2$, then $\tilde{b}_n^2$ and $\tilde{\tilde{b}}_n^2$ are not consistent estimators. This can be checked directly, but is also implied by a more general fact that in this case no consistent estimator of b^2 exists (cf. Sect. 4).

Remark 3.14. The limiting distributions of $\tilde{b}_n^2$ and $\tilde{\tilde{b}}_n^2$ in the case $m=1$ and $2\lambda \leq \sigma^2$ are not known.

To complete our study of the conditional least squares estimators we need to consider the supercritical case. We will first formulate a lemma closely related to Theorem 3.5 of Wei and Winnicki (1989).

Lemma 3.15. *If $m>1$, $\sigma^2<\infty$ and $b^2<\infty$, then*

$$\sum_{k=1}^{n} U_k^2 \Big/ \sum_{k=1}^{n} X_{k-1} \xrightarrow{d} \frac{m-1}{m} \sum_{k=0}^{\infty} m^{-k} \xi_k^2 \tag{3.29}$$

and

$$\sum_{k=1}^{n} X_{k-1} U_k^2 \Big/ \sum_{k=1}^{n} X_{k-1}^2 \xrightarrow{d} \frac{m^2-1}{m^2} \sum_{k=0}^{\infty} m^{-2k} \xi_k^2, \tag{3.30}$$

where $\{\xi_k, k=0, 1, \ldots\}$ are i.i.d. $N(0, 1)$ random variables.

The proof of Lemma 3.15 can be carried out using the methods of the last mentioned paper.

Theorem 3.16. *If $m>1$, $\sigma^2<\infty$ and $b^2<\infty$, then $\hat{\sigma}_n^2$ and $\tilde{b}_n^2$ are not consistent estimators.*

Proof. By (2.8), $m^{-(n-1)} \sum_{k=1}^{n} X_{k-1} \to mZ/(m-1)$ a.s. and $m^{-2(n-1)} \sum_{k=1}^{n} X_{k-1}^2 \to m^2 Z^2/(m^2-1)$ a.s. Together with Lemma 3.15, this implies that

$$\hat{\sigma}_n^2 \xrightarrow{d} \frac{m^2-1}{m^2} \sum_{k=0}^{\infty} m^{-2k} \xi_k^2$$

and the limit is clearly a nondegenerate random variable.

Using similar methods we obtain inconsistency of $\hat{b}_n^2$, but this also follows from a more general fact that if $m>1$, then no parameters of the immigration distribution have consistent estimators [Wei and Winnicki (1990)].

Remark 3.17. It can be shown, using Lemma 3.15 as above and the properties of $\hat{m}_n$ and $\hat{\lambda}_n$ [cf. Wei and Winnicki (1990)] that $\hat{\hat{\sigma}}_n^2$ is not a consistent estimator.

The last two theorems in this section give the asymptotic properties of the weighted conditional least squares estimators in the supercritical case.

Theorem 3.18. *If* $m>1$, $\sigma^2<\infty$ *and* $b^2<\infty$, *then*

$$\tilde{\sigma}_n^2 \to \sigma^2 \quad a.s. \tag{3.31}$$

If, in addition, $E(\xi_{n,i}^4)<\infty$ *and* $E(\eta_n^4)<\infty$, *then*

$$\sqrt{\frac{n}{2\sigma^4}}(\tilde{\sigma}_n^2-\sigma^2)\xrightarrow{d} N(0,1). \tag{3.32}$$

Proof. We will use (3.19). Since under the assumption $m>1$

$$\sum_{k=1}^{\infty}(X_{k-1}+1)^{-\alpha}<\infty \tag{3.33}$$

for any $\alpha>0$ (cf. Remark 2.5), it follows by the local martingale convergence theorem (Hall and Heyde (1980), Theorem 2.17) that $\sum_{k=1}^{n}\frac{V_k}{(X_{k-1}+1)^2}$ converges almost surely. Hence

$$\sqrt{n}B_n \to 0 \text{ a.s.} \quad \text{and} \quad C_n \to 0 \text{ a.s.} \tag{3.34}$$

An application of the strong law for martingales (Hall and Heyde (1980), Theorem 2.18) shows that $A_n \to 0$ a.s., which completes the proof of (3.31).

Remark 2.5 implies that $\sqrt{n}A_n \xrightarrow{d} N(0, 2\sigma^4)$, which together with (3.34) gives (3.32).

To prove a corresponding theorem for $\tilde{\tilde{\sigma}}_n^2$ we need a lemma characterizing the almost sure order of magnitude or $\tilde{\lambda}_n-\lambda$ and $\tilde{m}_n-m$ in the supercritical case.

Lemma 3.19. *If* $m>1$, $\sigma^2<\infty$ *and* $b^2<\infty$, *then*

$$\tilde{m}_n-m=O\left(\left(\log n \Big/ \sum_{k=1}^{n}(X_{k-1}+1)\right)^{\frac{1}{2}}\right) \quad a.s. \tag{3.35}$$

and

$$\tilde{\lambda}_n-\lambda=O(1) \quad a.s. \tag{3.36}$$

Proof. It is shown in the proof of Theorem 2.2 of Wei and Winnicki (1990) that

$$\tilde{m}_n - m = O\left(\sum_{k=1}^{n} U_k \Big/ \sum_{k=1}^{n} (X_{k-1}+1)\right) + o(1) \qquad \text{a.s.}$$

By Theorem 4 of Heyde and Leslie (1971),

$$\text{(3.37)} \quad \sum_{k=1}^{n} U_k = O\left(\left[\sum_{k=1}^{n}\left((X_{k-1}+1)\log k\right)\right]^{\frac{1}{2}}\right) = O\left(\left[\log n \sum_{k=1}^{n} (X_{k-1}+1)\right]^{\frac{1}{2}}\right) \quad \text{a.s.}$$

Hence (3.35) follows.
Similarly, using the proof of Theorem 2.5 of Wei and Winnicki (1990) and (3.37) it can be seen that

$$\tilde{\lambda}_n - \lambda = O\left(\sum_{k=1}^{n} \frac{U_k}{X_{k-1}+1} \Big/ \sum_{k=1}^{n} \frac{1}{X_{k-1}+1}\right) + o(1) \qquad \text{a.s.}$$

But by (3.33) and the local martingale convergence theorem,

$$\text{(3.38)} \qquad \sum_{k=1}^{n} \frac{U_k}{X_{k-1}+1} \quad \text{converges almost surely to a finite limit.}$$

This completes the proof of (3.36).

Theorem 3.20. *If* $m > 1$, $\sigma^2 < \infty$ *and* $b^2 < \infty$, *then*

$$\text{(3.39)} \qquad \tilde{\tilde{\sigma}}_n^2 \to \sigma^2 \qquad \textit{a.s.}$$

If, in addition, $E(\xi_{n,i}^4) < \infty$ *and* $E(\eta_n^4) < \infty$, *then*

$$\text{(3.40)} \qquad \sqrt{\frac{n}{2\sigma^4}}\,(\tilde{\tilde{\sigma}}_n^2 - \sigma^2) \xrightarrow{d} N(0, 1).$$

Proof. By Theorem 3.18, it is enough to show

$$\text{(3.41)} \qquad \sqrt{n}(\tilde{\tilde{\sigma}}_n^2 - \tilde{\sigma}_n^2) \to 0 \qquad \text{a.s.}$$

Using (3.22), we only have to prove that $\sqrt{n}D_n \to 0$ a.s. and $\sqrt{n}E_n \to 0$ a.s. But under the present assumption these relations are easily verified using Lemma 3.18, (3.33), (3.37) and (3.38).

Remark 3.21. As already mentioned in the proof of Theorem 3.16, there does not exist a consistent estimator for b^2 if $m > 1$.

4. Nonexistence of consistent estimators in the critical case

In this section we will show that with the exception of the first four moments of the offspring distribution and the mean of the immigration distribution, no parameters of the critical, transient branching process with immigration may have consistent estimators.

Our approach follows Wei and Winnicki (1990). Consider a time homogeneous Markov chain X_n on a countable state space $\mathscr{S}$. Its probability measure P is determined by an initial distribution ρ and a transition probability function p. We will assume that ρ can be any probability function on $\mathscr{S}$ and that $p \in \mathscr{P}$, where $\mathscr{P}$ is a given family of transition probability functions on $\mathscr{S}$. We will write $P=(\rho, p)$. Let $\theta(p)$ be a parameter of the transition function, i.e. $\theta: \mathscr{P} \to \Theta$, where Θ is a metric space with a metric d. We will say that θ admits a consistent estimator if there exists a sequence of random variables $\hat{\theta}_n$ such that for each $n \in N$, $\hat{\theta}_n$ is measurable with respect to $\sigma(X_0, \ldots, X_n)$ and $d(\hat{\theta}_n, \theta(p)) \xrightarrow{P} 0$ for any $P=(\rho, p)$, $p \in \mathscr{P}$.

The following necessary condition for existence of a consistent estimator was given by Wei and Winnicki (1990).

Proposition 4.1. *If there exists a consistent estimator $\hat{\theta}_n$ for θ, then*

$$\prod_{k=n_0}^{n_0+n} \frac{q(X_{k+1} \mid X_k)}{p(X_{k+1} \mid X_k)} \xrightarrow[n \to \infty]{} 0 \quad \text{a.s.} \; -(\rho, p) \tag{4.1}$$

for any $n_0 \in \mathbf{N}_0$ and $p, q \in \mathscr{P}$ such that $\theta(p) \neq \theta(q)$.

Now suppose that $\{X_n\}$ is a critical transient branching process with immigration, i.e. we assume that $m=1$ and $2\lambda > \sigma^2$. We will use subscripts $E_P(X)$, $\text{supp}_P(X)$ etc. to denote expectation, support and other parameters of the distribution of a random variable X under the probability measure P.

Let $h_P = \max\{h \in \mathbf{N}_0: P(\xi_{n,i} = u (\text{mod}\, h)) = 1$ for some $u \in \mathbf{N}_0\}$ be the span of the offspring distribution under P and let $u_P = \min\{u \in \mathbf{N}_0: P(\xi_{n,i} = u (\text{mod}\, h_P)) = 1\}$ be the corresponding offset.

For $r \in \{0, \ldots, h_P - 1\}$, let

$$T_P(r) = \sum_{i=0}^{\infty} P(\eta_1 = r + i h_P),$$

$$U_P(r) = \sum_{i=0}^{\infty} (r + i h_P) P(\eta_1 = r + i h_P),$$

$$V_P(r) = \sum_{i=0}^{\infty} (r + i h_P)^2 P(\eta_1 = r + i h_P).$$

Also denote $E_P(\xi_{n,i}^k)$ by $\alpha_{P,k}$, $E_P(\eta_n^k)$ by $\beta_{P,k}$ $(k=1, 2, \ldots)$ and let $m_P = \alpha_{P,1}$, $\sigma_P^2 = \alpha_{P,2} - m_P^2$, $\lambda_P = \beta_{P,1}$, $b_P^2 = \beta_{P,2} - \lambda_P^2$. We are now ready to formulate a basic asymptotic expansion of the transition function of $\{X_n\}$.

Lemma 4.2. *Consider two pairs $P=(\rho, p)$ and $Q=(\delta, q)$ of initial and transition probability functions of a branching process with immigration $\{X_n\}$. Assume that $\alpha_{P,7}, \alpha_{Q,7} < \infty$, $\beta_{P,5}, \beta_{Q,5} < \infty$, $m_P = m_Q = 1$, $2\lambda_P > \sigma_P^2$, $2\lambda_Q > \sigma_Q^2$.*

If

$$\text{supp}_P(X_n) \subseteq \text{supp}_Q(X_n) \tag{4.2}$$

for all sufficiently large n, then

$$(4.3)\qquad q(X_{n+1}\mid X_n)=\frac{h_Q\, e^{-\frac{1}{2}(X_{n+1}-X_n)^2/(\sigma_Q\sqrt{X_n})^2}}{(2\pi\sigma_Q^2 X_n)^{\frac{1}{2}}}\, M_Q(X_n, X_{n+1}) \qquad a.s.\ -P,$$

where

$$M_Q(X_n, X_{n+1})=T_Q(r_n)\left(1+\sum_{v=1}^{5}\frac{1}{X_n^{v/2}}\, g_{Q,v}(\Delta_n)\right)$$
$$+U_Q(r_n)\left(\frac{\Delta_n}{\sigma_Q^2\sqrt{X_n}}+\frac{1}{X_n}\,\frac{\alpha_{Q,3}}{6\alpha_{Q,2}^{3/2}\,\sigma_Q}(\Delta_n^4-6\Delta_n^2+3)\right)$$
$$+V_Q(r_n)\frac{1}{X_n}\,\frac{1}{2\sigma_Q^2}(\Delta_n^2-1)+O\left(\frac{(\log X_n)^7}{X_n^{3/2}}\right),$$
$$r_n=(X_{n+1}-u_Q X_n)(\mathrm{mod}\ h_Q),$$
$$\Delta_n=\frac{X_{n+1}-X_n}{\sigma_Q\sqrt{X_n}},$$

and $O\left(\frac{(\log X_n)^7}{X_n^{3/2}}\right)$ *denotes terms of the order of* $\frac{(\log X_n)^7}{X_n^{3/2}}$ *a.s.* $-P$*, which depend on* q*, while* $g_{Q,v}$ *are polynomials of degree* $3v$ *with coefficients depending on* $\alpha_{Q,1}, \dots, \alpha_{Q,v+2}$. $\Bigl($*Explicitly,* $g_{Q,v}=\sum H_{v+2s}(x)\prod_{j=1}^{v}\frac{1}{k_j!}\left(\frac{\gamma_{Q,j+1}}{\alpha_{Q,2}^{(j+2)/2}(j+2)!}\right)^{k_j}$*, where the summation is carried over all non-negative integer solutions of the equation* $k_1+2k_2+\dots+vk_v=v$*,* $s=k_1+k_2+\dots+k_v$*,* H_v *are the Hermite polynomials of order* v *and* $\gamma_{Q,j}$ *is the cumulant of order* j *of* $\xi_{n,i}$ *under* Q*; in particular,*

$$g_{Q,1}(x)=\frac{\alpha_{Q,3}}{6\alpha_{Q,2}^{3/2}}(x^3-3x),$$
$$g_{Q,2}(x)=\left(\frac{\alpha_{Q,3}^2}{72\alpha_{Q,2}^3}+\frac{\alpha_{Q,4}-3\alpha_{Q,2}^2}{24\alpha_{Q,2}^2}\right)(x^6-14x^4+39x^2-12).\Bigr)$$

Proof. Notice that the assumption $m_P=m_Q=1$ implies $u_P=u_Q=0$. For any $l\in\mathbf{N}_0$ and $n=kh_Q+r$, where $k\in\mathbf{N}_0$ and $T_Q(r)>0$, we have

$$(4.4)\qquad q(n\mid l)=Q\left(\sum_{j=1}^{l}\xi_{1,j}+\eta_1=kh_Q+r\right)$$
$$=\sum_{i=0}^{k}Q\left(\sum_{j=1}^{l}\xi_{1,j}=(k-i)\,h_Q\right)Q(\eta_1=ih_Q+r).$$

Using asymptotic expansion in local limit theorem [cf. Petrov (1975)],

$$Q\left(\sum_{j=1}^{l}\xi_{1,j}=ih_Q\right)=\frac{h_Q}{\sigma_Q\sqrt{l}}\,\varphi(x_i)\left(1+\sum_{v=1}^{5}\frac{1}{l^{v/2}}\,g_{Q,v}(x_i)\right)+O\left(\frac{1}{l^3}\right),$$

where $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $x_i=\frac{ih_Q-l}{\sigma\sqrt{l}}$, $\left|O\left(\frac{1}{l^3}\right)\right|\leqq\frac{c}{l^3}$ and c is a constant which does not depend on l and i. Here $o(\cdot)$ is a generic notation for 'terms of smaller order', which depend only on q. Similarly, $O(\cdot)$ is a generic notation for 'terms of the order'. Subscripts $o_1(\cdot)$, $o_2(\cdot)$, ... etc. will be added when several o or O terms appear in the same equation. By (4.4),

$$\begin{aligned}
q(n\,|\,l)=&\sum_{i=0}^{k}\frac{h_Q}{\sigma_Q\sqrt{l}}\varphi(x_{k-i})\left(1+\sum_{\nu=1}^{5}\frac{1}{l^{\nu/2}}g_{Q,\nu}(x_{k-i})\right)Q(\eta_1=ih_Q+r)+O\left(\frac{1}{l^3}\right)\\
=&\frac{h_Q}{\sigma_Q\sqrt{l}}\varphi(\Delta)\,T_Q(r)\left(1+\sum_{\nu=1}^{5}\frac{1}{l^{\nu/2}}g_{Q,\nu}(\Delta)\right)\\
&-\frac{h_Q}{\sigma_Q\sqrt{l}}\varphi(\Delta)\left(1+\sum_{\nu=1}^{5}\frac{1}{l^{\nu/2}}g_{Q,\nu}(\Delta)\right)\sum_{i=k+1}^{\infty}Q(\eta_1=ih_Q+r)\\
&+\sum_{i=0}^{k}\frac{h_Q}{\sigma_Q\sqrt{l}}[\varphi(x_{k-i})-\varphi(\Delta)]\,Q(\eta_1=ih_Q+r)\\
&+\sum_{i=0}^{k}\frac{h_Q}{\sigma_Q\sqrt{l}}\frac{1}{\sqrt{l}}\left[\varphi(x_{k-i})\,g_{Q,1}(x_{k-i})-\varphi(\Delta)\,g_{Q,1}(\Delta)\right]Q(\eta_1=ih_Q+r)\\
&+\sum_{i=1}^{k}\frac{h_Q}{\sigma_Q\sqrt{l}}\left[\varphi(x_{k-i})\sum_{\nu=2}^{5}\frac{1}{l^{\nu/2}}g_{Q,\nu}(x_{k-i})\right.\\
&\left.-\varphi(\Delta)\sum_{\nu=2}^{5}\frac{1}{l^{\nu/2}}g_{Q,\nu}(\Delta)\right]Q(\eta_1=ih_Q+r)+O\left(\frac{1}{l^3}\right),
\end{aligned}$$

where $\Delta=\frac{n-l}{\sigma_Q\sqrt{l}}$.

Since $\beta_{Q,5}<\infty$, $\sum_{i=k}^{\infty}i^{\alpha}Q(\eta_1=i)=o\left(\frac{1}{k^{5-\alpha}}\right)$, $0\leqq\alpha\leqq 5$. Hence,

$$\sum_{i=k+1}^{\infty}i^{\alpha}Q(\eta_1=ih_Q+r)=o\left(\frac{1}{k^{5-\alpha}}\right),\qquad 0\leqq\alpha\leqq 5. \tag{4.5}$$

Using a five term Taylor expansion,

$$\begin{aligned}
\varphi(x_{k-i})-\varphi(\Delta)=\varphi(\Delta)\Bigg\{&\Delta\frac{1}{\sigma_Q\sqrt{l}}(ih_Q+r)+\frac{1}{2\sigma_Q^2}\frac{1}{l}(\Delta^2-1)(ih_Q+r)^2\\
&+\frac{1}{6\sigma_Q^3}\frac{1}{l^{3/2}}H_3(\Delta)(ih_Q+r)^3+\frac{1}{24\sigma_Q^4}\frac{1}{l^2}H^4(\Delta)(ih_Q+r)^4\Bigg\}\\
&+\frac{\varphi(\theta_i)\,H_5(\theta_i)(ih_Q+r)^5}{120\sigma_Q^5}\frac{1}{l^{5/2}},
\end{aligned}$$

where θ_i is a point between x_{k-i} and Δ.

It follows that

$$\sum_{i=0}^{k} \frac{h_Q}{\sigma_Q \sqrt{l}} [\varphi(x_{k-i}) - \varphi(\Delta)]\, Q(\eta_1 = i h_Q + r)$$

$$= \frac{h_Q}{\sigma_Q \sqrt{l}} \varphi(\Delta) \Bigg\{ \Delta \frac{1}{\sigma_Q \sqrt{l}} U_Q(r) + \frac{1}{2\sigma_Q^2} \frac{1}{l} (\Delta^2 - 1) V_Q(r)$$

$$+ \frac{1}{6\sigma_Q^3} \frac{1}{l^{3/2}} H_3(\Delta) \sum_{i=1}^{k} (i h_Q + r)^3 Q(\eta_1 = i h_Q + r)$$

$$+ \frac{1}{24\sigma_Q^4} \frac{1}{l^2} H_4(\Delta) \sum_{i=1}^{k} (i h_Q + r)^4 Q(\eta_1 = i h_Q + r)$$

$$+ \Delta \frac{1}{\sigma_Q \sqrt{l}} \sum_{i=k+1}^{\infty} (i h_Q + r)\, Q(\eta_1 = i h_Q + r)$$

$$+ \frac{1}{2\sigma_Q^2} \frac{1}{l} (\Delta^2 - 1) \sum_{i=k+1}^{\infty} (i h_Q + r)^2 Q(\eta_1 = i h_Q + r) \Bigg\} + O\left(\frac{1}{l^3}\right),$$

where we used the fact that for any polynomial g,

$$\sup_{-\infty < x < \infty} |g(x)|\, \varphi(x) < \infty. \tag{4.6}$$

Using the assumption $\alpha_{Q,k} < \infty$, $k = 1, \ldots, 7$ and (4.5) we obtain

$$\sum_{i=0}^{k} \frac{h_Q}{\sigma_Q \sqrt{l}} [\varphi(x_{k-i}) - \varphi(\Delta)]\, Q(\eta_1 = i h_Q + r) \tag{4.7}$$

$$= \frac{h_Q}{\sigma_Q \sqrt{l}} \varphi(\Delta) \Bigg\{ \Delta \frac{1}{\sigma_Q \sqrt{l}} U_Q(r) + \frac{1}{2\sigma_Q^2} \frac{1}{l} (\Delta^2 - 1) V_Q(r) + \frac{1}{6\sigma_Q^3} \frac{1}{l^{3/2}} H_3(\Delta)\, O_1(1)$$

$$+ \frac{1}{24\sigma_Q^4} \frac{1}{l^2} H_4(\Delta)\, O_2(1) + \frac{1}{\sqrt{l}} \Delta\, o_1\left(\frac{1}{k^4}\right) + \frac{1}{l} (\Delta^2 - 1)\, o_2\left(\frac{1}{k^3}\right) \Bigg\} + O\left(\frac{1}{l^3}\right).$$

Similarly,

$$\sum_{i=0}^{k} \frac{h_Q}{\sigma_Q \sqrt{l}} \frac{1}{\sqrt{l}} \Big[\varphi(x_{k-i})\, g_{Q,1}(x_{k-i}) - \varphi(\Delta)\, g_{Q,1}(\Delta) \Big] Q(\eta_1 = i h_Q + r) \tag{4.8}$$

$$= \frac{h_Q}{\sigma_Q \sqrt{l}} \varphi(\Delta) \Bigg\{ \frac{\alpha_{Q,3}}{6\alpha_{Q,2}^{3/2} \sigma_Q} \frac{1}{l} (\Delta^4 - 6\Delta^2 + 3)\, U_Q(r) + \frac{1}{l^{3/2}} g_1(\Delta)\, O_3(1)$$

$$+ \frac{1}{l^2} g_2(\Delta)\, O_4(1) + \frac{1}{l} (\Delta^4 - 6\Delta^2 + 3)\, o_3\left(\frac{1}{k^4}\right) \Bigg\} + O\left(\frac{1}{l^3}\right),$$

where g_1 and g_2 are polynomials of order 5 and 6, respectively, whose coefficients depend on the moments $\alpha_{Q,1}, \ldots, \alpha_{Q,7}$.

Finally,

$$(4.9)\quad \sum_{i=1}^{k} \frac{h_Q}{\sigma_Q\sqrt{l}}\left[\varphi(x_{k-i})\sum_{\nu=1}^{5}\frac{1}{l^{\nu/2}}\,g_{Q,\nu}(x_{k-i})-\varphi(\Delta)\sum_{\nu=2}^{5}\frac{1}{l^{\nu/2}}\,g_{Q,\nu}(\Delta)\right]Q(\eta_1=ih_Q+r)$$
$$=\frac{h_Q}{\sigma_Q\sqrt{l}}\,\varphi(\Delta)\left\{\frac{1}{l^{3/2}}\,g_3(\Delta)\,O_5(1)+\frac{1}{l^2}\,g_4(\Delta)\,O_6(1)\right\}+O\left(\frac{1}{l^3}\right),$$

where g_3 and g_4 are polynomials of order 7 and 10, respectively.

By (4.4)–(4.9),

$$(4.10)\quad g(n\,|\,l)=\frac{h_Q}{\sigma_Q\sqrt{l}}\,\varphi(\Delta)\left\{T_Q(r)\left(1+\sum_{\nu=1}^{5}\frac{1}{l^{\nu/2}_{Q,\nu}}(\Delta)\right)\right.$$
$$+U_Q(r)\left[\frac{1}{\sigma_Q\sqrt{l}}\,\Delta+\frac{\alpha_{Q,3}}{6\sigma_{Q,2}^2}\,\frac{1}{l}(\Delta^4-6\Delta^2+3)\right]$$
$$+\frac{1}{2\sigma_Q^2}\,\frac{1}{l}\,V_Q(r)(\Delta^2-1)+\left(1+\sum_{\nu=1}^{5}\frac{1}{l^{\nu/2}}\,g_{Q,\nu}(\Delta)\right)o_4\left(\frac{1}{k^5}\right)$$
$$+\left[\frac{1}{\sqrt{l}}\,\Delta+\frac{1}{l}(\Delta^4-6\Delta^2+3)\right]o_5\left(\frac{1}{k^4}\right)+\frac{1}{l}(\Delta^2-1)\,O_6\left(\frac{1}{k^3}\right)$$
$$\left.+\frac{1}{l^{3/2}}\,g_5(\Delta)\,O_7(1)+\frac{1}{l^2}\,g_6(\Delta)\,O_8(1)\right\}+O\left(\frac{1}{l^3}\right),$$

where g_5 and g_6 are polynomials of degree 7 and 10, respectively.

Thus, by (4.2),

$$q(X_{n+1}\,|\,X_n)=\frac{h_Q}{\sqrt{2\pi\sigma_Q^2X_n}}\,e^{-\frac{1}{2}\Delta_n^2}\left\{T_Q(r_n)\left(1+\sum_{\nu=1}^{5}\frac{1}{X_n^{\nu/2}}\,g_{Q,\nu}(\Delta_n)\right)\right.$$
$$+U_Q(r_n)\left[\frac{\Delta_n}{\sigma_Q\sqrt{X_n}}+\frac{1}{X_n}\,\frac{\alpha_{Q,3}}{6\alpha_{Q,2}^{3/2}\sigma_Q}(\Delta_n^4-6\Delta_n^2+3)\right]$$
$$+V_Q(r_n)\frac{1}{X_n}\,\frac{1}{2\sigma_Q^2}(\Delta_n^2-1)+\left(1+\sum_{\nu=1}^{5}\frac{1}{X_n^{\nu/2}}\,g_{Q,\nu}(\Delta_n)\right)o\left(\frac{1}{X_{n+1}^5}\right)$$
$$+\left[\frac{\Delta_n}{\sqrt{X_n}}+\frac{1}{X_n}(\Delta_n^4-6\Delta_n^2+3)\right]o\left(\frac{1}{X_{n+1}^4}\right)$$
$$+\frac{1}{X_n}(\Delta_n^2-1)\,o\left(\frac{1}{X_{n+1}^3}\right)+\frac{1}{X_n^{3/2}}\,g_5(\Delta_n)\,O_7(1)$$
$$\left.+\frac{1}{X_n^2}\,g_6(\Delta_n)\,O_8(1)\right\}+O\left(\frac{1}{X_n^3}\right).$$

To complete the proof, we observe that by Lemma 2.14 and Theorem 2.15 of Wei and Winnicki (1989),

$$X_{n+1}/X_n \to 1 \quad \text{a.s.} \ -P,$$

(4.11)
$$\frac{X_{n+1}-X_n}{\sqrt{X_n}} = O(\sqrt{\log X_n}) \quad \text{a.s.} \ -P,$$

and

$$\frac{1}{X_n^3} = \frac{e^{-\frac{1}{2}(X_{n+1}-X_n)^2/(\sigma_Q\sqrt{X_n})^2}}{\sqrt{X_n}} \, O\left(\frac{1}{X_n^{3/2}}\right) \quad \text{a.s.} \ -P.$$

Remark 4.3. If the span h_Q of the offspring distribution divides the span of the immigration distribution under Q, then $T_Q(r_n)=1$ a.s.-Q. Under the assumption (4.2) the above relations hold also a.s.-P.

Corollary 4.4. *Suppose that, in addition to the assumptions of Lemma* 4.2, *the span of the offspring distribution divides the span of the immigration distribution under both P and Q, that $h_Q \geqq h_P$ and that*

(4.12)
$$\sigma_Q^2=\sigma_P^2, \qquad \lambda_Q=\lambda_P, \qquad \alpha_{Q,3}=\alpha_{P,3}, \qquad \alpha_{Q,4}=\alpha_{P,4}.$$

Then there exists an integer valued random variable N such that

(4.13)
$$\lim_{n\to\infty} \prod_{k=N}^{N+n} \frac{q(X_{k+1}\,|\,X_k)}{p(X_{k+1}\,|\,X_k)} > 0 \quad \textit{a.s.} \ -P.$$

Proof. Clearly, Lemma 4.2 remains valid when $Q=P$. It follows that

$$N=\inf\{n\colon p(X_{k+1}\,|\,X_k)\, q(X_{k+1}\,|\,X_k)>0 \quad \text{for all } \ k \geqq n\}$$

is finite a.s.-P. Notice that by Remark 4.3 and (4.3)

(4.14)
$$\begin{aligned} M_Q(X_k, X_{k+1}) = {} & 1+\sum_{\nu=1}^{5} \frac{1}{X_k^{\nu/2}}\, g_{Q,\nu}(\Delta_k)) \\ & + \left[\frac{\Delta_k}{\sigma_Q\sqrt{X_k}} + \frac{\alpha_{Q,3}}{6\alpha_{Q,2}^{3/2}\sigma_Q}\, \frac{1}{X_k}(\Delta_k^4-6\Delta_k^2+3)\right]\lambda_Q \\ & + \frac{1}{2\sigma_Q^2}\, \frac{1}{X_k}(\Delta_k^2-1)(b_Q^2+\lambda_Q^2) + O\left(\frac{(\log X_k)^7}{X_k^{3/2}}\right) \quad \text{a.s.} \ -P. \end{aligned}$$

Since $\sigma_Q^2=\sigma_P^2$, it is enough to show that

$$\lim_{n\to\infty} \sum_{k=N}^{N+n} \frac{M_Q(X_k, X_{k+1})}{M_P(X_k, X_{k+1})} > 0 \quad \text{a.s.} \ -P.$$

(The limit in (4.13) is infinity if $h_Q > h_P$).

Hence it is sufficient to show that the series

$$\sum_{k=N}^{\infty}\left(\frac{M_Q(X_k, X_{k+1})}{M_P(X_k, X_{k+1})}-1\right) \quad \text{converges a.s. } -P \tag{4.15}$$

and

$$\sum_{k=N}^{\infty}\left(\frac{M_Q(X_k, X_{k+1})}{M_P(X_k, X_{k+1})}-1\right)^2<\infty \quad \text{a.s. } -P. \tag{4.16}$$

By (4.11), (4.12), (4.14) and the fact that $g_{Q,1}$ and $g_{Q,2}$ depend only on $\alpha_{Q,1}, \ldots, \alpha_{Q,4}$ we have

$$\frac{M_Q(X_k, X_{k+1})}{M_P(X_k, X_{k+1})}-1=\frac{\dfrac{b_Q^2-b_P^2}{2\sigma_Q^2}\dfrac{1}{X_k}(\Delta_k^2-1)+O\left(\dfrac{(\log X_k)^9}{X_k^{3/2}}\right)}{1+O\left(\dfrac{(\log X_k)^{3/2}}{X_k^{1/2}}\right)}. \tag{4.17}$$

Using (4.11) it follows that

$$\left(\frac{M_Q(X_k, X_{k+1})}{M_P(X_k, X_{k+1})}-1\right)^2=O\left(\left(\frac{\log X_k}{X_k}\right)^2\right) \quad \text{a.s. } -P.$$

Hence, (4.16) obtains by Lemma 2.13 of Wei and Winnicki (1989). The same result implies that $\sum_{k=N}^{\infty}\frac{(\log X_k)^9}{X_k^{3/2}}<\infty$ a.s.-P and it remains to show that

$$\sum_{k=N}^{\infty}\frac{\dfrac{1}{X_k}(\Delta_k^2-1)}{1+O\left(\dfrac{(\log X_k)^{3/2}}{X_k^{1/2}}\right)} \tag{4.18}$$

converges a.s.-P. Using (4.11) again, it is easy to see that we only have to show convergence of $\sum_{k=N}^{\infty}\frac{1}{X_k}(\Delta_k^2-1)$. Now,

$$E_P\left[\frac{1}{X_k}(\Delta_k^2-1)\,|\,\mathscr{F}_k\right]=\frac{b_P^2+\lambda_P^2}{\sigma_P^2 X_k^2},$$

so

$$E_P\left[\frac{1}{X_k}(\Delta_k^2-1)\,|\,\mathscr{F}_k\right] \quad \text{converges a.s. } -P. \tag{4.19}$$

Furthermore,

$$E_P\left\{\left[\frac{1}{X_k}(\Delta_k^2-1)-\frac{b_P^2+\lambda_P^2}{\sigma_P^2 X_k^2}\right]^2 \,\middle|\, \mathscr{F}_k\right\}=O\left(\frac{1}{X_k^2}\right), \tag{4.20}$$

so by the local martingale convergence theorem (Hall and Heyde (1980), Theorem 2.17) the martingale

$$\sum_{k=N}^{N+n} \left\{ \frac{1}{X_k}(\Delta_k^2 - 1) - \frac{b_P^2 + \lambda_P^2}{\sigma_P^2 X_k^2} \right\} \quad \text{converges a.s.} \ -P. \tag{4.21}$$

From (4.19) and (4.21) we conclude the convergence of $\sum_{k=N}^{\infty} \frac{1}{X_k}(\Delta_k^2 - 1)$ and Corollary 4.4 is proved.

The following result on nonexistence of consistent estimators for the parameters of the branching process with immigration is an immediate consequence of Proposition 4.1 and Corollary 4.4.

Theorem 4.5. *Consider the class $\mathscr{D} = \mathscr{D}(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \lambda, h, v)$ of transition functions corresponding to branching processes with immigration satisfying the assumptions of Lemma* 4.2, *such that the span of the offspring distribution divides the span of the immigration distribution, having common span h of the offspring distribution, common offset v of the immigration distribution, common first four moments $\alpha_1, \ldots, \alpha_4$ of the offspring distribution and common mean λ of the immigration distribution. Then no parameter of the critical, transient branching process with immigration which takes at least two values on $\mathscr{D}$ has a consistent estimator.*

Remark 4.6. Theorem 4.5 shows that the only parameters of the critical, transient branching process with immigration that may have consistent estimators are the span h of the offspring distribution, offset v of the immigration distribution, the first four moments $\alpha_1, \ldots, \alpha_4$ of the offspring distribution, the mean λ of the immigration distribution and functions of $\alpha_1, \ldots, \alpha_4, \lambda, h, v$. In order to show that the lattice parameters h and v do not have consistent estimators in general, we notice that the conclusion of Corollary 4.4 holds also under the following assumptions:

(1) The assumptions of Lemma 4.2
(2) (4.12)
(3) $h_Q T_Q(r_n) = h_P T_P(r_n)$, $h_Q U_Q(r_n) = h_P U_P(r_n)$ a.s. $-P$.

Essentially the same proof works. The relation (4.17) now reads

$$\frac{M_Q(X_k, X_{k+1})}{M_P(X_k, X_{k+1})} - 1 = \frac{\dfrac{V_Q(r_n) - V_P(r_n)}{2\sigma_Q^2} \dfrac{1}{X_k}(\Delta_k^2 - 1) + O\left(\dfrac{(\log X_k)^9}{X_k^{3/2}}\right)}{1 + O\left(\dfrac{(\log X_k)^{3/2}}{X_k^{1/2}}\right)}$$

and we find that a.s. $-P$ convergence of the series

$$\sum_{k=N}^{\infty} \frac{V_Q(r_k) - V_P(r_k)}{X_k}(\Delta_k^2 - 1)$$

has to be established. This follows, as in the proof of Corollary 4.4, by an application of the martingale convergence theorem, but we have to notice that

$$E\left[\frac{V_Q(r_k)-V_P(r_k)}{X_k}(\Delta_k^2-1)\,|\,\mathscr{F}_k\right]=\frac{c_{Q,P}}{\sigma_Q^2 X_k^2},$$

where $c_{Q,P}=E[(V_Q(\eta_n)-V_P(\eta_n))\,\eta_n^2]$ and similarly adjust the proof of (4.20).

It is easy to construct examples of pairs of probability measures P, Q satisfying the assumptions (1)–(3) and such that $h_Q \neq h_P$ and $v_Q \neq v_P$.

Let $\operatorname{supp}_Q(\xi_{n,i})=\{0, 2, 4, \ldots\}$, $\operatorname{supp}_P(\xi_{n,i})=\{0, 1, 2, \ldots\}$, $Q(\eta_n=1)=\frac{1}{4}$, $Q(\eta_n=2)=\frac{1}{2}$, $Q(\eta_n=3)=\frac{1}{4}$, $P(\eta_n=0)=\frac{1}{2}$, $P(\eta_n=2)=\frac{1}{2}$.

Then $\operatorname{supp}_Q(X_n)=\operatorname{supp}_P(X_n)$ for all $n \geqq 1$, $h_Q=2$, $v_Q=1$, $T_Q(0)=\frac{1}{2}$, $T_Q(1)=\frac{1}{2}$, $U_Q(0)=1$, $U_Q(1)=1$, $h_P=1$, $v_P=0$, $T_P(0)=1$, $U_P(0)=1$.

Acknowledgement. I would like to thank Professor C.Z. Wei for suggesting this topic to me and for providing several important ideas used in this paper.

References

Athreya, K.B., Ney, P.E.: Branching processes. New York Heidelberg Berlin: Springer 1972
Hall, P., Heyde, C.C.: Martingale limit theory and its application. New York: Wiley 1980
Helland, I.S.: Central limit theorems for martingales with discrete or continuous time. Scand. J. Stat. **9**, 79–94 (1982)
Heyde, C.C.: On estimating the variance of the offspring distribution in a simple branching process. Adv. Appl. Probab. **6**, 421–433 (1974)
Heyde, C.C., Leslie, J.R.: Improved classical limit analogues for Galton-Watson process with or without immigration. Bull. Austral. Math. Soc. **5**, 145–155 (1971)
Heyde, C.C., Seneta, E.: Estimation theory for growth and immigration rates in a multiplicative process. J. Appl. Probab. **9**, 235–258 (1972)
Klimko, L.A., Nelson, P.I.: On conditional least squares estimation for stochastic processes. Ann. Stat. **6**, 629–642 (1978)
Nelson, P.I.: A note on strong consistency of least squares estimators in regression models with martingale difference errors. Ann. Stat. **5**, 1057–1064 (1980)
Pakes, A.G.: Non-parametric estimation in the Galton-Watson process. Math. Biosci. **24**, 71–92 (1975)
Petrov, V.V.: Sums of independent random variables. New York Heidelberg Berlin: Springer 1975
Quine, M.P.: Asymptotic results for estimators in a subcritical branching process with immigration. Ann. Probab. **4**, 319–325 (1976); Correction **5**, 318
Seneta, E.: A note on the supercritical Galton-Watson process with immigration. Math. Biosci. **6**, 305–311 (1970)
Straser, H.: Martingale difference arrays and stochastic integrals. Probab. Th. Rel. Fields **72**, 83–98 (1986)
Wei, C.Z., Winnicki, J.: Some asymptotic results for the branching process with immigration. Stochastic Processes Appl. **31**, 261–282 (1989)
Wei, C.Z., Winnicki, J.: Estimation of the means in the branching process with immigration. Ann. Stat. **18**, 757–773 (1990)
Yanev, N.M., Tchoukova-Dantcheva, S.: On the statistics of branching processes with immigration. C.R. Acad. Bulg. Sci. **33**, 463–471 (1980)