Sankhyā : The Indian Journal of Statistics
2003, Volume 65, Part 1, pp 107-121

A NOTE ON ESTIMATION IN MULTITYPE SUPERCRITICAL BRANCHING PROCESSES WITH IMMIGRATION

By SANJAY SHETE
University of Texas M.D. Anderson Cancer Center, Houston, USA
and
T.N. SRIRAM
University of Georgia, Athens, USA

SUMMARY. For multitype branching processes with immigration, weighted conditional least squares estimator of the mean matrix M and the maximal eigenvalue ρ of M are developed based on little more information about the process than just the generation sizes. For the supercritical case, strong consistency and asymptotic normality of the estimators are established. Comparisons in terms of asymptotic variances show that the weighted conditional least squares estimators derived here are as good as the maximum likelihood estimators obtained by Asmussen and Keiding (1978) under the full family tree information.

1. Introduction

Consider a multitype branching process with immigration which admits p different types of particles. An individual of type i is assumed to produce offspring of p different types. Let $\boldsymbol{Z}_n = (Z_n^{(1)}, \ldots, Z_n^{(p)})$ where $Z_n^{(j)}$, $j = 1, \ldots, p$, denote the number of type j particles in the n-th generation. Also, let $\boldsymbol{Z}_0 = \boldsymbol{e}_k$, where $\boldsymbol{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ with 1 in the k-th position. Then, for each $n \geq 1$ we can write

$$Z_n^{(j)} = \sum_{i=1}^{p} \sum_{k=1}^{Z_{n-1}^{(i)}} \xi_{n-1,i,k}^{(j)} + Y_n^{(j)}, \quad j = 1, \ldots, p \tag{1.1}$$

where $\xi_{n-1,i,k}^{(j)}$ denotes the number of type j offspring produced by the k-th individual who is of type i belonging to the $(n-1)$th generation and $Y_n^{(j)}$ denotes the number of type j immigrants in the n-th generation.

Paper received April 1999; revised August 2001.
AMS (2000) *subject classifications.* Primary 60J80; secondary 62M05.
Keywords and phrases. Supercritical, weighted conditional least squares, maximal eigenvalue, martingales, invariance principle.

Throughout we will assume that for each $i = 1, \ldots, p$, $\{\boldsymbol{\xi}^i_{n,k} = (\xi^{(1)}_{n,i,k}, \ldots, \xi^{(p)}_{n,i,k});\ k, n \geq 1\}$ and $\{(Y^{(1)}_n, \ldots, Y^{(p)}_n); n \geq 1\}$ are two independent sequences of i.i.d. random vectors. Furthermore, for each $n, k \geq 1$, the vectors $\{\boldsymbol{\xi}^1_{n,k}, \ldots, \boldsymbol{\xi}^p_{n,k}\}$ are independent. Let m_{ij} be the mean number of type j offspring of a type i individual and $M = ((m_{ij}))$ be the mean matrix. Let $\lambda = (\lambda_1, \ldots, \lambda_p)$ be the mean of the immigration random vector $(Y^{(1)}_1, \ldots, Y^{(p)}_1)$. It is assumed that the means m_{ij} and λ_i are positive and finite for each $i, j = 1, \ldots, p$. Let Σ_i denote the variance-covariance of the vector $(\xi^{(1)}_{1,i,1}, \ldots, \xi^{(p)}_{1,i,1})$, $i = 1, \ldots, p$, and Δ be the variance-covariance matrix of the immigration vector.

It is well known (see, Athreya and Ney (1972), Chapter V.2) that if the mean matrix M is positively regular (that is, there exists a positive integer n such that all entries of M^n are strictly positive), then, by the Perron-Frobenius theorem, M has a maximal eigenvalue which is positive. Assume throughout that M is positively regular and let ρ denote the maximal eigenvalue of M. The parameter ρ plays the same role as the offspring mean of a single-type branching process in that the growth behavior of the process $Z^{(i)}_n$ is governed by the value of ρ. More precisely, let $q^i = P[||\boldsymbol{Z}_n|| \to 0 | \boldsymbol{Z}_0 = \boldsymbol{e}_i]$, where $||\boldsymbol{Z}_n|| = \max(Z^{(1)}_n, \ldots, Z^{(p)}_n)$. Then, for $\rho \leq 1$, $q^i = 1$ for $i = 1, \ldots, p$ and for $\rho > 1$, $q^i < 1$ for all i. The multitype process $\boldsymbol{Z}_n$ is said to be subcritical if $\rho < 1$, critical if $\rho = 1$ and supercritical if $\rho > 1$.

The study of estimation of ρ and the elements of M has been considered in the literature. For two-type Galton-Watson process, Badalbaev (1976) proposed a ratio-type estimator of ρ. For the case $\rho > 1$, Becker (1977) proposed three different ratio-type estimators of ρ and established their strong consistency on the set of non-extinction. Nanthi (1978, 1980) also proposed a strongly consistent estimator of ρ for the case $\rho > 1$.

Asmussen and Keiding (1978), however, considered the maximum likelihood (m.l.) estimators of the elements of M and ρ, based upon the knowledge of entire parent-offspring combinations in the first n generations. For the case $\rho > 1$, they established the almost sure (a.s.) limit behavior and asymptotic normality of their m.l. estimators. In addition, Asmussen and Keiding (1978) also studied the limit distribution of Becker's estimator of ρ for $\rho > 1$.

For the subcritical case $\rho < 1$, Quine and Durham (1977) used a formal analogy between the multitype branching process with immigration and the vector-valued autoregression to derive strongly consistent and asymptotically normal estimators of M and ρ. The probabilistic aspects of critical case $\rho = 1$ has been studied by Athreya and Ney (1972) and Joffe and Metivier

(1986), and the statistical estimation of ρ in this case has been considered by Badalbaev and Mukhitdinov (1989). See Dion (1993) for a more detailed discussion on some of the results mentioned above.

The previously mentioned results do not solve the problem of how to estimate the elements of M and ρ if we do not know whether $\rho < 1$, $\rho = 1$ or $\rho > 1$ and no unified estimation theory for M and ρ, such as the one in Wei and Winnicki (1990) for single-type case, is available in the literature. In an attempt to provide such a unified estimation theory, we use the idea of weighted conditional least squares (as discussed in Wei and Winnicki (1990)) and propose estimators of elements of M and ρ.

2. Estimators and Their Properties

Before we develop a method of estimation it should be mentioned that the question of estimability of elements of M, unfortunately, depends on what is and what is not observable. In fact, if we only observe the generation sizes of each type, then it can be shown that the elements of M are *not* consistently estimable. Such a result can be established using methods similar to those in Guttorp and Siegel (1985) and Lockhart (1982); see, for instance, Shete (1998), section 4.2 for details. However, ρ can be estimated based on the generation sizes alone, namely, using Becker's (1977) estimate.

¿From these it is clear that one needs a little more information about the process than just the generation sizes (not necessarily the entire parent-offspring combinations as assumed in Asmussen and Keiding (1978)) in order to obtain consistent estimators and hence develop a unified estimation theory for M as well as ρ. To this end, we assume that $Z_{k,i}^{(j)}$, the number of type j individuals in the k-th generation whose parents were of type i, is observable. Incidentally, Keiding and Lauritzen (1978) observed that the m.l. estimators of m_{ij} obtained in Asmussen and Keiding (1978) under full family tree information are also the m.l. estimators of m_{ij} and ρ based upon $Z_{k,i}^{(j)}$'s only. We will show below that our weighted conditional least squares estimators of m_{ij} and ρ based on $Z_{k,i}^{(j)}$ have the same asymptotic properties as the m.l. estimators in Asmussen and Keiding (1978); see end of this section.

In this article we only present the results for the case $\rho > 1$. Asymptotic properties of weighted conditional least squares estimator of M and ρ for the cases $\rho < 1$ and $\rho = 1$ are not known yet. Henceforth, for notational simplicity, we consider the case $p = 2$, the two-type branching process with immigration. Let $\tilde{\boldsymbol{Z}}_k = (Z_{k,1}^{(1)}, Z_{k,1}^{(2)}, Z_{k,2}^{(1)}, Z_{k,2}^{(2)})$ be observable for each $k \geq 1$. Note then that $Z_n^{(j)} = Z_{n,1}^{(j)} + Z_{n,2}^{(j)}$, $j = 1, 2$. Let $\tilde{\boldsymbol{Y}}_n = (Y_{n,1}^{(1)}, Y_{n,1}^{(2)}, Y_{n,2}^{(1)}, Y_{n,2}^{(2)})$, where $Y_{n,i}^{(j)}$ denotes the number of type j immigrants in the n-th generation whose parents were of type i. It is assumed that the sequence of offspring random vectors and $\{\tilde{\boldsymbol{Y}}_n\}$ are independent. However, the vector $\tilde{\boldsymbol{Y}}_n$ is *not* assumed to be observable. Let $\tilde{\boldsymbol{\Lambda}} = (\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22})$ denote the mean of $\tilde{\boldsymbol{Y}}_n$ and $\tilde{\Delta} = \begin{bmatrix} \Delta_1 & \mathbf{0} \\ \mathbf{0} & \Delta_2 \end{bmatrix}$ denote the variance-covariance of $\tilde{\boldsymbol{Y}}_n$ where Δ_i and $\mathbf{0}$ are 2×2 matrices. Using these notations and the ones in Section 1, we can write

$$Z_{n,i}^{(j)} = \sum_{k=1}^{Z_{n-1}^{(i)}} \xi_{n-1,i,k}^{(j)} + Y_{n,i}^{(j)}, \tag{2.1}$$

$i, j = 1, 2$. Then the vector $\tilde{\boldsymbol{Z}}_n$ can be rewritten as

$$\tilde{\boldsymbol{Z}}_n = \tilde{\boldsymbol{Z}}_{n-1}\tilde{M} + \tilde{\boldsymbol{\Lambda}} + \tilde{\boldsymbol{\epsilon}}_n \tag{2.2}$$

where $\tilde{\boldsymbol{\Lambda}}$ is as defined above, $\tilde{\boldsymbol{\epsilon}}_n = (\epsilon_{n,1}^{(1)}, \epsilon_{n,1}^{(2)}, \epsilon_{n,2}^{(1)}, \epsilon_{n,2}^{(2)})$ with

$$\epsilon_{n,i}^{(j)} = Z_{n,i}^{(j)} - m_{ij}(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)}) - \lambda_{ij}, \tag{2.3}$$

for $i, j = 1, 2$ and $\tilde{M}$ is defined by

$$\tilde{M} = \begin{bmatrix} m_{11} & m_{12} & 0 & 0 \\ 0 & 0 & m_{21} & m_{22} \\ m_{11} & m_{12} & 0 & 0 \\ 0 & 0 & m_{21} & m_{22} \end{bmatrix}. \tag{2.4}$$

Let $\mathcal{F}_n$ denote a σ-algebra defined by

$$\mathcal{F}_n = \sigma\{\xi_{l-1,i,k}^{(j)}, Y_{l,i}^{(j)}; i, j = 1, 2, 1 \le l \le n, k \ge 1\}. \tag{2.5}$$

Then, $\{\tilde{\boldsymbol{\epsilon}}_n\}$ is a 4-dimensional martingale difference sequence, that is, $E\{\tilde{\boldsymbol{\epsilon}}_n | \mathcal{F}_{n-1}\} = \mathbf{0}$ almost surely (a.s.). The conditional variance-covariance matrix of $\tilde{\boldsymbol{\epsilon}}_n$ is

$$Var(\tilde{\boldsymbol{\epsilon}}_n | \mathcal{F}_{n-1}) = \begin{bmatrix} e_{11} & e_{1,12} & 0 & 0 \\ e_{1,12} & e_{12} & 0 & 0 \\ 0 & 0 & e_{21} & e_{2,12} \\ 0 & 0 & e_{2,12} & e_{22} \end{bmatrix}, \tag{2.6}$$

where $e_{ij} = \sigma_{ij}^2(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)}) + d_{ij}^2, e_{i,12} = \sigma_{i,12}(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)}) + d_{i,12}$, $i, j = 1, 2$, and σ_{ij}^2, $\sigma_{i,12}$, d_{ij}^2 and $d_{i,12}$ are variance and covariance elements of Σ_i and Δ_i, respectively, defined above. Note that $Var(\tilde{\boldsymbol{\epsilon}}_n | \mathcal{F}_{n-1})$

is unbounded if $\tilde{\boldsymbol{Z}}_n \to \infty$ a.s. as $n \to \infty$. In such cases the classical Gauss-Markov theorem suggests that the weighted conditional least squares approach may bring some improvement over the conditional least squares approach.

To construct weighted conditional least squares estimators, define a diagonal matrix

$$B_n = diag((Z_{n-1,1}^{(1)} + Z_{n-1,2}^{(1)} + 1)^{-1/2}, (Z_{n-1,1}^{(1)} + Z_{n-1,2}^{(1)} + 1)^{-1/2},$$
$$(Z_{n-1,1}^{(2)} + Z_{n-1,2}^{(2)} + 1)^{-1/2}, (Z_{n-1,1}^{(2)} + Z_{n-1,2}^{(2)} + 1)^{-1/2}). \tag{2.7}$$

Post-multiplying (2.2) by B_n on both sides we can rewrite each component in (2.2) as

$$\frac{Z_{n,i}^{(j)}}{(Z_{n-1,1}^{(i)} + +Z_{n-1,2}^{(i)} + 1)^{1/2}} = m_{ij}(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)} + 1)^{1/2}$$
$$+ (\lambda_{ij} - m_{ij})(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)} + 1)^{-1/2} + \delta_{n,i}^{(j)} \tag{2.8}$$

for $i, j = 1, 2$, where $\delta_{n,i}^{(j)} = \epsilon_{n,i}^{(j)}/(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)} + 1)^{1/2}$ with $\epsilon_{n,i}^{(j)}$ as defined in (2.3). Let $\boldsymbol{\delta}_n = (\delta_{n,1}^{(1)}, \delta_{n,1}^{(2)}, \delta_{n,2}^{(1)}, \delta_{n,2}^{(2)})$. Then, we have $E(\boldsymbol{\delta}_n|\mathcal{F}_{n-1}) = \mathbf{0}$ and from (2.6)

$$Var(\boldsymbol{\delta}_n|\mathcal{F}_{n-1}) = \begin{bmatrix} e_{11}^* & e_{1,12}^* & 0 & 0 \\ e_{1,12}^* & e_{12}^* & 0 & 0 \\ 0 & 0 & e_{21}^* & e_{2,12}^* \\ 0 & 0 & e_{2,12}^* & e_{22}^* \end{bmatrix} \tag{2.9}$$

where $e_{ij}^* = e_{ij}/(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)} + 1)$ and $e_{i,12}^* = e_{i,12}/(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)} + 1)$, $i, j = 1, 2$. Note that the elements of the matrix $Var(\boldsymbol{\delta}_n|\mathcal{F}_{n-1})$ stays bounded in n as $\tilde{\boldsymbol{Z}}_n \to \infty$. This implies that the conditional variance of the "error" term $\boldsymbol{\delta}_n$ would not fluctuate too much even if $\tilde{\boldsymbol{Z}}_n$ is unbounded.

These considerations lead us to study the weighted conditional least squares estimators of $\tilde{M}$ and $\tilde{\Lambda}$ in (2.2) which are obtained by minimizing $\sum_{i=1}^n \boldsymbol{\delta}_i \cdot \boldsymbol{\delta}_i$, where '.' denotes the dot product. The weighted conditional least squares estimators of m_{ij} and λ_{ij} are given by

$$\hat{m}_{ij} = \left\{ \sum_{k=1}^n Z_{k,i}^{(j)} \sum_{k=1}^n \frac{1}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} - n \sum_{k=1}^n \frac{Z_{k,i}^{(j)}}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} \right\}$$
$$\times \left\{ \sum_{k=1}^n (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \sum_{k=1}^n \frac{1}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} - n^2 \right\}^{-1} \tag{2.10}$$

and

$$\hat{\lambda}_{ij} = \left\{ \sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)}) \sum_{k=1}^{n} \frac{Z_{k,i}^{(j)}}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} - \sum_{k=1}^{n} Z_{k,i}^{(j)} \sum_{k=1}^{n} \frac{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)})}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} \right\}$$
$$\times \left\{ \sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \sum_{k=1}^{n} \frac{1}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} - n^2 \right\}^{-1}, \quad (2.11)$$

for $i, j = 1, 2$. Here, $\hat{\lambda}_{ij} = \hat{m}_{ij} + \hat{p}_{ij}$, where $\hat{p}_{ij}$ is the least squares estimator of $p_{ij} = (\lambda_{ij} - m_{ij})$ in (2.8). ¿From these, the weighted conditional least squares estimator of the mean matrix M is given by

$$\hat{M} = \begin{pmatrix} \hat{m}_{11} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{22} \end{pmatrix} \quad (2.12)$$

where $\hat{m}_{ij}$ are as defined in (2.10) and the weighted conditional least squares estimator of the maximal eigenvalue ρ of M is given by $\hat{\rho}$, the maximal eigenvalue of $\hat{M}$ defined in (2.12). Next, we state some asymptotic properties of $\hat{M}$ and $\hat{\rho}$ defined above.

In the rest of the paper, it is assumed that the mean matrix M is positive, that is, $m_{ij} > 0$ for all i and j (This is equivalent to assuming that $\tilde{M}$ in (2.4) is positively regular), the maximal eigenvalue $\rho > 1$ and that the random variable $\tilde{W}$ in Proposition A.1 is such that $P[\tilde{W} = 0] = 0$ (see appendix). The last assumption is merely a notational convenience in order to avoid making trivial exceptions on the set of non-extinction. Also, we do not discuss below the asymptotic properties of $\hat{\lambda}_{ij}$ defined above because we conjecture that $\hat{\lambda}_{ij}$'s are not even weakly consistent estimators of λ_{ij}. This conjecture is motivated by the results obtained by Wei and Winnicki (1990, see Section 3) for supercritical single-type branching processes. Issues such as non-existence of consistent estimators for multitype branching processes will be discussed elsewhere.

Before we state the main theorem we introduce some notations. Define a block diagonal matrix

$$D = \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{bmatrix} \quad (2.13)$$

where Σ_i, $i = 1, 2$, are the variance-covariance matrices defined in Section 1, and $\mathbf{0}$ is a 2×2 zero-matrix. Note that the maximal eigenvalue ρ of M is given by

$$\rho = [m_1 + m_{22} + \{(m_{11} - m_{22})^2 + 4m_{12}m_{21}\}^{1/2}]/2 \tag{2.14}$$

for the two-type branching processes. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be the right and left eigenvectors corresponding to ρ which are normalized so that

$$\boldsymbol{u} \cdot \boldsymbol{v} = 1 = \boldsymbol{u} \cdot \mathbf{1} = |\boldsymbol{u}|, \tag{2.15}$$

where $\mathbf{1} = (1, \ldots, 1)$. Let

$$V = \sum_{i=1}^{2} (v^i)^{-1} \left[\left(\frac{\partial \rho}{\partial m_{i1}} \right)^2 \sigma_{i1}^2 + 2 \left(\frac{\partial \rho}{\partial m_{i1}} \frac{\partial \rho}{\partial m_{i2}} \right) \sigma_{i,12} + \left(\frac{\partial \rho}{\partial m_{i2}} \right)^2 \sigma_{i2}^2 \right] \tag{2.16}$$

where ∂ denotes partial derivative, v^i is the i-th component of the vector $\boldsymbol{v}$, and σ_{i1}^2 and $\sigma_{i,12}$ are the elements of Σ_i defined in Section 1. Finally, for $i, j = 1, 2$ let

$$X_{n,i}^{(j)} = \left[\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \right]^{1/2} (\hat{m}_{ij} - m_{ij}) \tag{2.17}$$

for $\hat{m}_{ij}$ defined in (2.10) and

$$\boldsymbol{X}_n = (X_{n,1}^{(1)}, X_{n,1}^{(2)}, X_{n,2}^{(1)}, X_{n,2}^{(2)}). \tag{2.18}$$

THEOREM. *For the weighted conditional least squares estimators $\hat{M}$ and $\hat{\rho}$ defined above the following results hold as $n \to \infty$:*

(a) $\hat{M} \to M$ a.s. and $\hat{\rho} \to \rho$ a.s.,

(b) $\boldsymbol{X}_n \xrightarrow{D} N_4(\mathbf{0}, D)$,

and

(c) For $\tilde{W}$ defined in Proposition A.1,

$$[\tilde{W}(1 + \rho + \ldots + \rho^{n-1})]^{1/2} (\hat{\rho} - \rho) \xrightarrow{D} N(0, V).$$

Consequently, for $i = 1, 2$

(d) $$\left[\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)/v^i \right]^{1/2} (\hat{\rho} - \rho) \xrightarrow{D} N(0, V).$$

Proof of the above Theorem depends on some asymptotic properties of the process $\{\tilde{\boldsymbol{Z}}_n\}$ and these are proved in the Appendix. The Theorem is proved in the next Section.

By (3.9) and (3.10) below, and assertion (d) of Theorem we have that, in terms of asymptotic variance, the weighted conditional least squares estimators of m_{ij} and ρ based on $\{Z_{k,i}^{(j)}, 1 \leq k \leq n, i, j = 1,2\}$ are as good as the m.l. estimators of the same obtained in Asmussen and Keiding (1978) [see Theorems 5.1 and Corollary 5.1 of Asmussen and Keiding (1978)] under full family tree information.

3. Proof

Proof of the Theorem. The strong consistency of $\hat{M}$ follows from the results in Corollary A.2. From this and (2.14) we have that $\hat{\rho}$ is also strongly consistent. This proves assertion (a). Next, note that $\hat{\rho}$ is a smooth function of $\hat{m}_{ij}$, $i, j = 1, 2$. Therefore, by continuous mapping theorem it is enough to establish the asymptotic normality of $\boldsymbol{X}_n$, which is assertion (b).

Now, for assertion (b), substitute (2.3) into (2.10) and use algebraic manipulations to rewrite $X_{n,i}^{(j)}$ defined in (2.17) as

$$X_{n,i}^{(j)} = (C_{n,i}^{(j)} - E_{n,i}^{(j)})/V_{n,i}, \quad i, j = 1, 2, \tag{3.1}$$

where $C_{n,i}^{(j)}$ is as defined before Lemma A.3 below,

$$E_{n,i}^{(j)} = \left\{ n \sum_{k=1}^{n} \epsilon_{k,i}^{(j)} / (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \right\} \times$$

$$\left\{ \sum_{k=1}^{n} 1/(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) [\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)]^{1/2} \right\}^{-1} \tag{3.2}$$

and

$$V_{n,i} = 1 - n^2 \left[\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \sum_{k=1}^{n} 1/(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \right]^{-1} . \tag{3.3}$$

¿From these it is easy to see that

$$\boldsymbol{X}_n = (\boldsymbol{C}_n - \boldsymbol{E}_n) \boldsymbol{V}_n^{-1} \tag{3.4}$$

where $\boldsymbol{E}_n = (E_{n,1}^{(1)}, E_{n,1}^{(2)}, E_{n,2}^{(1)}, E_{n,2}^{(2)})$ and $\boldsymbol{V}_n = diag(V_{n,1}, V_{n,1}, V_{n,2}, V_{n,2})$ is a diagonal matrix. By Corollary A.2 we have that $\boldsymbol{V}_n \to \boldsymbol{I}$ a.s. as $n \to \infty$

where $\boldsymbol{I}$ is an identity matrix. By Lemma A.3 (iii), $\boldsymbol{C}_n \xrightarrow{D} N_4(\mathbf{0}, D)$ as $n \to \infty$. Therefore, it only remains to show that

$$\boldsymbol{E}_n \to \mathbf{0} \ \text{ a.s. as } \ n \to \infty. \tag{3.5}$$

To show this, first note that for each $i, j = 1, 2$, $\{\epsilon_{n,i}^{(j)}/(Z_{n-1,1}^{(i)} + Z_{n-1,2}^{(i)} + 1)\}$ is a martingale difference sequence with respect to $\mathcal{F}_n$ defined in (2.5). Moreover, as seen below (2.6),

$$\sum_{k=1}^{\infty} E\{\epsilon_{k,i}^{(j)^2}/(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)^2 \Big| \mathcal{F}_{i-1}\}$$

$$= \sum_{k=1}^{\infty} \{\sigma_{ij}^2 (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)}) + d_{ij}^2\}/(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)^2 < \infty \ \text{ a.s.} \tag{3.6}$$

where we used Corollary A.2. Therefore, by the martingale convergence theorem [see Hall and Heyde (1980), Theorem 2.17] we have that

$$\sum_{k=1}^{n} \epsilon_{k,i}^{(j)}/(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \ \text{ converges a.s.,} \tag{3.7}$$

as $n \to \infty$. Hence, from Corollary A.2 and (3.7) we have that $E_{n,i}^{(j)} \to 0$ a.s. as $n \to \infty$ for each $i, j = 1, 2$. The result in (3.5) now follows. Assertion (b) now follows from (3.4) and (3.5).

As for assertion (c), first note from Corollary A.2 that for each $i = 1, 2$

$$\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) / \sum_{k=1}^{n} \rho^{k-1} \to \tilde{W}(v_1^i + v_2^i) \ \text{ a.s.} \tag{3.8}$$

as $n \to \infty$. This, together with assertion (b), implies that

$$\tilde{\boldsymbol{X}}_n = (\tilde{X}_{n,1}^{(1)}, \tilde{X}_{n,1}^{(2)}, \tilde{X}_{n,2}^{(1)}, \tilde{X}_{n,2}^{(2)}) \xrightarrow{D} N_4(\mathbf{0}, \tilde{D}) \tag{3.9}$$

as $n \to \infty$, where

$$\tilde{X}_{n,i}^{(j)} = [\tilde{W} \sum_{k=1}^{n} \rho^{k-1}]^{1/2} (\hat{m}_{ij} - m_{ij})$$

and

$$\tilde{D} = \begin{bmatrix} (v_1^1 + v_2^1)^{-1}\Sigma_1 & 0 \\ 0 & (v_1^2 + v_2^2)^{-1}\Sigma_2 \end{bmatrix}$$

with Σ_1 and Σ_2 as defined in D (see (2.13)). Now, since $\hat{\rho}$ is a smooth function of $\hat{m}_{ij}$ (see (2.14)), apply Theorem 3.3.A. of Serfling (1980) to the sequence $\tilde{\boldsymbol{X}}_n$ in (3.9) to get the required result.

Assertion (d) follows from (3.8) and the fact that the left eigenvector $\boldsymbol{v} = (v^1, v^2)$ of M defined in (2.15) and the left eigenvector $\tilde{\boldsymbol{v}} = (v_1^1, v_1^2, v_2^1, v_2^2)$ of $\tilde{M}$ defined in the Appendix satisfy the following:

$$v^i = v_1^i + v_2^i \quad \text{for} \quad i = 1, 2. \tag{3.10}$$

Hence the Theorem. □

4. Appendix

Here we establish some asymptotic properties of the process $\{\tilde{\boldsymbol{Z}}_n\}$ defined in (2.2). These results are crucial for the proof of the main theorem.

Let $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ denote the right and left eigenvectors of $\tilde{M}$ corresponding to the maximal eigenvalue ρ such that $\tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{v}} = 1 = \tilde{\boldsymbol{u}} \cdot \mathbf{1}$. The following proposition concerns the growth rate of $\tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{Z}}_n$ and $\tilde{\boldsymbol{Z}}_n$ for $\tilde{\boldsymbol{Z}}_n$ defined in (2.2). These results are exactly the same as those available in the literature for $\boldsymbol{u} \cdot \boldsymbol{Z}_n$ and $\boldsymbol{Z}_n$ for $\boldsymbol{Z}_n$ defined in (1.1); see Athreya and Ney (1972), for instance.

Proposition A.1. *Let $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ be as defined above. Then there exists a nonnegative random variable $\tilde{W}$ such that*

$$\lim_{n\to\infty} \rho^{-n} \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{Z}}_n = \tilde{W} \quad a.s. \,. \tag{A.1}$$

Also, on the set $\{\tilde{\boldsymbol{Z}}_n \to \infty \quad as \quad n \to \infty\}$,

$$\lim_{n\to\infty} \tilde{\boldsymbol{Z}}_n / \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{Z}}_n = \tilde{\boldsymbol{v}} \quad a.s. \tag{A.2}$$

and

$$\lim_{n\to\infty} \rho^{-n} \tilde{\boldsymbol{Z}}_n = \tilde{W} \tilde{\boldsymbol{v}} \quad a.s. \tag{A.3}$$

where $P[\tilde{W} > 0] > 0$ if and only if $E\{\xi_{1,i,1}^{(j)} \log \xi_{1,i,1}^{(j)}\} < \infty$ for each $i, j = 1, 2$.

Proof. Let $\tilde{W}_n = \rho^{-n} \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{Z}}_n$. Then, from (2.2) and (2.5) it follows that $E\{\tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{\epsilon}}_{n+1} | \mathcal{F}_n\} = 0$. From this we have that

$$E\{\tilde{W}_{n+1} | \mathcal{F}_n\} = \rho^{-(n+1)} \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{Z}}_n \tilde{M} + \rho^{-(n+1)} \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{\Lambda}} = W_n + \zeta_n,$$

where we used the fact that $\tilde{M}\tilde{\boldsymbol{u}} = \rho\tilde{\boldsymbol{u}}$ and let $\zeta_n = \rho^{-(n+1)}\tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{\Lambda}}$. Therefore, $\{W_n; \mathcal{F}_n\}$ is an "almost supermartingale" sequence (see Robbins and Siegmund (1971) for a definition). Then, by Theorem 1 of Robbins and Siegmund (1971), $\lim_{n\to\infty} W_n = \tilde{W}$, for some r.v. $\tilde{W} < \infty$ on the set $\{\sum_{n=1}^{\infty} \zeta_n < \infty\}$. Since $\rho > 1$, $\sum_{n=1}^{\infty} \rho^{-(n+1)}\tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{\Lambda}} < \infty$ follows trivially. Hence the assertion (A.1).

That convergence in (A.2) holds in probability can be proved using arguments similar to Theorem V.6.3, of Athreya and Ney (1972). For the required a.s. convergence in (A.2) see Remark 1 of section V.6 of Athreya and Ney (1972). Assertion (A.3) follows from (A.1) and (A.2). Finally, the "if and only if" assertion in (A.3) follows from arguments similar to Theorem 8.2 of Mode (1971). □

The following Corollary follows from Proposition A.1 and the fact that ordinary convergence implies Césaro convergence.

COROLLARY A.2. *Under the conditions of Proposition A.1, the following results hold almost surely as $n \to \infty$: For each $i, j = 1, 2$*

$$\sum_{k=1}^{\infty} \frac{1}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)} < \infty$$

$$\rho^{-n} \sum_{k=1}^{n} Z_{k,i}^{(j)} \to [\rho/(\rho - 1)]\tilde{W} v_i^j$$

$$\rho^{-n} \sum_{k=1}^{n} Z_{k-1,i}^{(j)} \to (\rho - 1)^{-1}\tilde{W} v_i^j$$

$$n^{-1} \sum_{k=1}^{n} Z_{k,i}^{(j)}/(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \to \rho v_i^j/(v_1^i + v_2^i)$$

$$n^2 \left[\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1) \sum_{k=1}^{n} \frac{1}{(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)}\right]^{-1} \to 0.$$

The next lemma is crucial for the proof of the main theorem. The proof of the results in the lemma are very similar to those in Section 3 of Wei and Winnicki (1989). Therefore, our proofs will be brief. Before we prove the lemma we will introduce some notations.

Let $(\boldsymbol{R}^4)^{\infty}$ denote the space of real valued vector sequences $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots)$ with the metric $d(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{\infty} 2^{-i} ||\boldsymbol{x}_i - \boldsymbol{y}_i||/(1 + ||\boldsymbol{x}_i - \boldsymbol{y}_i||)$. Let $\mathcal{B}$ be the

Borel sigma algebra generated by d. Consider the random elements $\Gamma_n = \{\boldsymbol{\gamma}_{nk}\}$ and $\Gamma = \{\boldsymbol{\gamma}_k\}$ on $((\boldsymbol{R}^4)^\infty, \mathcal{B})$, where $\boldsymbol{\gamma}_{nk} = (\gamma_{nk,1}^{(1)}, \gamma_{nk,1}^{(2)}, \gamma_{nk,2}^{(1)}, \gamma_{nk,2}^{(2)})$ are defined by

$$\gamma_{nk,i}^{(j)} = \frac{(Z_{n-k+1,i}^{(j)} - m_{ij}[Z_{n-k,1}^{(i)} + Z_{n-k,2}^{(i)}] - Y_{n-k+1,i}^{(j)})}{(Z_{n-k,1}^{(i)} + Z_{n-k,2}^{(i)} + 1)^{1/2}} I_{\{1 \le k \le n\}}, \tag{A.4}$$

for $i, j = 1, 2$ and $\{\boldsymbol{\gamma}_k\}$ is a sequence of i.i.d. random vectors distributed as $N_4(0, D)$ for D defined in (2.13). Define a vector $\boldsymbol{C}_n = (C_{n,1}^{(1)}, C_{n,1}^{(2)}, C_{n,2}^{(1)}, C_{n,2}^{(2)})$ where

$$C_{n,i}^{(j)} = \sum_{k=1}^{n} \epsilon_{k,i}^{(j)} / [\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)]^{1/2},$$

for $i, j = 1, 2$ and $\epsilon_{k,i}^{(j)}$ is as defined in (2.3).

LEMMA A.3. *The following results hold as $n \to \infty$:*

(i) $\Gamma_n \to \Gamma$ *weakly in* $((\boldsymbol{R}^4)^\infty, \mathcal{B})$,

(ii) For a real sequence $\{c_n\}$ such that $\sum_{i=1}^{\infty} |c_i| < \infty$

$$\sum_{k=1}^{n} c_k \boldsymbol{\gamma}_{nk} \xrightarrow{D} N_4\left(\mathbf{0}, D \sum_{k=1}^{\infty} c_k^2\right),$$

where $\{\boldsymbol{\gamma}_{nk}\}$ is as defined in (A.4) and D is as in (2.13). For $\boldsymbol{C}_n$ defined above,

(iii) $\boldsymbol{C}_n \xrightarrow{D} N_4(\mathbf{0}, D)$.

PROOF. It is known that on $((\boldsymbol{R}^4)^\infty, \mathcal{B})$, $\Gamma_n \to \Gamma$ weakly is equivalent to the weak convergence of $\{\boldsymbol{\gamma}_{nk}, 1 \le k \le r\}$ for all integers $r \ge 1$ (see Billingsley (1968), p. 19). The arguments below are similar to those in Theorem 1 of Heyde and Brown (1971). For $r \ge 1$ and $(\boldsymbol{\theta}_1, \ldots \boldsymbol{\theta}_r)$, where $\boldsymbol{\theta}_k = (\theta_{k,1}^{(1)}, \theta_{k,1}^{(2)}, \theta_{k,2}^{(1)}, \theta_{k,2}^{(2)})$ consider

$$\phi_n(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_r) = E\left[e^{i \sum_{k=1}^{r} \boldsymbol{\theta}_k \cdot \boldsymbol{\gamma}_{nk}}\right] = E\left[\prod_{k=1}^{r} A_{nk}\right] \tag{A.5}$$

where $A_{nk} = e^{i \boldsymbol{\theta}_k \cdot \boldsymbol{\gamma}_{nk}}$. Next, define

$$R_r^{(n)} = \prod_{k=1}^{r} A_{nk} - \prod_{k=1}^{r} e^{-1/2 \boldsymbol{\theta}_k D \boldsymbol{\theta}_k'}. \tag{A.6}$$

Now use arguments very similar to those in the proof of Theorem 1 of Heyde and Brown (1971) and (A.3) to conclude that $ER_r^{(n)} \to 0$ as $n \to \infty$. ¿From this and (A.5) we have that

$$\phi_n(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_r) \to e^{-1/2 \sum_{k=1}^{r} \boldsymbol{\theta}_k D \boldsymbol{\theta}'_k} \quad \text{as} \quad n \to \infty.$$

This proves assertion (i).

As for assertion (ii), for each r, define $\boldsymbol{U}_{rn} = \sum_{k=1}^{r} c_k \boldsymbol{\gamma}_{nk}$, $\boldsymbol{U}_r = \sum_{k=1}^{r} c_k \boldsymbol{\gamma}_k$ and $\boldsymbol{U} = \sum_{k=1}^{\infty} c_k \boldsymbol{\gamma}_k$ where $\{\boldsymbol{\gamma}_k\}$ is the sequence of i.i.d. N_4 $(\mathbf{0},\, D)$ vectors defined below (A.4). By part (i) of the Lemma, for each r, $\boldsymbol{U}_{rn} \xrightarrow{D} \boldsymbol{U}_r$ as $n \to \infty$. Also, $\boldsymbol{U}_r \xrightarrow{D} \boldsymbol{U}$ as $r \to \infty$, where $\boldsymbol{U}$ is N_4 $(\mathbf{0},\, D \sum_{k=1}^{\infty} c_k^2)$. Note from (A.4) that for $r \geq n$ we have that $\boldsymbol{U}_{rn} = \boldsymbol{U}_{nn}$ and for $1 \leq r \leq n$

$$\begin{aligned} P\{||\boldsymbol{U}_{rn} - \boldsymbol{U}_{nn}|| > \eta\} &\leq \eta^{-2} E||\boldsymbol{U}_{rn} - \boldsymbol{U}_{nn}||^2 \\ &= \eta^{-2} \sum_{i=1}^{2} \sum_{j=1}^{2} E\left(\sum_{k=r+1}^{n} c_k \gamma_{nk,i}^{(j)}\right)^2 \leq \eta^{-2} \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{ij}^2 \sum_{k=r+1}^{\infty} c_k^2. \quad \text{(A.7)} \end{aligned}$$

Hence, $\lim_{r\to\infty} \limsup_{n\to\infty} P\{||\boldsymbol{U}_{rn} - \boldsymbol{U}_{nn}|| > \eta\} = 0$. Therefore, by a result of Billinglsey (1968), p. 28, we have that $\boldsymbol{U}_{nn} \xrightarrow{D} \boldsymbol{U}$ as $n \to \infty$. Hence the assertion (ii).

As for assertion (iii), let

$$H_{n,i}^{(j)} = \sum_{k=1}^{n} [\rho^{(k-1)} \tilde{W}(v_1^i + v_2^i)]^{1/2} \epsilon_{k,i}^{(j)} / (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)^{1/2} \qquad \text{(A.8)}$$

and

$$\begin{aligned} L_{n,i}^{(j)} &= \sum_{k=1}^{n} [(Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)^{1/2} - (\rho^{(k-1)} \tilde{W}(v_1^i + v_2^i))^{1/2}] \epsilon_{k,i}^{(j)} \\ &\qquad / (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)^{1/2}. \end{aligned} \qquad \text{(A.9)}$$

Then, it is easy to check that

$$C_{n,i}^{(j)} = \{H_{n,i}^{(j)} + L_{n,i}^{(j)}\} / \left[\sum_{k=1}^{n} (Z_{k-1,1}^{(i)} + Z_{k-1,2}^{(i)} + 1)\right]^{1/2} \qquad \text{(A.10)}$$

for $i, j = 1, 2$. Set $\boldsymbol{H}_n = (H_{n,1}^{(1)}, H_{n,1}^{(2)}, H_{n,2}^{(1)}, H_{n,2}^{(2)})$ and $\boldsymbol{L}_n = (L_{n,1}^{(1)}, L_{n,1}^{(2)}, L_{n,2}^{(1)}, L_{n,2}^{(2)})$. In view of Corollary A.2 and from (A.10) it suffices to show that

$$\boldsymbol{L}_n = o_p(\rho^{n/2}) \quad \text{as} \quad n \to \infty \qquad \text{(A.11)}$$

and the vector $\tilde{\boldsymbol{H}}_n = (\tilde{H}^{(1)}_{n,1}, \tilde{H}^{(2)}_{n,1}, \tilde{H}^{(1)}_{n,2}, \tilde{H}^{(2)}_{n,2})$ where

$$\tilde{H}^{(j)}_{n,i} = (\rho-1)^{1/2}\sum_{k=1}^{n}\rho^{-(n-k+1)/2}\epsilon^{(j)}_{k,i}/(Z^{(i)}_{k-1,1}+Z^{(i)}_{k-1,2}+1)^{1/2}$$

is such that

$$\tilde{\boldsymbol{H}}_n \xrightarrow{D} N_4(\mathbf{0}, D) \quad \text{as} \quad n\to\infty. \tag{A.12}$$

Since $E|Y^{(j)}_{n,i} - \lambda_{ij}| = o(\rho^{n/2})$ we have from assertin (ii) above that (A.12) holds.

As for (A.11), define

$$D^{(i)}_n = \sum_{k=1}^{n}\rho^{(k-1)/2}[\{(Z^{(i)}_{k-1,1}+Z^{(i)}_{k-1,2}+1)/\rho^{(k-1)}\}^{1/2} - \{\tilde{W}(v^i_1+v^i_2)\}^{1/2}]^2$$

and

$$G^{(j)}_{n,i} = \sum_{k=1}^{n}\rho^{(k-1)/2}[\epsilon^{(j)}_{k,i}/(Z^{(i)}_{k-1,1}+Z^{(i)}_{k-1,2}+1)^{1/2}]^2.$$

Then, by (A.3) we have that $D^{(i)}_n = o(\rho^{n/2})$ a.s. and $E[G^{(j)}_{n,i}] = O(\rho^{n/2})$ which implies that $G^{(j)}_{n,i} = O_p(\rho^{n/2})$ for $i,j=1,2$. Therefore, by the Cauchy-Schwarz inequality $|L^{(j)}_{n,i}| \le (D^{(i)}_n)^{1/2}(G^{(j)}_{n,i})^{1/2} = o_p(\rho^{n/2})$ as $n\to\infty$. Hence we have (A.11). The Lemma now follows from the above arguments. □

Acknowledgements. The authors would like to thank the referee for a careful reading of the manuscript. Referee's suggestions led to a significant improvement in the presentation of the contents.

References

ASMUSSEN, S. and KEIDING, N. (1978). Martingale central limit theorems and asymptotic estimation theory for multitype branching processes. *Adv. Appl. Prob.* **10**, 109-129.

ATHREYA, K.B. and NEY, P.E. (1972). *Branching Processes.* Springer, Berlin.

BADALBAEV, I.S. (1976). A certain estimate for the eigenvalue of the mean value matrix of a branching random process with two types of particles. *Izv. Akad. Nauk USSR Ser. Fiz.-Mat. Nauk* **5**, 8-13 (in Russian).

BADALBAEV, I.S. and MUKHITDINOV, A. (1989). Limit theorems for some functionals in critical multitype branching processes. *Theory Prob. Appl.* **34**, 129-133.

BECKER, N. (1977). Estimation for discrete-time branching processes with applications to epidemics. *Biometrics* **33**, 515-522.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

DION, J-P. (1993). *Statistical Inference for Discrete Time Branching Processes.* Lecture notes, Universite du Quebec a Montreal.

GUTTORP, P. and SIEGEL, A.F. (1985). Consistent estimation in partially observed random walks. *Ann. Statist.* **13**, 958-969.

HALL, P. and HEYDE, C.C. (1980). *Martingale Limit Theory and its Applications.* New York, Academic Press.

HARRIS, T.E. (1963). *The Theory of Branching Processes.* Springer, Berlin.

HEYDE, C.C. and BROWN, B.M. (1971). An invariance principle and some convergence rate results for branching processes. *Z. Wahrsch. Verw. Geb* **20**, 271-278.

JOFFE, A. and METIVIER, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. Appl. Prob.* **18**, 20-65.

KEIDING, N. and LAURITZEN, S. (1978). Marginal maximum likelihood estimates and estimation of the offspring mean in a branching process. *Scand. J. Statist.* **5**, 106-110.

KESTEN, H. and STIGUM, B.P. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.* **37**, 1211-1223.

LOCKHART, R.A. (1982). On the nonexistence of consistent estimates in Galton-Watson branching processes. *J. Appl. Prob.* **19**, 842-846.

LÜTKEPOHL, H. (1996). *Handbook of Matrices.* Wiley, New York.

MODE, C.J. (1971). *Multitype Branching Processes: Theory and Applications*, American Elsevier, New York.

NANTHI, K. (1978). Estimation of the offspring mean/growth rate in a supercritical Galton-Watson process/multitype Galton-Watson process. *J. Madras Univ. B* **41**, 3, 25-28.

NANTHI, K. (1980). Some limit theorems on Galton-Watson type branching processes. *Gujarat. Statist. Rev.* **7**, 1-6.

QUINE, M.P. and DURHAM, P. (1977). Estimation for mulitype branching processes. *J. Appl. Prob.* **14**, 829-835.

ROBBINS, H.E. and SIEGMUND, D. (1971). A convergence theorem for nonnegative almost super martingales and some applications. In *Optimizing Methods in Statistics*, J.S. Rustagi, ed., Academic, New York, 223-258.

SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

SHETE, S. (1998). Estimation problems in single and multitype branching processes and in physical mapping of a chromosome. Ph.D. Thesis, Department of Statistics, The Univ. of Georgia.

WEI, C.Z. and WINNICKI, J. (1989). Some asymptotic results for branching processes with immigration. *Stoch. Proc. Appl.* **31**, 261-282.

WEI, C.Z. and WINNICKI, J. (1990). Estimation of the means in the branching process with immigration. *Ann. Statist.* **18**, 1757-1773.

SANJAY SHETE
DEPARTMENT OF EPIDEMIOLOGY
THE UNIVERSITY OF TEXAS
M.D. ANDERSON CANCER CENTER
HOUSTON, TEXAS 77030
E-mail: sshete@mail.mdanderson.org

T.N. SRIRAM
DEPARTMENT OF STATISTICS
UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602
E-mail: tn@stat.uga.edu