Søren Asmussen

On the Role of a Certain Eigenvalue in Estimating the Growth Rate of a Branching Process

Preprint
January
1981
2

Institute of Mathematical Statistics
University of Copenhagen

Søren Asmussen*

ON THE ROLE OF A CERTAIN EIGENVALUE IN ESTIMATING THE GROWTH RATE OF A BRANCHING PROCESS

Preprint 1981 No. 2

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

January 1981

Summary

In many situations, the data given on a p-type Galton-Watson process $\underline{Z}_n \in \mathbb{N}^p$ will consist of the total generation sizes $|\underline{Z}_n|$ only. In that case, the maximum likelihood estimator $\hat{\rho}_{ML}$ of the growth rate ρ is not observable, and the asymptotic properties of the most obvious estimators of ρ based on the $|\underline{Z}_n|$, as studied by Asmussen and Keiding (1978), show a crucial dependence on $|\rho_1|$, ρ_1 being a certain other eigenvalue of the offspring mean matrix. In fact, if $|\rho_1|^2 \geq \rho$, then the speed of convergence compares badly with $\hat{\rho}_{ML}$. In the present note, it is pointed out that recent results of Heyde (1981b) on so-called Fibonacci branching processes provide further examples of this phenomenon, and an estimator with the same speed of convergence as $\hat{\rho}_{ML}$ and based on the $|\underline{Z}_n|$ alone is exhibited for the case $p = 2$, $\rho_1^2 \geq \rho$.

1. Introduction

Consider a two-type Galton-Watson process $\underline{Z}_n = (Z_n(1)\ Z_n(2))$, i.e.

$$\underline{Z}_{n+1} = \sum_{i=1}^{2} \sum_{k=1}^{Z_n(i)} \underline{Z}_{n,k}^{(i)} \tag{1.1}$$

with the $Z_{n,k}^{(i)} \in \mathbb{N}^2$ independent for all n,i,k and with the same distribution $F^{(i)}$ for fixed i. We assume throughout that the offspring mean matrix $\underline{M} = (m(i,j))$ is positively regular. It is then a well-known consequence of the Perron-Frobenius theorem that $\underline{M}$ has two real eigenvalues ρ, ρ_1 with $\rho > 0$, $\rho > |\rho_1|$ and, letting $\underline{u}, \underline{v}, \underline{u}_1, \underline{v}_1$ be the corresponding left and right eigenvectors, that $u(i) > 0$, $v(i) > 0$. Letting $\langle \cdot , \cdot \rangle$ denote inner product, we have $\langle \underline{u}_1, \underline{v} \rangle = 0$ and we normalize by $\langle \underline{u}, \underline{v} \rangle = |\underline{v}| = \langle \underline{v}, \underline{1} \rangle = 1$ so that $\underline{u} + \underline{u}_1 = \underline{1}$. We consider only the supercritical case $\rho > 1$ with finite offspring variances and extinction probability zero so that $\lim_{n\to\infty} \rho^{-n} \underline{Z}_n = W\underline{v}$ with $0 < W < \infty$ a.s.

The present note is concerned with asymptotic properties of various estimators of ρ. A systematic account was given by Asmussen and Keiding (1978), henceforth referred to as AK, who derived the maximum likelihood estimator $\hat{\rho}_{ML}$ based on the first N generation as the largest eigenvalue of the empirical mean matrix

$$\underline{\hat{M}} = (\hat{m}(i,j)) = \Big(\sum_{n=0}^{N-1} \sum_{k=1}^{Z_n(i)} Z_{n,k}^{(i)}(j) \Big/ \sum_{n=0}^{N-1} Z_n(i) \Big)$$

(see also Keiding and Lauritzen (1978)) and obtained similar asymptotic properties as for the one-type case, in particular the relation

$$\lim_{N\to\infty} P((W\rho^N)^{-\frac{1}{2}} U_N \le y) = \Phi\left(\frac{y}{\sigma}\right) \quad \text{for some} \quad 0 < \sigma^2 < \infty \tag{1.2}$$

with $U_N = W\rho^N(\hat{\rho}_{ML} - \rho)$. Based on earlier work by Becker (1977) who was concerned with the robustness of estimation procedures for the one-type case, AK also studied

$\hat{\rho}_B = (|Z_1| + \ldots + |\underline{Z}_N|)/(|\underline{Z}_0| + \ldots + |\underline{Z}_{N-1}|)$ which in the one-type case coincides with $\hat{\rho}_{ML}$, and a more complicated behaviour was discovered. In fact, the relation (1.2) with $U_N = W\rho^N(\hat{\rho}_B - \rho)$ is only valid if $\rho_1^2 < \rho$, whereas for $\rho_1^2 = \rho$ and $\rho_1^2 > \rho$ one has two other types of asymptotics, viz.

$$\lim_{N\to\infty} P((WN\rho^N)^{-\frac{1}{2}}U_N \le y) = \Phi(\frac{y}{\sigma}) \quad \text{for some} \quad 0 < \sigma^2 < \infty \tag{1.3}$$

$$\lim_{N\to\infty} \rho_1^{-N} U_N \quad \text{exists a.s.} \tag{1.4}$$

respectively. Surprising at a first look, these results turn out to be closely related to other aspects of the limit theory of the process. In fact, just the same trichotomy holds for $U_N = \langle \underline{Z}_N, \underline{u}_1 \rangle$ and if $\rho_1^2 \ge \rho$, the behaviour of $\hat{\rho}_B$ can even be directly deduced from that of such linear functionals, cf. AK Example 6.1.

We refer to Dion and Keiding (1978) for estimation theory for branching processes in general and to Kesten and Stigum (1966a,b) for background material on the process. If $\rho_1^2 > \rho$, the a.s. limit of $\rho_1^{-n} \langle \underline{Z}_n, \underline{u}_1 \rangle$ is denoted by W_1 throughout and we let E^i, Var^i etc. refer to one initial particle of type i. The proofs of AK of relations like (1.2) and in particular (1.3) are somewhat simpler than those of Kesten and Stigum (1966b) and reduce essentially to a routine check of the conditions of

Lemma 1 *Let* $\underline{a}_1, \underline{a}_2, \ldots$ *be vectors and* $\gamma_1, \gamma_2, \ldots$ *constants, such that* $0 < \gamma_n < \infty$ *and* $\{\underline{a}_n/\gamma_n\}$ *is relatively compact, and define*

$$\alpha_N^2 = \sum_{n=1}^{N} \rho^{-n}\gamma_n^2, \quad \beta_N^2 = \sum_{n=1}^{N} \rho^{-n} \sum_{i=1}^{2} v(i)\mathrm{Var}^i \langle \underline{Z}_1, \underline{a}_n \rangle$$

Then if $\gamma_N^2 = o(\rho^N \alpha_N^2)$ *and* $\varliminf_{N\to\infty} \beta_N/\alpha_N > 0$, *the limiting distribution of*

$$(W\rho^N\beta_N^2)^{-\frac{1}{2}}\sum_{n=0}^{N-1}\langle \underline{Z}_{n+1} - \underline{Z}_n M, \underline{a}_{N-n}\rangle \tag{1.5}$$

exists and is standard normal.

[Note that AK has a superfluous condition and that their argument on pg.115 contains an error; for the correction and discussion of the minimality of the conditions, see Asmussen and Hering (1980) Ch.II. Sect.3 and also Ch.VIII]. This lemma will be used also in the present note as well as

Lemma 2 *For any fixed vector* $\underline{a}$, $\langle \underline{Z}_N - \underline{Z}_{N-1}M, \underline{a}\rangle = O((\rho^N \log N)^{\frac{1}{2}})$ *a.s.*, cf. Asmussen (1977) and Heyde and Leslie (1971).

2. Remarks on a paper by Heyde

In a recent paper, Heyde (1981b) studied so-called Fibonacci branching processes, viz. (one-type) time-lagged processes $\{X_n\}$ defined recursively by

$$X_n = X_{n-1} + \sum_{k=1}^{X_{n-2}} Y_{n,k} \tag{2.1}$$

with the $Y_{n,k}$ i.i.d. With $m = EY_{n,k}$ and $\rho = (1+(1+4m)^{\frac{1}{2}})/2$, he showed the a.s. existence of $\lim_{n\to\infty} X_n/\rho^n$ as well as non-degeneracy subject to $EY_{n,k}\log Y_{n,k} < \infty$. Note that if $m = 1$, then $EX_n = EX_{n-1} + EX_{n-2}$ is the Fibonacci sequence and X_n mimics similar stochastic properties, e.g. $X_n/X_{n-1} \to (1+\sqrt{5})/2$ (the golden ratio). See also Heyde (1981a). Heyde also studied the problem of estimating ρ, to which end he considered $\hat{\rho}_{LN} = X_N/X_{N-1}$ and noted a surprising dependence of the size of ρ. In fact, if $(1-\rho)^2 < \rho$, then (1.2) holds while if $(1-\rho)^2 > \rho$, then (1.4) holds with $\rho_1 = 1-\rho$. Here $U_N = W\rho^N(\hat{\rho}_{LN} - \rho)$ and the notation refers to the names of Lotka and Nagaev connected with the one type analogue.

We should like to point out here, that the relation of these results to those for multitype Galton-Watson processes goes further than just formal similarity. In fact, define $\underline{Z}_0 = (X_0 \; 0)$, $\underline{Z}_1 = (X_1 - X_0 \; X_0)$ $\underline{Z}_n = (X_n - X_{n-1} \; X_{n-1})$. Then (1.1) holds for $n \geq 1$ with $\underline{X}^{(1)}_{n,k} = (0 \; 1)$, $\underline{X}^{(2)}_{n,k} = (Y_{n+1,k} \; 1)$ and also for all $n \geq 0$ if, as is inherent in Heyde's interpretation of his model, $X_0 = X_1$. That is, $X_n = |\underline{Z}_n|$ is simply the total population size of a multitype Galton-Watson process. The offspring mean matrix is

$$M = \begin{pmatrix} 0 & 1 \\ & \\ m & 1 \end{pmatrix} ,$$

the eigenvalues of which are indeed $\rho = (1+(1+4m)^{\frac{1}{2}})/2, \rho_1 = 1-\rho$.

For example, the remark of AK pg.125 to the effect that the analysis of $\hat{\rho}_B$ and $\hat{\rho}_{LN}$ follows similar lines, leads to the following approach to the asymptotics. Recall that $\underline{u} + \underline{u}_1 = \underline{1}$ and write

$$\hat{\rho}_{LN} - \rho = (S_N + T_N)/|\underline{Z}_{N-1}| , \qquad (2.2)$$

$$S_N = \langle \underline{Z}_N - \underline{Z}_{N-1}\underline{M}, \underline{1} \rangle , \; T_N = (\rho_1 - \rho)\langle \underline{Z}_{N-1} , \underline{u}_1 \rangle$$

Here the a.s. magnitude of S_N is given by Lemma 2, while T_N is a linear functional referred to in Section 1. If $\rho_1^2 \geq \rho$, it follows that T_N dominates S_N so that the behaviour of $U_N = W\rho^N(\hat{\rho}_{LN} - \rho)$ is that of ρT_N . In particular, if $\rho_1^2 > \rho$, it follows that the limit obtained by Heyde in (1.4) identifies with $\rho\rho_1 \; (\rho_1 - \rho)W_1$, while in the case $\rho_1^2 = \rho$ not considered by him, (1.3) holds. If $\rho_1^2 < \rho$, write $S_N + T_N$ as

$$\langle \underline{Z}_N - \underline{Z}_{N-1}\underline{M}, \underline{1} \rangle + \sum_{n=0}^{N-2} \langle \underline{Z}_{n+1} - \underline{Z}_n\underline{M}, (\rho_1-\rho)\rho_1^{N-2-n}\underline{u}_1 \rangle + (\rho_1-\rho)\rho_1^{N-1} \langle \underline{Z}_0, \underline{u}_1 \rangle$$

and check that Lemma 1 applies to prove (1.2).

It should also be noticed that for Heyde's model maximum likelihood estimation based on total generation sizes only is possible. In fact, in the same way as in the one-type case it is easily seen that

$$\hat{m}_{ML} = \Big(\sum_{n=2}^{N} \sum_{k=1}^{X_{n-2}} Y_{n,k}\Big) / \sum_{n=2}^{N} X_{n-2} = (X_N - X_1) / \sum_{n=0}^{N-2} X_n$$

is the MLE of m and that $U_N = W\rho^N(\hat{m}_{ML} - M)$ has the property (1.2). Hence $\rho(\hat{m}_{ML}) = (1+(1+4\hat{m}_{ML})^{\frac{1}{2}})/2$ is the MLE of ρ and $U_N = W\rho^N(\rho(\hat{m}_{ML}) - \rho)$ has the property (1.2) by a standard transformation theorem.

3. Eliminating ρ_1

For the general two-type Galton-Watson process, the evaluation of $\hat{\rho}_{ML}$ requires a rather detailed observational scheme which could hardly be assumed in general. Frequently, the observations will consist of total generation sizes only and the results of AK and Heyde then show poor asymptotic properties of the estimation procedures considered so far if $\rho_1^2 \geq \rho$. We shall here give an affirmative answer to the obvious question whether in the case $\rho_1^2 \geq \rho$ it is possible to eliminate ρ_1 so as to produce an estimator $\tilde{\rho}$ with the property (1.2) and based on the $|\underline{Z}_n|$ only.

The idea is to first recall that T_N in (2.2) dominates S_N and next to note that $T_N \cong \rho_1 T_{N-1} \cong \rho_1^2 T_{N-2}$. Hence the three equations (2.2) for $N, N-1, N-2$ involve only the three unknown ρ, ρ_1, T_{N-2} and terms of smaller magnitude and should thus asymptotically determine the unknown. More precisely, define

$$a_i = \frac{|\underline{Z}_{N-3+i}|}{|\underline{Z}_{N-3}|} \; i=1,2,3, \quad \Theta = \frac{(\rho_1-\rho)\langle \underline{Z}_{N-3}, \underline{u}_1\rangle}{|\underline{Z}_{N-3}|} \qquad \text{and write}$$

$$a_1 = \rho + \Theta + \delta_1 \qquad (3.1a)$$

$$a_2 = a_1\rho + \rho_1\Theta + \delta_2 \qquad (3.1b)$$

$$a_3 = a_2\rho + \rho_1^2\Theta + \delta_3 \qquad (3.1c)$$

where $\delta_1 = \langle \underline{Z}_{N-2} - \underline{Z}_{N-3}\underline{M} , \underline{1}\rangle / |\underline{Z}_{N-3}|$,

$$\delta_2 = \{ \langle \underline{Z}_{N-1}-\underline{Z}_{N-2}\underline{M},\underline{1}\rangle + (\rho_1-\rho)\langle \underline{Z}_{N-2}-\underline{Z}_{N-3}\underline{M},\underline{u}_1\rangle\}/|\underline{Z}_{N-3}| ,$$

$$\delta_3 = \{ \langle \underline{Z}_N-\underline{Z}_{N-1}\underline{M},\underline{1}\rangle + (\rho_1-\rho)\langle Z_{N-1}-\underline{Z}_{N-2}\underline{M},\underline{u}_1\rangle +$$

$$(\rho_1-\rho)\rho_1\langle \underline{Z}_{N-2}-\underline{Z}_{N-3}\underline{M},u_1\rangle\}/|\underline{Z}_{N-3}| .$$

Neglecting the δ_i leads to the equation

$$a_1 = \tilde{\rho} + \tilde{\Theta} \qquad (3.2a)$$

$$a_2 = a_1\tilde{\rho} + \tilde{\rho}_1\tilde{\Theta} \qquad (3.2b)$$

$$a_3 = a_2\tilde{\rho} + \tilde{\rho}_1^2\tilde{\Theta} . \qquad (3.2c)$$

<u>Proposition 1</u> *Suppose that* $\rho_1^2 > \rho$ *and that* $\mathrm{Var}^i\langle \underline{Z}_1,\underline{u}_1\rangle > 0$ *for some* i . *Then:* (i) *With probability one it holds for all large enough* N *that* (3.2) *admits solutions* $\tilde{\rho},\tilde{\rho}_1,\tilde{\Theta}$ *with* $\tilde{\rho} > \tilde{\rho}_1$. *The solutions can be computed by taking first* $\tilde{\rho},\tilde{\rho}_1$ *as the solutions in descending order of*

$$Az^2 + Bz + C = 0 \quad (A = a_1^2 - a_2, B = a_3 - a_1a_2, C = a_2^2 - a_1a_3) \qquad (3.3)$$

and next determine $\tilde{\Theta}$ *by* (3.2a); (ii) $\tilde{\rho}$ *is strongly consistent for* ρ *and* $\tilde{\rho}_1$ *strongly consistent for* ρ_1 ; (iii) $U_N = W\rho^N(\tilde{\rho}-\rho)$ *has the property* (1.2); (iv) *The property* (1.2) *holds also for either of*

$$W\rho^N\tilde{\Theta}(\tilde{\rho}_1-\rho_1) \quad or \quad W\rho_1^n(\tilde{\rho}_1-\rho_1)/W_1 .$$

<u>*Proof*</u>. We first remark that $P(W_1 = 0) = 0$. In fact, the non-degeneracy

condition is well-known to imply $\mathrm{Var}^i W_1 > 0 \quad i = 1,2$ so that $P(W_1 = x) \leq p$ for some $p < 1$ and all x. Writing $W_1 = \rho_1^{-n} \sum_1^{|\underline{Z}_n|} W_1^{n,k}$, with $W_1^{n,k}$ the W_1-functional evaluated in the line of descent initiated by the k^{th} individual at time n, it follows that

$$P(W_1 = 0) = EP(W_1 = 0 | \underline{Z}_n) \leq Ep^{|\underline{Z}_n|} \to 0 .$$

Next note that multiplying (3.1a) by a_1 and subtracting (3.1b) yields $A = (\rho - \rho_1)\Theta + \varepsilon$, $\varepsilon = a_1\delta_1 - \delta_2 + \Theta^2 + \Theta\delta_1$. Here $\Theta \cong (\rho_1 - \rho)(\rho_1/\rho)^{N-3} W_1/W$ is non-zero for N large and $\varepsilon/\Theta \to 0$ a.s. as $N \to \infty$, cf. Lemma 2. Thus $A \cong \Theta(\rho - \rho_1)$ and similarly

$$B \cong \Theta(\rho_1^2 - a_2) \cong \Theta(\rho_1^2 - \rho^2) , \quad C \cong \Theta(\rho_1 a_2 - \rho_1^2 a_1) \cong \Theta\rho\rho_1(\rho - \rho_1) .$$

Now write (3.2) in the form

$$\tilde{\Theta} = a_1 - \tilde{\rho} \tag{3.4a}$$

$$a_2 = a_1(\tilde{\rho} + \tilde{\rho}_1) - \tilde{\rho}\tilde{\rho}_1 \tag{3.4b}$$

$$a_3 = a_2(\tilde{\rho} + \tilde{\rho}_1) - a_1\tilde{\rho}\tilde{\rho}_1 \tag{3.4c}$$

Here (3.4b) follows by inserting (3.4a) in (3.2b) while (3.4c) results upon solving (3.2b) for $\tilde{\rho}_1\tilde{\Theta}$ and insert in (3.2c). Now (3.4b), (3.4c) are a set of linear equations in $\tilde{\rho} + \tilde{\rho}_1, \tilde{\rho}\tilde{\rho}_1$. The determinant $-a_1^2 + a_2 = -A$ is non-zero for N large and solving yields $\tilde{\rho} + \tilde{\rho}_1 = -B/A$, $\tilde{\rho}\tilde{\rho}_1 = C/A$ so that indeed for N large the solution of (3.2) is equivalent to the procedure described in (i).

The above estimates for A,B,C show that the discriminant of (3.3) satisfies

$$D = B^2 - 4AC \cong \Theta^2\{(\rho_1^2 - \rho^2)^2 - 4\rho\rho_1(\rho - \rho_1)^2\} = \Theta^2(\rho - \rho_1)^4$$

Hence the asymptotic form of the solutions is

$$\frac{-B \pm D^{\frac{1}{2}}}{2A} = \frac{\rho^2-\rho_1^2 \pm (\rho-\rho_1)^2}{2(\rho-\rho_1)} = \frac{\rho+\rho_1 \pm (\rho-\rho_1)}{2} = \begin{cases}\rho \\ \rho_1\end{cases}$$

and the strong consistency (ii) follows. For (iii), (iv), subtract (3.1) from (3.2) and perform some elementary manipulations to get

$$Q^* \begin{pmatrix}\tilde{\rho} - \rho \\ \tilde{\rho}_1 - \rho_1 \\ \tilde{\Theta} - \Theta\end{pmatrix} = \begin{pmatrix}\delta_1 \\ \delta_2 \\ \delta_3\end{pmatrix} \tag{3.5}$$

where

$$\underline{Q}^* = \begin{pmatrix}1 & 0 & 1 \\ a_1 & \Theta & \tilde{\rho}_1 \\ a_2 & \Theta(\tilde{\rho}_1 + \rho_1) & \tilde{\rho}_1^2\end{pmatrix} \cong \begin{pmatrix}1 & 0 & 1 \\ \rho & \Theta & \rho_1 \\ \rho^2 & 2\Theta\rho_1 & \rho_1^2\end{pmatrix} = \underline{Q}$$

(say). Furthermore, straightforward algebra shows that $\det\underline{Q}^* \cong \det\underline{Q} = -\Theta(\rho-\rho_1)^2$ so that for N large $\underline{Q}^{*-1}$ exists, and solving (3.5) shows after some calculations that

$$\tilde{\rho} - \rho \cong \frac{\rho_1^2}{(\rho-\rho_1)^2}\delta_1 - \frac{2\rho_1}{(\rho-\rho_1)^2}\delta_2 + \frac{1}{(\rho-\rho_1)^2}\delta_3 \tag{3.6a}$$

$$\tilde{\rho}_1 - \rho_1 \cong -\frac{\rho\rho_1}{\Theta(\rho-\rho_1)}\delta_1 + \frac{\rho+\rho_1}{\Theta(\rho-\rho_1)}\delta_2 - \frac{1}{\Theta(\rho-\rho_1)}\delta_3 \tag{3.6b}$$

$$\tilde{\Theta} - \Theta \cong \frac{\rho^2-2\rho\rho_1}{(\rho-\rho_1)^2}\delta_1 + \frac{2\rho_1}{(\rho-\rho_1)^2}\delta_2 - \frac{1}{(\rho-\rho_1)^2}\delta_3 \tag{3.6c}$$

(up to terms of smaller magnitude). The expressions defining the δ_i show that (up to normalizing constants) the r.h.s. of (3.6a) is of the form considered in Lemma 1, the conditions of which are automatic with $\gamma_n = 1$ in view of $\underline{a}_n = \underline{0}$ $n > 3$, and it follows that the limiting distribution of $(\tilde{\rho}-\rho)|\underline{Z}_{N-3}|/(W\rho^N)^{\frac{1}{2}}$ exists and is normal with mean zero. This is equivalent to (iii). For (iv), exactly the same argument shows that $(\tilde{\rho}_1-\rho_1)\theta|\underline{Z}_{N-3}|/(W\rho^N)^{\frac{1}{2}}$ is asymptotically normal with mean zero and finite variance, which is equivalent to (iv) in view of $\tilde{\theta}/\theta \to 1$ a.s., as follows from (3.6c) by similar estimates as in the study of A . □

Remark 1 The same procedure works also for the case $\rho_1^2 = \rho$, only is it necessary to reformulate (i), (ii) slightly. In fact, the central limit theorem (1.3) for $U_N = \langle \underline{Z}_N, \underline{u}_1 \rangle$ yields now $\delta_i/\theta \to 0$ in probability rather than a.s. It follows from Asmussen (1977) that with probability one $\langle \underline{Z}_N, \underline{u}_1 \rangle / (\rho^N N \log\log N)^{\frac{1}{2}}$ has as its set of limit point a compact interval comprising the origin, but this neither contradicts nor proves $\delta_i/\theta \to 0$ a.s. Thus without further investigations we must rephrase (i) to the probability of (3.2) to be solvable for a given N in the way described to tend to one as $N \to \infty$ and (ii) to $\tilde{\rho} \overset{P}{\to} \rho$, $\tilde{\rho}_1 \overset{P}{\to} \rho_1$.

Remark 2 Explicit expressions for the variance in (1.2) can in principle be deduced from the proof but will not be stated here, since they are complicated and do not apply to produce confidence intervals before offspring variances have been estimated. It should be noted from (iv) that the speed of convergence of $\tilde{\rho}_1$ is slower as for the MLE, which has asymptotic properties similar to $\hat{\rho}_{ML}$, cf. the obvious extension of AK Corollary 5.1.

Remark 3 Though the above considerations certainly do not pretend neither to be of great practical applicability nor to present anything than a first tentative suggestion, it seems reasonable to ask for robustness properties of the procedure. Again, this presents a complicated problem and we shall only give some remarks. Apriori information on the relative sizes of ρ and ρ_1 will seldom be available, although one may try to base a preliminary survey on fluctuations of successive values of $\hat{\rho}_{LN}$, cf. Becker (1977). If $\rho_1^2 < \rho$, one has to replace (3.2) by

$$a_1 = \rho + \varepsilon_1 \ , \ a_2 = a_1\rho + \varepsilon_2 \ , \ a_3 = a_2\rho + \varepsilon_3$$

where $(W\rho^N)^{-\frac{1}{2}}(\varepsilon_1\varepsilon_2\varepsilon_3)$ is asymptotically normal with mean vector $\underline{0}$ and a covariance matrix which is regular except for offspring distributions with special dependence structures. In that case,

$$D = B^2 - 4AC = (\varepsilon_3 - a_2\varepsilon_1)^2 - 4(a_1\varepsilon_1 - \varepsilon_2)(a_2\varepsilon_2 - a_1\varepsilon_3) \cong$$

$$(\varepsilon_3 - \rho^2\varepsilon_1)^2 - 4(\rho\varepsilon_1 - \varepsilon_2)(\rho^2\varepsilon_2 - \rho\varepsilon_3) = (\rho^2\varepsilon_1 - 2\rho\varepsilon_2 + \varepsilon_3)^2$$

and the asymptotic form of the solutions of (3.3) is

$$\frac{-B \pm D^{\frac{1}{2}}}{2A} \cong \frac{\rho^2\varepsilon_1 - \varepsilon_3 \pm (\rho^2\varepsilon_1 - 2\rho\varepsilon_2 + \varepsilon_3)}{2(\rho\varepsilon_1 - \varepsilon_2)} = \begin{cases} \rho \\ \dfrac{\rho\varepsilon_2 - \varepsilon_3}{\rho\varepsilon_1 - \varepsilon_2} \end{cases}$$

so that indeed one root is consistent for ρ while the other has a non-degenerate limit distribution. Consistency will also still hold in some simple cases if the number of types is $p > 2$, but in view of the complicated algebra involved we shall not go into that.

As a final illustration, we shall compute $\tilde{\rho}$ for some smallpox data (Rodrigues-da-Silva et al. (1963)) for which Becker (1977) has suggested the p-type Galton-Watson process as model. Unfortunately the sample size is very small, $N = 4$ with successive generation sizes 1,5,3,12,24. Thus $a_1 = 3/5$, $a_2 = 12/5$, $a_3 = 24/5$ and (3.3) becomes $17z^2 - 28z - 24 = 0$ which yields $\tilde{\rho} = 2.27$, $\tilde{\rho}_1 = -0.62$. Note that the estimates fail to satisfy $\rho_1^2 > \rho$ and that not even the meaning of types nor the value of p is clear in this example. Thus the value of $\tilde{\rho}$ fits surprisingly nicely with other estimates like $\hat{\rho}_B = 2.10$ or

$\hat{\rho}_{LN} = 2.00$ considered by Becker.

Acknowledgements

I would like to thank Chris Heyde for stimulating discussions.

References

Asmussen, S. (1977). Almost sure behavior of linear functionals of supercritical branching processes. *Trans. Amer. Math. Soc.* *231*, 233-248.

Asmussen, S. and Hering, H. (1980). *Lectures on Branching Processes.* (forthcoming).

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

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

Dion, J.-P. and Keiding, N. (1978). Statistical inference in branching processes. *Advances in Probability* *5*, (Joffe-Ney ed.). New York: Academic Press.

Heyde, C.C. (1981a). On a probabilistic analogue of the Fibonacci sequence. *J. Appl. Probability* (to appear).

Heyde, C.C. (1981b). On Fibonacci (or lagged Bienaymé-Galton-Watson) branching processes. *J. Appl. Probability* (to appear).

Heyde, C.C. and Leslie, J.R. (1971). Improved classical limit analogues for Galton-Watson processes with or without immigration. *Bull. Austr. Math. Soc.* *5*, 145-155 (1971).

Keiding, N. and Lauritzen, S.L. (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. (1966a). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.* *37*, 1221-1223.

Kesten, H. and Stigum, B.P. (1966b). Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Ann. Math. Statist.* *37*, 1463-1481.

Rodrigues-da-Silva, G., Rabello, S.I. and Angulo, J.J. (1963). Epidemic of *variola minor* in a suburb of São Paulo. *Public Health Report, Washington* *78* , 165-171.

PREPRINTS 1980

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Olkin, Ingram and Væth, Michael: Maximum Likelihood Estimation in a Two-way Analysis of Variance with Correlated Errors in One Classification.

No. 2 Skovgaard, Ib M.: Transformation of an Edgeworth Expansion by a Sequence of Smooth Functions.

No. 3 Asmussen, Søren: Equilibrium Properties of the M/G/1 Queue.

No. 4 Johansen, Søren and Keiding, Susanne: A Family of Models for the Elimination of Substrate in the Liver.

No. 5 Skovgaard, Ib M.: Edgeworth Expansions of the Distributions of Maximum Likelihood Estimators.

No. 6 Tjur, Tue: A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model.

No. 7 Lauritzen, Steffen L. : Examples of Extreme Point Models in Statistice.

PREPRINTS 1981

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Johansen, Søren: Asymptotic Inference in Random Coefficient Regression Models.

No. 2 Asmussen, Søren: On the Role of a Certain Eigenvalue in Estimating the Growth Rate of a Branching Process.