THE INSTITUTE OF STATISTICS

THE CONSOLIDATED UNIVERSITY OF NORTH CAROLINA

A UNIFIED LIMIT THEORY VIA BOOTSTRAP FOR BRANCHING PROCESSES WITH IMMIGRATION

by

S. Datta and T.N. Sriram

August 1993

Mimeo Series #2306

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

MIMEO DATTA, S.
SERIES SRIRAM, T.N.
#2306

A UNIFIED LIMIT THEORY VIA BOOTSTRAP FOR BRANCHING PROCESSES WITH IMMIGRATION

NAME	DATE

A UNIFIED LIMIT THEORY VIA BOOTSTRAP FOR BRANCHING PROCESSES WITH IMMIGRATION

SOMNATH DATTA and T.N. SRIRAM*

Department of Statistics
University of Georgia
Athens, GA 30602 U.S.A.

ABSTRACT

In this paper we consider bootstrap approximation to the sampling distribution of the maximum likelihood estimator (m.l.e.) of the offspring mean m in a branching process with immigration. A clever modification of the standard parametric bootstrap procedure is shown to eliminate the invalidity of the standard bootstrap for the case m=1, as reported in Sriram (1992). Furthermore, the modified bootstrap is shown to provide valid approximations for other values of m ($\neq 1$) as well. Thus, in this example, the modified bootstrap provides a unified solution while the form of the limit distribution of the m.l.e. via classical asymptotic theory depends on m. It is argued that similar modifications will be useful more generally.

AMS(1990) Subject Classifications: Primary 60J80,62G09.

Keywords and phrases: Bootstrap, branching processes with immigration, asymptotic validity, limit theory.

*A major portion of the work was done while the second author was visiting the Department of Statistics, University fo North Carolina at Chapel Hill.

1. Introduction.

Consider a branching process with immigration which can be defined recursively by the following equation:

$$Z_i = \sum_{k=1}^{Z_{i-1}} \xi_{i-1,k} + Y_i, \quad i = 1, 2, \dots \quad , \tag{1.1}$$

with $Z_0 = 1$. Here Z_i is the size of the i-th generation of a population, $\xi_{i-1,k}$ is the offspring size of the k-th individual belonging to the $(i-1)$-th generation and Y_i is the number of immigrants contributing to the population's i-th generation. Throughout the paper, we assume that $\{\xi_{i-1,k}\}$ and $\{Y_i\}$ in (1.1) are independent sequences of i.i.d. nonnegative, integer valued random variables (r.v.'s) with finite means m and λ, and finite and positive variances σ^2 and b^2, respectively.

Suppose that a sample $\{(Z_i, Y_i), i = 1, \dots, n\}$ is available. Then a natural estimator of the offspring mean m and the immigration mean λ are given respectively by

$$\hat{m}_n = \left(\sum_{i=1}^{n} Z_{i-1}\right)^{-1} \sum_{i=1}^{n} (Z_i - Y_i)$$

and

$$\hat{\lambda}_n = n^{-1} \sum_{i=1}^{n} Y_i. \tag{1.2}$$

It is also possible to estimate m and λ based only on the partial information on $\{Z_i\}$ and study their properties; see Heyde and Seneta (1972, 1974), Wei and Winnicki (1990) and the references therein.

For statistical inference about parameter m, one may consider the pivot given by

$$V_n = (\sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}} (\hat{m}_n - m). \tag{1.3}$$

Generally, knowing the distribution of the pivot permits forming confidence intervals, setting up tests of hypotheses about m etc. However, here, the form of the limit distribution of V_n depends on m. More specifically, it is known that, as $n \to \infty$,

$$V_n \xrightarrow{D} \begin{cases} N(0, \sigma^2) & \text{if} \quad m \neq 1 \\ \{X(1) - \lambda\} / \{\int_0^1 X(t)dt\}^{\frac{1}{2}} = V & \text{if} \quad m = 1, \end{cases} \tag{1.4}$$

where $\{X(t)\}$ is a nonnegative diffusion process with generator $\mathcal{A}h(x) = \lambda h'(x) + (\frac{1}{2}) x \sigma^2 h''(x)$, for $h \in C_c^\infty[0, \infty)$, and is obtained as a weak limit of the process $Z_{[nt]}/n$, as $n \to \infty$. Here

$C_c^\infty[0,\infty)$ is the space of all infinitely differentiable functions on $[0,\infty)$ which have compact supports, and $'$ and $''$ denote the first and the second derivative, respectivley. For the result (1.4), see Sriram, Basawa and Huggins (1991) for the case $m < 1$ (subcritical) and $m = 1$ (critical), and Wei and Winnicki (1989) for $m > 1$ (supercritical).

In an attempt to approximate the sampling distribution of V_n, Sriram (1992) considered the standard parametric bootstrap. However, it was shown that it does not lead to an asymptotically valid approximation for the case $m = 1$. See Sriram (1992) for details. Because of the failure of standard parametric bootstrap at $m = 1$, the investigation of the same for other values of m (namely $m \neq 1$) was not carried out in Sriram (1992).

In this paper, we propose a clever modification of the standard parametric bootstrap and show that the modified procedure provides an asymptotically valid approximation to the sampling distribution of V_n, not only for the case $m = 1$, but also for the case $m \neq 1$. Thus, in this example, the modified parametric bootstrap serves as a unifying inference tool.

The basic idea of modifying a standard bootstrap so as to provide a unified solution can be used in other critical cases known in the literature as well. For example, it has been shown by Basawa et al. (1991) and Datta (1992) that for autoregression a similar invalidity results from the use of standard bootstrap, when the autoregressive parameter is $\pm$ 1. It is possible to propose a similar modified bootstrap scheme for autoregression as done here and establish its asymptotic validity for all values of the autoregressive parameter, but it will be considered in a forthcoming article.

Bootstrap methods have received considerable attention since the pioneering work of Efron (1979). A good survey of results for the i.i.d. setup is provided in a review article by Babu (1989); one may also consult the recent book by Hall (1992). Bootstrap schemes for various dependent models have been proposed by Bose (1988), Künsch (1989), Basawa et al. (1989), Lahiri (1991), Athreya and Fuh (1992), Liu and Singh (1992), Politis and Romano (1992), Datta and McCormick (1992 and 1993), among others.

2. The Bootstrap and Summary of Results.

For the purpose of bootstrap, we assume that the offspring and the immigration r.v.'s have a power series distribution with respective probability mass functions (p.m.f.'s) given by

$$P_\theta[\xi = u] = a(u)\theta^u/A(\theta), \qquad u = 0, 1, \ldots,$$

and

$$Q_\phi[Y = y] = b(y)\phi^y / B(\phi), \qquad y = 0, 1, \ldots \quad . \tag{2.1}$$

Here, $\{a(u)\}$ and $\{b(y)\}$ are known nonnegative sequences, $A(\theta) = \sum_{u=0}^{\infty} a(u)\theta^u$ for $0 < \theta < \theta^*$ and $B(\phi) = \sum_{y=0}^{\infty} b(y)\phi^y$ for $0 < \phi < \phi^*$, where θ^* and ϕ^* are the radii of the two power series. Note that, under the parametric model (2.1), $\hat{m}_n$ and $\hat{\lambda}_n$ given in (1.2) are maximum likelihood estimators of m and λ, respectively; see, for instance, Bhat and Adke (1981).

It can be easily shown that $m = E_\theta(\xi) = \theta A'(\theta)/A(\theta)$, $\quad \lambda = E_\phi(Y) = \phi B'(\phi)/B(\phi)$ and the variances $\sigma^2 = V_\theta(\xi) = \theta(\partial m/\partial\theta)$, and $b^2 = V_\phi(Y) = \phi(\partial\lambda/\partial\phi)$. Here ∂ denotes a partial derivative. Since θ, ϕ, σ^2 and b^2 are all assumed to be positive, we have that m and λ are strictly increasing functions of θ and ϕ, respectively. Let $m = f(\theta)$ and $\lambda = g(\phi)$, where f and g are known functions.

A (parametric) bootstrap procedure to approximate the sampling distribution of V_n in (1.3) can be described as follows. Given a sample $\mathcal{X}_n = \{(Z_i, Y_i), i = 1, \ldots, n\}$, estimate the offspring mean m by some estimator $\tilde{m}_n$ based on $\mathcal{X}_n$ and the immigration mean λ by $\hat{\lambda}_n$ defined in (1.2). Replace θ and ϕ in (2.1) by their respective estimates $\tilde{\theta}_n = f^{-1}(\tilde{m}_n)$ and $\hat{\phi}_n = g^{-1}(\hat{\lambda}_n)$. Conditional on $\mathcal{X}_n$, let $\{\xi^*_{i,j}\}$ be a sequence of i.i.d. r.v.'s having p.m.f. $P_{\tilde{\theta}_n}$ and $\{Y_i^*\}$ be a sequence of i.i.d. r.v.'s having p.m.f. $Q_{\hat{\phi}_n}$. The bootstrap sample $\mathcal{X}_n^* = \{(Z_i^*, Y_i^*), i = 1, \ldots, n\}$ is then obtained recursively from

$$Z_i^* = \sum_{k=1}^{Z_{i-1}^*} \xi^*_{i-1,k} + Y_i^*, \quad i = 1, 2, \ldots, \tag{2.2}$$

with $Z_0^* = 1$. Define the bootstrap analogue of $\hat{m}_n$ and V_n by

$$\hat{m}_n^* = (\sum_{i=1}^n Z_{i-1}^*)^{-1} \sum_{i=1}^n (Z_i^* - Y_i^*), \tag{2.3}$$

and

$$V_n^* = (\sum_{i=1}^n Z_{i-1}^*)^{\frac{1}{2}} (\hat{m}_n^* - \tilde{m}_n), \tag{2.4}$$

respectively. The (conditional) distribution of V_n^* in (2.4) (given the original sample $\mathcal{X}_n$) constitutes a bootstrap approximation to the distribution of V_n.

Note that, so far we have left the selection of the estimator $\tilde{m}_n$ for parametric bootstrap quite arbitrary. Clearly, the natural choice for it is the m.l.e $\hat{m}_n$ in (1.2) itself. This choice of $\tilde{m}_n$ corresponds to the standard bootstrap mentioned in the introduction. However, as mentioned earlier, with $\tilde{m}_n = \hat{m}_n$ Sriram (1992) showed that the conditional limit distribution of the bootstrap pivot V_n^* ***does not*** coincide with the limit distribution of V_n in (1.4), when $m = 1$. In other words, the asymptotic validity does not hold for the standard parametric bootstrap. A deeper analysis shows that the main reason for its failure is that, when $m = 1$, the estimated distribution $P_{f^{-1}(\hat{m}_n)}$ no longer serves as a good choice for the bootstrap population, since $\hat{m}_n$ does not converge to m fast enough.

In this paper we propose the following selection of $\tilde{m}_n$ which converges sufficiently fast to m, when $m = 1$, and still remains consistent for other values of m. The idea behind it is that of an adaptive (data-dependent) shrinkage towards $m = 1$ (similar in spirit to the Hodges estimator, see LeCam (1953)). In order to describe our selection of $\tilde{m}_n$, let $\{\eta_n\}$ be a non-random sequence satisfying $n\eta_n \to 0$. Let $\delta_n = h(\sum_{i=1}^n Z_{i-1})$ where h is a positive function on $[0, \infty)$ such that $\lim_{z\to\infty}(z/\log\log z)^{\frac{1}{2}} h(z) = \infty$, $\lim_{z\to\infty} h(z) = 0$; eg. $h(z) = z^{-\frac{1}{3}}$. For $\hat{m}_n$ in (1.2), define

$$\tilde{m}_n = \begin{cases} 1 - \eta_n & \text{if} \quad 1 - \delta_n \le \hat{m}_n \le 1 - \eta_n \\ 1 + \eta_n & \text{if} \quad 1 + \eta_n \le \hat{m}_n \le 1 + \delta_n \\ \hat{m}_n & \text{otherwise.} \end{cases} \tag{2.5}$$

The role of δ_n and η_n in the definition of $\tilde{m}_n$ is explained further in Remark 2.1 below.

The main result of this paper is that, with the above selection of $\tilde{m}_n$, the (conditional) distribution of bootstrap pivot V_n^* in (2.4) is asymptotically close to the distribution of the pivot V_n given in (1.3). In order to describe it formally, we let P^* denote the probability corresponding to the bootstrap sample $\mathcal{X}_n^*$, conditional on the original sample $\mathcal{X}_n$.

Theorem 2.1. Consider the model given in (1.1) and assume (2.1). For V_n in (1.3) and V_n^* in (2.4), where $\tilde{m}_n$ is given by (2.5), we have

$$\sup_x |P\{V_n \le x\} - P^*\{V_n^* \le x\}| \to 0 \quad \text{as} \quad n \to \infty, \tag{2.6}$$

almost surely (a.s.), for all $m > 0$.

Remark 2.1. The construction of $\tilde{m}_n$ ensures that it is within η_n distance of 1 whenever $\hat{m}_n$ is within δ_n distance of 1 (an indication that the true m is one or very close to one). A possible, perhaps the simplest, choice of η_n is zero. However, use of a non-zero η_n, eg., $\eta_n = (nlogn)^{-1}$ may be better for small sample properties of the bootstrap when m is very close to one.

In order to prove Theorem 2.1 we obtain a series of results which are of independent interest as well. Different techniques are used for the three cases $m < 1, m = 1$, and $m > 1$. In all the cases, properties of the estimator $\tilde{m}_n$ in (2.5) (which in turn uses that of $\hat{m}_n$) are required. These are presented in the next section. Necessary limit theorems for array of branching processes for the three cases are obtained in Section 4. Finally, Theorem 2.1 is proved in Section 5 using results obtained in Sections 3 and 4.

3. Properties of the Modified Estimator.

Let $\tilde{m}_n$ be the estimator given in (2.5). In this section we obtain a number of properties of $\tilde{m}_n$ which will be useful in proving Theorem 2.1. First, we state and prove a law of the iterated logarithm (LIL) for $\hat{m}_n$ in (1.2), when $m = 1$. The LIL result will be used in the proposition following the proof of the lemma. Note that for Lemma 3.1 and Proposition 3.1 below we do not require the distributional assumption in (2.1), hence we state the necessary moment conditions.

Lemma 3.1. Suppose $m = 1$ in model (1.1) and $E|\xi_{1,1}|^{2+s} < \infty$ for some $s > 0$. Then for $\hat{m}_n$ defined in (1.2),

$$(\hat{m}_n - 1) = O(\left(\frac{\log\log \sum_{i=1}^n Z_{i-1}}{\sum_{i=1}^n Z_{i-1}}\right)^{\frac{1}{2}}) \quad \text{a.s.,} \quad \text{as} \quad n \to \infty. \tag{3.1}$$

Proof. It will be shown below that conclusion (3.1) follows from Lemma 2 (result (2.4)) of Wei (1985). To this end, let

$$\begin{aligned} w_i = (Z_i - Z_{i-1} - Y_i) &= \sum_{k=1}^{Z_{i-1}} (\xi_{i-1,k} - 1) \\ &= u_i \epsilon_i, \end{aligned} \tag{3.2}$$

where $u_i = \sqrt{Z_{i-1}}$, and $\epsilon_i = w_i/u_i$, if $u_i \neq 0$ and 0 otherwise. Clearly, $\{\epsilon_i\}$ is a sequence of martingale differences with respect to the σ-field $\mathcal{F}_n = \sigma\{\xi_{i-1,k}, Y_i, 1 \le i \le n, k \ge 1\}$. Moreover, by Lemma 2.1 of Lai and Wei (1983), there is a constant C such that

$$E\{|\epsilon_i|^{2+s}|\mathcal{F}_{i-1}\} \le CE|\xi_{1,1} - 1|^{2+s} \quad \text{a.s..} \tag{3.3}$$

Hence, condition (1.2) of Lemma 2 in Wei (1985) holds. It only remains to check condition (2.3) of Lemma 2 in Wei (1985), which amounts to showing that for some $0 < c < 1$,

$$Z_{n-1} = o((\sum_{i=1}^n Z_{i-1})^c) \quad \text{a.s.,} \quad \text{as} \quad n \to \infty. \tag{3.4}$$

It can be shown using arguments similar to the proof of Lemma A of Sriram, Basawa and Huggins (1991) that for $\gamma > \frac{1}{2}$,

$$\lim_{k \to \infty} P(Z_n > \delta(\sum_{i=1}^{n} Z_{i-1})^{\gamma} \quad \text{for some} \quad n \geq k) = 0, \tag{3.5}$$

for any $\delta > 0$. This implies that for $\gamma > \frac{1}{2}$,

$$Z_n = o((\sum_{i=1}^{n} Z_{i-1})^{\gamma}) \quad \text{a.s.,} \quad \text{as} \quad n \to \infty. \tag{3.6}$$

Since $\sum_{i=1}^{n} Z_{i-1} \leq \sum_{i=1}^{n+1} Z_{i-1}$ we have from (3.6) that for $\gamma > \frac{1}{2}$, $Z_n = o((\sum_{i=1}^{n+1} Z_{i-1})^{\gamma})$ a.s. as $n \to \infty$. This yields (3.4). Hence, as $n \to \infty$,

$$\sum_{i=1}^{n} (Z_i - Z_{i-1} - Y_i) = o(\left((\sum_{i=1}^{n} Z_{i-1}) \log\log(\sum_{i=1}^{n} Z_{i-1}) \right)^{\frac{1}{2}}) \quad \text{a.s.,}$$

which by the definition of $\hat{m}_n$ yields the required result. ■

Proposition 3.1. Consider the model (1.1). The following hold for the estimator $\tilde{m}_n$ in (2.5):

Case $m \neq 1$:

$$\tilde{m}_n \to m \quad \text{a.s., as} \quad n \to \infty. \tag{3.7}$$

Case $m = 1$: Suppose for some $s > 0, E|\xi_{1,1}|^{2+s} < \infty$. Then

$$n(\tilde{m}_n - 1) \to 0 \quad \text{a.s.,} \quad \text{as} \quad n \to \infty. \tag{3.8}$$

Case $m > 1$:

$$n(\tilde{m}_n - m) \to 0 \quad \text{a.s.,} \quad \text{as} \quad n \to \infty, \tag{3.9}$$

and hence

$$(\frac{\tilde{m}_n}{m})^n \to 1 \quad \text{a.s.,} \quad \text{as} \quad n \to \infty. \tag{3.10}$$

Proof. Let $m \neq 1$. Note that (3.7) follows readily if we show for $\hat{m}_n$ in (1.2) that

$$\hat{m}_n \to m \quad \text{a.s., as} \quad n \to \infty \tag{3.11}$$

and

$$P(\tilde{m}_n \neq \hat{m}_n \quad i.o.) = 0. \tag{3.12}$$

For (3.11), note that

$$\hat{m}_n - m = \sum_{i=1}^{n}(Z_i - mZ_{i-1} - Y_i) / \sum_{i=1}^{n} Z_{i-1} \to 0 \quad \text{a.s.,} \quad \text{as} \quad n \to \infty, \tag{3.13}$$

by the strong law of large numbers for martingales (Hall and Heyde (1980), [HH] hereafter, Theorem 2.18) applied to the martingale sequence $\{\sum_{i=1}^{n}(Z_i - mZ_{i-1} - Y_i), \quad \mathcal{F}_n\}$ for $\mathcal{F}_n$ defined in Lemma 3.1. As for (3.12),

$$\begin{aligned} P(\tilde{m}_n \neq \hat{m}_n \quad i.o.) &\leq P(|\hat{m}_n - 1| \leq \delta_n \quad i.o.) \\ &\leq P(|\hat{m}_n - m| \geq |m-1| - \delta_n \quad i.o.) \\ &= 0, \end{aligned}$$

because $m \neq 1, \delta_n \downarrow 0$ a.s. and (3.13) holds. Hence the case $m \neq 1$.

Let $m = 1$ and $\eta > 0$. Then, there exists N such that $n\eta_n < \eta$ for all $n \geq N$. Therefore

$$\begin{aligned} P(n|\tilde{m}_n - 1| > \eta \quad i.o.) &\leq P(|\hat{m}_n - 1| > \delta_n \quad i.o.) \\ &= P\left(\frac{(\sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}}|\hat{m}_n - 1|}{(\log\log \sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}}} > \frac{\delta_n(\sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}}}{(\log\log \sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}}} \quad i.o.\right) \\ &= 0 \end{aligned}$$

by the assumptions on δ_n (see (2.5)) and Lemma 3.1. Hence the case $m = 1$.

Finally, let $m > 1$. Observe that (3.9) follows from (3.12) provided

$$n(\hat{m}_n - m) \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty. \tag{3.14}$$

For $\epsilon \in (0, \frac{1}{2})$, write using the definition of $\hat{m}_n$ that

$$n(\hat{m}_n - m) = [n/(\sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}-\epsilon}][\sum_{i=1}^{n}(Z_i - mZ_{i-1} - Y_i)/(\sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}+\epsilon}]. \tag{3.15}$$

Now (3.14) follows since the strong law for martingales (see [HH], Theorem 2.18) implies that $\sum_{i=1}^{n}(Z_i - mZ_{i-1} - Y_i)/(\sum_{i=1}^{n} Z_{i-1})^{\frac{1}{2}+\epsilon} \to 0$ a.s., and a result of Seneta (1970) [also see Heyde (1970), Theorem 3] implies that

$$m^{-n}\sum_{i=1}^{n} Z_{i-1} \to (m-1)^{-1}W \quad \text{a.s.} \quad \text{as} \quad n \to \infty, \tag{3.16}$$

where W is a positive random variable. It is easy to see that (3.9) implies (3.10). Hence the proposition. ■

4. Array of Branching Processes

Consider a general array of branching processes $\{Z_i^{(n)}\}$ given by

$$Z_i^{(n)} = \sum_{k=1}^{Z_{i-1}^{(n)}} \xi_{i-1,k}^{(n)} + Y_i^{(n)}, \quad i = 1, 2, \ldots, \tag{4.1}$$

where, for each n, $Z_0^{(n)} = 1$, $\{\xi_{i,j}^{(n)}\}$ is a sequence of i.i.d. r.v.'s with mean μ_n and variance σ_n^2, and $\{Y_i^{(n)}\}$ is a sequence of i.i.d. r.v.'s with mean λ_n and variance b_n^2; also, assume that $\{\xi_{i,j}^{(n)}\}$ and $\{Y_i^{(n)}\}$ are independent. The following condition is assumed throughout this section: For the model defined in (4.1)

(C-1)

$$\mu_n \to m, \qquad \lambda_n \to \lambda, \qquad \sigma_n^2 \to \sigma^2 \quad \text{as} \quad n \to \infty, \tag{4.2}$$

where m, λ, and σ^2 are all positive and finite. Define

$$\hat{\mu}_n = \Big(\sum_{i=1}^{n} Z_{i-1}^{(n)}\Big)^{-1} \sum_{i=1}^{n} (Z_i^{(n)} - Y_i^{(n)}) \tag{4.3}$$

and the pivot

$$\nu_n = \Big(\sum_{i=1}^{n} Z_{i-1}^{(n)}\Big)^{\frac{1}{2}} (\hat{\mu}_n - \mu_n). \tag{4.4}$$

In this section, we derive the limit distribution of ν_n for the cases $m < 1$, $m = 1$ and $m > 1$ under (C-1) and other regularity conditions. Incidentally, the limit distribution of ν_n when $m = 1$ in (C-1) has been derived by Sriram (1992) and it is stated next.

Theorem 4.1. For model (4.1), assume condition (C-1) in (4.2) with $m = 1$. Furthermore, assume that $b_n^2 \to b^2 \in (0, \infty)$ and for a real number α

$$\mu_n = 1 + \alpha n^{-1} + o(n^{-1}) \quad \text{as} \quad n \to \infty. \tag{4.5}$$

Suppose, for any sequence $\{x_n\}$ such that $x_n \to x \in (0, \infty)$,

$$\lim_{n\to\infty} \sigma_n^{-2} E|\xi_{1,1}^{(n)} - \mu_n|^2 I_{\{|\xi_{1,1}^{(n)} - \mu_n| \geq \epsilon \sigma_n \sqrt{n x_n}\}} = 0 \tag{4.6}$$

for all $\epsilon > 0$. Then for ν_n defined in (4.4),

$$\nu_n \xrightarrow{D} \nu(\alpha, \lambda, \sigma^2) \quad \text{as} \quad n \to \infty, \tag{4.7}$$

where

$$\nu(\alpha, \lambda, \sigma^2) = \left\{\int_0^1 X_\alpha(t)dt\right\}^{-\frac{1}{2}} \{X_\alpha(1) - \lambda\} - \alpha \left\{\int_0^1 X_\alpha(t)dt\right\}^{\frac{1}{2}} \tag{4.8}$$

and $\{X_\alpha(t)\}$ is a nonnegative diffusion process with generator

$$\mathcal{A}_\alpha h(x) = \alpha x h'(x) + \lambda h'(x) + (1/2)x\sigma^2 h''(x), \qquad h \in C_c^\infty[0, \infty). \tag{4.9}$$

Proof. See Sriram (1992), Corollary 3.1. ■

Next, we consider the case when $\mu_n \to m < 1$. First, we state (without proof) a L_1-convergence for martingale arrays, which will be used in the theorem proved below.

Lemma 4.1. For each $n \geq 1$, let $\{U_{ni}, \mathcal{G}_{ni}, 1 \leq i \leq n\}$ define a martingale difference sequence. If

$$\lim_{M\to\infty} \sup_n \sup_{1\leq i\leq n} E|U_{ni}|I_{\{|U_{ni}|>M\}} = 0 \tag{4.10}$$

then

$$n^{-1}E|\sum_{i=1}^n U_{ni}| \to 0. \tag{4.11}$$

The above lemma can be proved using exactly the same arguments as in the proof of Theorem 2.22 of [HH] with $p = 1$, even though the theorem in [HH] does not consider an array.

For the process in (4.1) define a sequence of σ-fields $\mathcal{F}_{ni}$, for each $n \geq 1$, by

$$\mathcal{F}_{ni} = \sigma\{\xi_{l-1,j}^{(n)}, Y_l^{(n)}, 1 \leq l \leq i, \quad j \geq 1\}. \tag{4.12}$$

Theorem 4.2. For model (4.1), assume condition (C-1) in (4.2) with $m < 1$ and also that $b_n^2 \to b^2 \in (0, \infty)$. Furthermore, assume that for some $\delta > 0$, $E|\xi_{1,1}^{(n)}|^{2+\delta} \leq B$ for all $n \geq 1$. Then, for ν_n defined in (4.4),

$$\nu_n \xrightarrow{D} N(0, \sigma^2) \quad \text{as} \quad n \to \infty. \tag{4.13}$$

Proof. Use (4.3) and (4.4) to write

$$\begin{aligned} \nu_n &= (\sum_{i=1}^n Z_{i-1}^{(n)})^{-\frac{1}{2}} \sum_{i=1}^n (Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - Y_i^{(n)}) \\ &= \left(n^{-1}\sum_{i=1}^n Z_{i-1}^{(n)}\right)^{-\frac{1}{2}} n^{-\frac{1}{2}} \sum_{i=1}^n (Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - Y_i^{(n)}). \end{aligned} \tag{4.14}$$

Let

$$X_{ni} = n^{-\frac{1}{2}}(Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - Y_i^{(n)}) \quad \text{and} \quad S_{nk} = \sum_{i=1}^{k} X_{ni}. \tag{4.15}$$

Then, for each $n \geq 1$, $\{S_{nk}, \mathcal{F}_{nk}, k \geq 1\}$ defines a martingale sequence, so that $\{S_{nk}, \mathcal{F}_{nk}, k \geq 1, n \geq 1\}$ defines a martingale array with the conditional variance given by

$$\begin{aligned} V_{nn}^2 &= \sum_{i=1}^{n} E(X_{ni}^2 | \mathcal{F}_{n,i-1}) \\ &= \sigma_n^2 n^{-1} \sum_{i=1}^{n} Z_{i-1}^{(n)}. \end{aligned} \tag{4.16}$$

Use (4.15) and (4.16) to write $\nu_n = \sigma_n(S_{nn}/V_{nn})$. Then, we will use the martingale array CLT given in Corollary 3.2 of [HH] to prove (4.13). To this end, it suffices to check the conditions of Corollary 3.2 of [HH] (also see Corollary 3.1 and Theorem 3.2 of [HH]). First we show that V_{nn}^2 converges in probability. For this, observe that

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}(Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - \lambda_n) &= n^{-1}\sum_{i=1}^{n}\{(1-\mu_n)Z_i^{(n)} + \mu_n(Z_i^{(n)} - Z_{i-1}^{(n)}) - \lambda_n\} \\ &= (1-\mu_n)n^{-1}\sum_{i=1}^{n} Z_i^{(n)} + \mu_n n^{-1}(Z_n^{(n)} - Z_0^{(n)}) - \lambda_n. \end{aligned} \tag{4.17}$$

Let $U_{ni} = (Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - \lambda_n)$, then $\{U_{ni}, \mathcal{F}_{ni}, i \geq 1\}$ defines a martingale difference sequence for each $n \geq 1$. Furthermore,

$$\begin{aligned} E|U_{ni}|^2 &= E\left(\sum_{j=1}^{Z_{i-1}^{(n)}}(\xi_{i-1,j}^{(n)} - \mu_n) + (Y_i^{(n)} - \lambda_n)\right)^2 \\ &= \sigma_n^2 E(Z_{i-1}^{(n)}) + b_n^2, \quad \text{by conditioning on} \quad \mathcal{F}_{n,i-1}, \\ &= \sigma_n^2\{\mu_n^{i-1} + \lambda_n(1 + \mu_n + \cdots + \mu_n^{i-2})\} + b_n^2 \end{aligned} \tag{4.18}$$

by repeated conditioning and the fact that $Z_0^{(n)} = 1$. Now use (C-1) in (4.2) and that $\mu_n \to m < 1$ to get

$$\sup_n \sup_{1 \leq i \leq n} E|U_{ni}|^2 < \infty. \tag{4.19}$$

This implies condition (4.10) of Lemma 4.1. Hence, by Lemma 4.1,

$$n^{-1}\sum_{i=1}^{n}(Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - \lambda_n) \xrightarrow{L_1} 0 \quad \text{as} \quad n \to \infty, \tag{4.20}$$

and therefore (4.20) holds in probability as well. Also, since

$$E(Z_n^{(n)}) = \mu_n^n + \lambda_n(1 + \mu_n + \cdots + \mu_n^{n-1}),$$

we have by (C-1) that

$$n^{-1} Z_n^{(n)} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty. \tag{4.21}$$

Using (C-1) once more, (4.17), (4.20) and (4.21) we get

$$n^{-1} \sum_{i=1}^{n} Z_i^{(n)} \xrightarrow{P} \lambda/(1-m) \quad \text{as} \quad n \to \infty. \tag{4.22}$$

Hence, from (4.16), (4.21) and (4.22)

$$V_{nn}^2 \xrightarrow{P} \sigma^2 \lambda/(1-m) \quad \text{as} \quad n \to \infty. \tag{4.23}$$

Let $\eta = \sigma^2\lambda/(1-m)$. Since η is a constant, condition (3.21) of Theorem 3.2 in [HH] can be dropped (see Remarks after Corollary 3.1 in [HH]). Now, it only remains to show for X_{ni} defined in (4.15) that for all $\epsilon > 0$,

$$\sum_{i=1}^{n} E\{X_{ni}^2 I_{\{|X_{ni}|>\epsilon\}} | \mathcal{F}_{n,i-1}\} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty. \tag{4.24}$$

For this, it suffices to show that

$$\sum_{i=1}^{n} E\{|X_{ni}|^{2+\delta} | \mathcal{F}_{n,i-1}\} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty, \tag{4.25}$$

where δ is as in the hypothesis of the theorem and $\delta \in (0,2)$. But, by a result of Chow and Teicher (1978) [see Corollary 10.3.2, p. 357] there exists a constant $K_\delta > 0$ such that

$$\begin{aligned} E\{|Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - Y_i^{(n)}|^{2+\delta} | \mathcal{F}_{n,i-1}\} &= E\{|\sum_{k=1}^{Z_{i-1}^{(n)}} (\xi_{i-1,k}^{(n)} - \mu_n)|^{2+\delta} | \mathcal{F}_{n,i-1}\} \\ &\leq K_\delta |Z_{i-1}^{(n)}|^{1+\delta/2} E|\xi_{1,1}^{(n)} - \mu_n|^{2+\delta}. \end{aligned}$$

Hence, by the assumptions that $E|\xi_{1,1}^{(n)}|^{2+\delta} \leq B$ for all $n \geq 1$, and $\mu_n \to m < 1$ we have for some constant B_1 that

$$\begin{aligned} \sum_{i=1}^{n} E\{|X_{ni}|^{2+\delta} | \mathcal{F}_{n,i-1}\} &\leq B_1 n^{-(1+\delta/2)} \sum_{i=1}^{n} |Z_{i-1}^{(n)}|^{1+\delta/2} \\ &\xrightarrow{P} 0 \quad \text{as} \quad n \to \infty, \end{aligned} \tag{4.26}$$

because, for $\delta \in (0,2), E|Z_{i-1}^{(n)}|^{1+\delta/2} \leq E(Z_{i-1}^{(n)})^2$ and it is possible to show using (C-1) in (4.2) and the assumption $m < 1$ that $\sum_{i=1}^{n} E(Z_{i-1}^{(n)})^2 = O(n)$, as $n \to \infty$. Hence, the theorem follows from Corollary 3.2 of [HH], (4.23) and (4.26). ■

Finally, consider the case when $\mu_n \to m > 1$. For this case, assume further that $\{\xi_{i,j}^{(n)}\}$ and $\{Y_i^{(n)}\}$ defined in (4.1) satisfy the following: For each $n \geq 1$, the r.v.

$$\xi_{1,1}^{(n)} \quad \text{has a power series distribution with p.m.f.} \quad P_{\theta_n},$$

and

$$Y_1^{(n)} \quad \text{has a power series distribution with p.m.f.} \quad Q_{\phi_n}, \tag{4.27}$$

where P_{θ_n} and Q_{ϕ_n} are as defined in (2.1) with θ and ϕ replaced by θ_n and ϕ_n, respectively. Let F_{θ_n} and G_{ϕ_n} be distribution functions associated with P_{θ_n} and Q_{ϕ_n}, respectively. Also, let F_θ and G_ϕ be distribution functions associated with P_θ and Q_ϕ defined in (2.1).

Theorem 4.3. For the model (4.1), assume condition (C-1) in (4.2) with $m > 1$. Furthermore, assume that (4.27) holds and

$$n(\mu_n - m) \to 0 \quad \text{as} \quad n \to \infty. \tag{4.28}$$

Then, for ν_n defined in (4.4),

$$\nu_n \xrightarrow{D} N(0, \sigma^2) \quad \text{as} \quad n \to \infty.$$

For the proof of Theorem 4.3, we construct an array process $\{\tilde{Z}_i^{(n)}\}$ and a process $\{\tilde{Z}_i\}$ on a common probability space, in the following way: Let $\{U_{i,j}\}$ and $\{V_i\}$ be two sequences of uniform (0,1) r.v.'s, all of which are independent and defined on a common probability space. Define, for $F_{\theta_n}, G_{\phi_n}, F_\theta$ and G_θ above,

$$\tilde{\xi}_{i,j}^{(n)} = F_{\theta_n}^{-1}(U_{i,j}), \quad \tilde{Y}_i^{(n)} = G_{\phi_n}^{-1}(V_i), \quad \tilde{\xi}_{i,j} = F_\theta^{-1}(U_{i,j}) \quad \text{and} \quad \tilde{Y}_i = G_\phi^{-1}(V_i), \tag{4.29}$$

$i, j \geq 1$. Here, for a distribution function H, $H^{-1}(w) = \inf\{x : H(x) \geq w\}$, for $0 < w < 1$. Now, define $\tilde{Z}_i^{(n)}$ by

$$\tilde{Z}_i^{(n)} = \sum_{k=1}^{\tilde{Z}_{i-1}^{(n)}} \tilde{\xi}_{i-1,k}^{(n)} + \tilde{Y}_i^{(n)}, \quad i = 1, 2, \ldots, \tag{4.30}$$

with $\tilde{Z}_0^{(n)} = 1$. Also, define $\tilde{Z}_i$ by

$$\tilde{Z}_i = \sum_{k=1}^{\tilde{Z}_{i-1}} \tilde{\xi}_{i-1,k} + \tilde{Y}_i, \quad i = 1, 2, \ldots, \tag{4.31}$$

with $\tilde{Z}_0 = 1$. Observe that, by (4.27) and the above construction, $\{\tilde{Z}_i^{(n)}\}$ has the same distribution as $\{Z_i^{(n)}\}$ defined in (4.1) for each $n \geq 1$, and $\{\tilde{Z}_i\}$ has the same distribution as $\{Z_i\}$ defined in (1.1) (with the assumption (2.1)). Define a sequence of σ-fields by

$$\mathcal{G}_i = \sigma\{U_{k-1,j}, V_k, 1 \leq k \leq i, j \geq 1\}, \quad i \geq 1. \tag{4.32}$$

By a result of Seneta (1970), there exists a positive r.v. W_1 such that

$$m^{-n}\tilde{Z}_n \to W_1 \quad \text{a.s. as} \quad n \to \infty, \tag{4.33}$$

for $\tilde{Z}_n$ defined in (4.31).

Next, we state two lemmas for the array $\{\tilde{Z}_i^{(n)}\}$ in (4.30) which will be used in the proof of Theorem 4.3 below.

Lemma 4.2. For the process $\{\tilde{Z}_i^{(n)}\}$ in (4.30), assume that $\mu_n \to m > 1$ and $\lambda_n \to \lambda$ as $n \to \infty$. Furthermore, assume that (4.28) holds. Then, for W_1 in (4.33),

$$\mu_n^{-n}\tilde{Z}_n^{(n)} \xrightarrow{P} W_1 \quad \text{as} \quad n \to \infty. \tag{4.34}$$

Moreover,

$$\mu_n^{-n}\sum_{i=1}^{n} \tilde{Z}_{i-1}^{(n)} \xrightarrow{P} (m-1)^{-1}W_1 \quad \text{as} \quad n \to \infty. \tag{4.35}$$

Lemma 4.3. For the process $\{\tilde{Z}_i^{(n)}\}$ in (4.30), assume that condition (C-1) in (4.2) holds with $m > 1$. Then, as $n \to \infty$,

$$(m-1)^{\frac{1}{2}} \sum_{j=1}^{n} m^{-j/2} \left[(\tilde{Z}_{n-j+1}^{(n)} - \mu_n \tilde{Z}_{n-j}^{(n)} - \tilde{Y}_{n-j+1}^{(n)})/(\tilde{Z}_{n-j}^{(n)} + 1)^{\frac{1}{2}}\right] \xrightarrow{D} N(0, \sigma^2).$$

Lemma 4.2 shows that a result similar to (4.33) holds for the array $\{\tilde{Z}_i^{(n)}\}$ in (4.30) as well. The proof of Lemma 4.2 is quite non-trivial and is given in the Section 6. Lemma 4.3

is an array version of Corollary 3.3 of Wei and Winnicki (1989) and can be proved using the same arguments as in their paper. Hence we omit its proof.

The proof of Theorem 4.3 is given next. The method of proof of Theorem 4.3 is similar to that of Theorem 3.5 of Wei and Winnicki (1989) for the model (1.1), although here one needs to work harder because of the array structure.

Proof of Theorem 4.3. For the process $\{\tilde{Z}_i^{(n)}\}$ in (4.30), let

$$\tilde{\mu}_n = (\sum_{i=1}^{n} \tilde{Z}_{i-1}^{(n)})^{-1} \sum_{i=1}^{n} (\tilde{Z}_i^{(n)} - \tilde{Y}_i^{(n)})$$

and

$$\tilde{\nu}_n = (\sum_{i=1}^{n} \tilde{Z}_{i-1}^{(n)})^{\frac{1}{2}} (\tilde{\mu}_n - \mu_n). \tag{4.36}$$

Since $\{\tilde{Z}_i^{(n)}\}$ and $\{Z_i^{(n)}\}$ defined in (4.1) have the same distribution for each $n \geq 1$, we have that $\tilde{\nu}_n$ has the same distribution as ν_n defined in (4.4). Therefore, it suffices to show that

$$\tilde{\nu}_n \xrightarrow{D} N(0, \sigma^2) \quad \text{as} \quad n \to \infty. \tag{4.37}$$

To this end, let $\tilde{\epsilon}_i^{(n)} = (\tilde{Z}_i^{(n)} - \mu_n \tilde{Z}_{i-1}^{(n)} - \tilde{Y}_i^{(n)})$. Use (4.36) to write

$$\begin{aligned}
\tilde{\nu}_n &= (\sum_{i=1}^{n} \tilde{Z}_{i-1}^{(n)})^{-\frac{1}{2}} \sum_{i=1}^{n} \tilde{\epsilon}_i^{(n)} \\
&= (\sum_{i=1}^{n} \tilde{Z}_{i-1}^{(n)})^{-\frac{1}{2}} \left\{ \sum_{i=1}^{n} \left[(\tilde{Z}_{i-1}^{(n)} + 1)^{\frac{1}{2}} - (m^{(i-1)} W_1)^{\frac{1}{2}} \right] \tilde{\epsilon}_i^{(n)} / (\tilde{Z}_{i-1}^{(n)} + 1)^{\frac{1}{2}} \right. \\
&\qquad \left. + W_1^{\frac{1}{2}} \sum_{i=1}^{n} m^{(i-1)/2} \tilde{\epsilon}_i^{(n)} / (\tilde{Z}_{i-1}^{(n)} + 1)^{\frac{1}{2}} \right\}
\end{aligned} \tag{4.38}$$

where W_1 is the r.v. defined in Lemma 4.2 above. Observe that (4.28) implies

$$(\mu_n / m)^n \to 1 \quad \text{as} \quad n \to \infty. \tag{4.39}$$

Therefore, by (4.39), (4.35) and (4.38) it suffices to show that

$$\sum_{i=1}^{n} \left[(\tilde{Z}_{i-1}^{(n)} + 1)^{\frac{1}{2}} - (m^{(i-1)} W_1)^{\frac{1}{2}} \right] \tilde{\epsilon}_i^{(n)} / (\tilde{Z}_{i-1}^{(n)} + 1)^{\frac{1}{2}} = o_p(m^{\frac{n}{2}}) \tag{4.40}$$

and

$$(m-1)^{\frac{1}{2}} \sum_{i=1}^{n} m^{-(n-i+1)/2} \tilde{\epsilon}_i^{(n)} / (\tilde{Z}_{i-1}^{(n)} + 1)^{\frac{1}{2}} \xrightarrow{D} N(0, \sigma^2). \tag{4.41}$$

But, (4.41) follows from Lemma 4.3 by setting $n-i+1=j$ and summing over j. As for (4.40), use $\tilde{Z}_i$ in (4.31) to rewrite the left side of (4.40) as

$$\sum_{i=1}^{n} m^{(i-1)/2}\left\{\left[(\tilde{Z}_{i-1}^{(n)}+1)/m^{(i-1)}\right]^{\frac{1}{2}}-\left[(\tilde{Z}_{i-1}+1)/m^{(i-1)}\right]^{\frac{1}{2}}\right\}\tilde{\epsilon}_i^{(n)}/(\tilde{Z}_{i-1}^{(n)}+1)^{\frac{1}{2}}$$

$$+\sum_{i=1}^{n} m^{(i-1)/2}\{[(\tilde{Z}_{i-1}+1)/m^{(i-1)}]^{\frac{1}{2}}-W_1^{\frac{1}{2}}\}\tilde{\epsilon}_i^{(n)}/(\tilde{Z}_{i-1}^{(n)}+1)^{\frac{1}{2}}$$

$= (I)+(II)$, say.

$$\tag{4.42}$$

Consider (II) first. Apply Cauchy-Schwarz inequality to get

$$|(II)| \leq A_n^{\frac{1}{2}} B_n^{\frac{1}{2}} \tag{4.43}$$

where

$$A_n = \sum_{i=1}^{n} m^{(i-1)/2}\{[(\tilde{Z}_{i-1}+1)/m^{(i-1)}]^{\frac{1}{2}}-W_1^{\frac{1}{2}}\}^2$$

and

$$B_n = \sum_{i=1}^{n} m^{(i-1)/2}[\tilde{\epsilon}_i^{(n)}]^2/(\tilde{Z}_{i-1}^{(n)}+1). \tag{4.44}$$

By (4.33) [or Heyde (1970), Theorem 3] we have that

$$A_n = o(\sum_{i=1}^{n} m^{(i-1)/2}) = o(m^{\frac{n}{2}}) \quad \text{a.s.} \tag{4.45}$$

Moreover, from the definition of $\tilde{\epsilon}_i^{(n)}$ and (4.30) we get by conditioning w.r.t. $\mathcal{G}_i$ in (4.32) that

$$E(B_n) \leq \sigma_n^2 \sum_{i=1}^{n} m^{(i-1)/2} = O(m^{\frac{n}{2}}), \tag{4.46}$$

since $\sigma_n^2 \to \sigma^2$. Hence

$$B_n = O_p(m^{\frac{n}{2}}). \tag{4.47}$$

From (4.43), (4.45) and (4.47),

$$|(II)| = o_p(m^{\frac{n}{2}}). \tag{4.48}$$

As for (I) in (4.42), use Cauchy-Schwarz inequality once again to get

$$|(I)| \leq C_n^{\frac{1}{2}} B_n^{\frac{1}{2}}, \tag{4.49}$$

where B_n is as in (4.44) and

$$C_n = \sum_{i=1}^{n} m^{(i-1)/2}\{[(\tilde{Z}_{i-1}^{(n)}+1)/m^{(i-1)}]^{\frac{1}{2}} - [(\tilde{Z}_{i-1}+1)/m^{(i-1)}]^{\frac{1}{2}}\}^2. \tag{4.50}$$

Now, choose n large enough such that $\mu_n > 1$. By the inequality $|\sqrt{x+1}-\sqrt{y+1}|^2 \leq |x-y|$, for $x, y \geq 0$, arguments similar to (6.3) in Section 6 and that $E(\tilde{Z}_i) \leq Km^i$ for some $K > 0$,

$$\begin{aligned}
E|C_n| &\leq \sum_{i=1}^{n} m^{-(i-1)/2} E|\tilde{Z}_{i-1}^{(n)} - \tilde{Z}_{i-1}| \\
&\leq \sum_{i=1}^{n} m^{-(i-1)/2}\{c_n(\sum_{j=0}^{i-2}\mu_n^j) + d_n(\sum_{j=1}^{i}\mu_n^{j-1}E(\tilde{Z}_{i-j-1}))\} \\
&\leq c_n(\mu_n-1)^{-1}[(\mu_n/m)\vee 1]^{n-1}\sum_{i=1}^{n} m^{(i-1)/2} + Km^{-1}d_n n[(\mu_n/m)\vee 1]^{n-1}\sum_{i=1}^{n} m^{(i-1)/2} \\
&= o(m^{\frac{n}{2}}),
\end{aligned} \tag{4.51}$$

because (4.39) holds, $c_n \to 0$ and $nd_n \to 0$ (see (6.6) in Section 6). Here $x \vee y$ denotes max(x,y). Therefore, by (4.49), (4.47) and (4.51) we have

$$|i_a| = o_p(m^{\frac{n}{2}}). \tag{4.52}$$

Assertion (4.40) now follows from (4.42), (4.48) and (4.52). From (4.40), (4.41) and (4.38), we have that (4.37) holds and hence the conclusion of the theorem. ∎

5. Asymptotic Validity of Bootstrap.

Proof of Theorem 2.1. Given a sample realization of $\{(Z_i, Y_i), i = 1, 2, \ldots, n\}$, observe that the bootstrap process $\{Z_i^*\}$ defined in (2.2) is an array process defined in (4.1) with

$$\mu_n = \tilde{m}_n, \quad \sigma_n^2 = Var^*(\xi_{1,1}^*), \quad \lambda_n = \hat{\lambda}_n \quad \text{and} \quad b_n^2 = Var^*(Y_1^*), \tag{5.1}$$

where $\tilde{m}_n$ is defined by (2.5) and $\hat{\lambda}_n$ is as in (1.2). For the power series distributions in (2.1), we have already seen that $m = f(\theta), \lambda = g(\phi), \sigma^2 = \theta f'(\theta)$ and $b^2 = \phi g'(\phi)$, where f and g are appropriately defined strictly increasing and smooth functions (see the paragraph below (2.1)). From this, since $\tilde{m}_n \to m$ a.s., for all $m > 0$ (see Proposition 3.1, displays (3.7) and (3.8)) and $\hat{\lambda}_n \to \lambda$ a.s. we have that for each $\omega \in \{\tilde{m}_n \to m \text{ and } \hat{\lambda}_n \to \lambda\}$,

$$\text{(C-1) in (4.2) is satisfied for } \{Z_i^*\} \text{ in (2.2) for all } m > 0. \tag{5.2}$$

Rest of the proof is divided into two cases.

Case $m = 1$: For this case, we will apply Theorem 4.1. By (3.8) of Proposition 3.1 we have that $n(\tilde{m}_n - 1) \to 0$ a.s.. Let $\Omega = \{\tilde{m}_n \to 1, \hat{\lambda}_n \to \lambda, \text{ and } n(\tilde{m}_n - 1) \to 0\}$. Then, clearly $P(\Omega) = 1$. For each $\omega \in \Omega$, the first two conditions of Theorem 4.1 and condition (4.5) (with $\alpha = 0$) are easily satisfied. Moreover, it is not hard to show that $n^{-\frac{1}{2}} E^*|\xi_{1,1}^* - \tilde{m}_n(\omega)|^3 \to 0$ as $n \to \infty$, which implies condition (4.6) of Theorem 4.1 (see Sriram (1992), display (4.7), for instance). Therefore, by Theorem 4.1, for each $\omega \in \Omega$,

$$\sup_{-\infty < x < \infty} |P^*(V_n^* \le x) - P(\nu(0, \lambda, \sigma^2) \le x)| \to 0 \quad \text{as} \quad n \to \infty, \tag{5.3}$$

where ν is defined in (4.8). But, $\nu(0, \lambda, \sigma^2)$ has the same distribution as V defined in (1.4). Hence, (2.6) follows from (1.4) and (5.3) for the case $m = 1$.

Case $m \neq 1$: Here, we will apply Theorem 4.2 for the case $m < 1$ and Theorem 4.3 for the case $m > 1$. For the case $m < 1$, let $\Omega_1 = \{\tilde{m}_n \to m, \hat{\lambda}_n \to \lambda\}$. Clearly, $P(\Omega_1) = 1$ and as before all the conditions of Theorem 4.2 are satisfied for each $\omega \in \Omega_1$ (note that it is easy to show that $E^*(\xi_{1,1}^*)^3 \le B$, for some $B > 0$, for all $n \ge 1$). Hence, (2.6) follows easily from Theorem 4.2 for the case $m < 1$. For the case $m > 1$, let $\Omega_2 = \{\tilde{m}_n \to m, \hat{\lambda}_n \to \lambda$ and $n(\tilde{m}_n - m) \to 0\}$. Note that from (3.9) we have that $n(\tilde{m}_n - m) \to 0$, if $m > 1$ (see Proposition 3.1). For $\omega \in \Omega_2$, argue as before and use Theorem 4.3 to conclude that (2.6) follows for the case $m > 1$. Hence the Theorem 2.1. ∎

6. Proof of Lemma 4.2.

Consider the processes defined in (4.30) and (4.31). In view of (4.33), for the assertion (4.34) it suffices to show that

$$E|(\tilde{Z}_n^{(n)}/\mu_n^n) - (\tilde{Z}_n/m^n)| \to 0 \quad \text{as} \quad n \to \infty. \tag{6.1}$$

Recall that $E\tilde{\xi}_{1,1}^{(n)} = \mu_n, E\tilde{\xi}_{1,1} = m$ and let

$$c_n = E|\tilde{Y}_1^{(n)} - \tilde{Y}_1| \quad \text{and} \quad d_n = E|\tilde{\xi}_{1,1}^{(n)} - \tilde{\xi}_{1,1}|. \tag{6.2}$$

Now use (4.30), (4.31), conditioning w.r.t. $\{\mathcal{G}_i\}$ defined in (4.32) and (6.2) to write

$$\begin{aligned}
E|\tilde{Z}_n^{(n)} - \tilde{Z}_n| &\le E|\tilde{Y}_1^{(n)} - \tilde{Y}_1| + E|\sum_{k=1}^{\tilde{Z}_{n-1}^{(n)}} \tilde{\xi}_{n-1,k}^{(n)} - \sum_{k=1}^{\tilde{Z}_{n-1}} \tilde{\xi}_{n-1,k}| \\
&\le c_n + \mu_n E|\tilde{Z}_{n-1}^{(n)} - \tilde{Z}_{n-1}| + d_n E(\tilde{Z}_{n-1}) \\
&\ \ \vdots \\
&\le c_n(\sum_{j=0}^{n-1} \mu_n^j) + \mu_n^n E|\tilde{Z}_0^{(n)} - \tilde{Z}_0| + d_n \sum_{j=0}^{n-1} \mu_n^j E(\tilde{Z}_{n-j-1}) \\
&= c_n(\sum_{j=0}^{n-1} \mu_n^j) + d_n \sum_{j=1}^{n} \mu_n^{j-1} E(\tilde{Z}_{n-j}), \qquad (6.3)
\end{aligned}$$

where we used $\tilde{Z}_0^{(n)} = \tilde{Z}_0 = 1$. By (4.39) (recall that (4.28) implies (4.39)) and $E\tilde{Z}_n = O(m^n)$ we have that

$$\begin{aligned}
E|(\tilde{Z}_n^{(n)}/\mu_n^n) - (\tilde{Z}_n/m^n)| &\le \mu_n^{-n} E|\tilde{Z}_n^{(n)} - \tilde{Z}_n| + |1 - (m/\mu_n)^n| E(\tilde{Z}_n/m^n) \\
&= \mu_n^{-n} E|\tilde{Z}_n^{(n)} - \tilde{Z}_n| + o(1). \qquad (6.4)
\end{aligned}$$

Since $\mu_n \to m > 1$, let n be large enough such that $\mu_n^{-1} < 1$. But, from (6.3) and since $E\tilde{Z}_{n-j} \le K_1 m^{n-j}$ for some $K_1 > 0$,

$$\begin{aligned}
\mu_n^{-n} E|\tilde{Z}_n^{(n)} - \tilde{Z}_n| &\le c_n \sum_{j=1}^{n} \mu_n^{-j} + d_n \mu_n^{-n} \sum_{j=1}^{n} \mu_n^{j-1} E(\tilde{Z}_{n-j}) \\
&\le c_n \sum_{j=1}^{\infty} \mu_n^{-j} + d_n K_1 (m/\mu_n)^n m^{-1} \sum_{j=1}^{n} (\mu_n/m)^{j-1} \\
&\le (\mu_n^{-1} - 1)^{-1} c_n + K_1 (m/\mu_n)^n m^{-1} \left[(\mu_n/m) \vee 1\right]^n (n d_n) \\
&\to 0 \qquad (6.5)
\end{aligned}$$

by $\mu_n \to m > 1, (\mu_n/m)^n \to 1$, and provided we show that

$$c_n \to 0 \quad \text{and} \quad n d_n \to 0 \quad \text{as} \quad n \to \infty. \qquad (6.6)$$

We will show below that $n d_n \to 0$ as $n \to \infty$. Similar, but simpler arguments show that $c_n \to 0$ as $n \to \infty$. To this end, observe that by (4.28) it suffices to show that

$$d_n = O(|\mu_n - m|) \quad \text{as} \quad n \to \infty. \qquad (6.7)$$

Recall the definitions of $\tilde{\xi}_{1,1}^{(n)}$ and $\tilde{\xi}_{1,1}$ from (4.30) and (4.31), respectively. Let $P_{\theta_n}[\tilde{\xi}_{1,1}^{(n)} = i] = a(i)\theta_n^i / A(\theta_n) = p_i(\theta_n)$ and $P_\theta[\tilde{\xi}_{1,1} = i] = a(i)\theta^i / A(\theta) = p_i(\theta)$ for $i \ge 0$. Also, for $i \ge 0$, let

$$q_i(\theta_n) = \sum_{j=0}^{i} p_j(\theta_n), \qquad q_i(\theta) = \sum_{j=0}^{i} p_j(\theta) \qquad (6.8)$$

and

$$r_i(\theta_n) = 1 - q_i(\theta_n) \quad \text{and} \quad r_i(\theta) = 1 - q_i(\theta).$$

Then, by the definition of $F_{\theta_n}^{-1}$ and F_θ^{-1} in (4.29) we have that

$$\begin{aligned} d_n = E|\tilde{\xi}_{1,1}^{(n)} - \tilde{\xi}_{1,1}| &= \int_0^1 |F_{\theta_n}^{-1}(x) - F_\theta^{-1}(x)|dx \\ &= \sum_{i=0}^{\infty} |q_i(\theta_n) - q_i(\theta)| \\ &= \sum_{i=0}^{\infty} |r_i(\theta_n) - r_i(\theta)|. \end{aligned} \tag{6.9}$$

Since $\mu_n = E\tilde{\xi}_{1,1}^{(n)} = f(\theta_n)$ and $m = E\tilde{\xi}_{1,1} = f(\theta)$ where f is a strictly increasing and smooth function (see Section 2), in order to show (6.7) it suffices to show that

$$\limsup_{n\to\infty} |\theta_n - \theta|^{-1} \sum_{i=0}^{\infty} |r_i(\theta_n) - r_i(\theta)| < \infty. \tag{6.10}$$

Now, as $n \to \infty$, $\theta_n \to \theta$ since $\mu_n \to m$. Let δ be a small positive number such that $I_\delta = (\theta - \delta, \theta + \delta) \in (0, \theta^*)$. Recall that θ^* is the radius of convergence of $\sum_{u=0}^{\infty} a(u)\theta^u$. Then, there exists N such that $\theta_n \in I_\delta$ for all $n \geq N$. Let $n \geq N$. Since $r_i(\theta)$ defined in (6.8) is a smooth function of θ, by the Mean value theorem

$$|r_i(\theta_n) - r_i(\theta)| \leq |\theta_n - \theta| \max_{\theta_0 \in [\theta-\delta, \theta+\delta]} |r_i'(\theta_0)| \tag{6.11}$$

Now, since $r_i(\theta) = \sum_{j=i+1}^{\infty} p_j(\theta)$ we have that

$$\begin{aligned} r_i'(\theta) &= \sum_{j=i+1}^{\infty} [A(\theta)]^{-2} \left\{A(\theta) j\theta^{j-1} a(j) - A'(\theta)\theta^j a(j)\right\} \\ &= [A(\theta)]^{-1} \sum_{j=i+1}^{\infty} j\theta^{j-1} a(j) - [A(\theta)]^{-2} A'(\theta) \sum_{j=i+1}^{\infty} \theta^j a(j) \end{aligned}$$

which implies that with

$$\tilde{\theta} = \theta + \delta, \quad A^* = \max_{\theta_0 \in [\theta-\delta, \theta+\delta]} |A(\theta_0)|^{-1} \quad \text{and} \quad B^* = \max_{\theta_0 \in [\theta-\delta, \theta+\delta]} |A'(\theta_0)/A^2(\theta_0)|$$

$$\max_{\theta_0 \in [\theta-\delta, \theta+\delta]} |r_i'(\theta_0)| \leq A^* \sum_{j=i+1}^{\infty} j(\tilde{\theta})^{j-1} a(j) + B^* \sum_{j=i+1}^{\infty} (\tilde{\theta})^j a(j). \tag{6.12}$$

Therefore, from (6.12) and (6.11)

$$|\theta_n-\theta|^{-1}\sum_{i=0}^{\infty}|r_i(\theta_n)-r_i(\theta)| \le A^*\sum_{i=0}^{\infty}\sum_{j=i+1}^{\infty} j(\tilde\theta)^{j-1}a(j)+B^*A(\tilde\theta)E_{\tilde\theta}(\tilde\xi_{1,1})$$
$$= A^*\sum_{j=1}^{\infty} j^2 a(j)\tilde\theta^{j-1}+B^*A(\tilde\theta)E_{\tilde\theta}(\tilde\xi_{1,1})$$
$$<\infty.$$

The assertion (6.10) now follows from above arguments. Hence, $nd_n \to 0$, as $n\to\infty$. Thus, (4.34) follows from (6.1), (6.4), (6.5) and (6.6).

As for assertion (4.35), write

$$\mu_n^{-n}\sum_{i=1}^{n}\tilde Z_i^{(n)} = \mu_n^{-n}[\sum_{i=0}^{n-1}(\tilde Z_i^{(n)}-\tilde Z_i)] + (m/\mu_n)^n m^{-n}\sum_{i=0}^{n-1}\tilde Z_i. \tag{6.13}$$

Then, by (4.34) and since $(m/\mu_n)^n \to 1$, it suffices to show that

$$\mu_n^{-n}\sum_{i=1}^{n}E|\tilde Z_i^{(n)}-\tilde Z_i|\to 0. \tag{6.14}$$

For each $1\le i\le n$, argue as in (6.3) to get

$$\mu_n^{-n}\sum_{i=1}^{n}E|\tilde Z_i^{(n)}-\tilde Z_i| \le c_n\mu_n^{-n}\sum_{i=1}^{n}(\sum_{j=0}^{i-1}\mu_n^j)+d_n\mu_n^{-n}\sum_{i=1}^{n}\sum_{j=1}^{i}\mu_n^{j-1}E(\tilde Z_{i-j})$$
$$\to 0, \tag{6.15}$$

by arguments similar to (6.5) and (6.6). Hence the assertion (4.35) and the Lemma. ∎

Acknowledgements

The second author would like to thank the Department of Statistics, University of North Carolina at Chapel Hill for providing financial support during his visit there in Fall 1992.

REFERENCES

Athreya, K.B. and Fuh, C.D. (1992). Bootstrapping Markov chains. In *Exploring the Limits of Bootstrap,* R. LePage and L. Billard, Eds, 49-64, Wiley.

Babu, G.J. (1989). Applications of Edgeworth expansions to bootstrap - a review. In *Statistical Data Analysis and Inference*, Y. Dodge, Ed, 223-237, North Holland.

Basawa, I.V., Mallik, A.K., McCormick, W.P. and Taylor, R.L. (1989). Bootstrapping explosive autoregressive processes. *Ann. Statist.* 17, 1479-1486.

Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H. and Taylor, R.L. (1991). Bootstrapping unstable first order autoregressive processes. *Ann. Statist.* 19, 1098-1101.

Bhat, B.R. and Adke, S.R. (1982). Maximum likelihood estimation for branching processes with immigration. *Adv. Appl. Prob.* 13, 498-509.

Bose, A. (1988). Edgeworth correction by bootstrap in autoregression. *Ann. Statist.* 16, 1709-1722.

Chow, Y.S. and Teicher, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales.* Springer, New York.

Datta, S. (1992). On asymptotic properties of bootstrap for unstable AR(1) processes. Tech. Report 92-29, Depart. of Statist., Univ. of Georgia, Athens.

Datta, S. and McCormick, (1992). Bootstrap for a finite state Markov chain based on i.i.d. resampling. In *Exploring the Limits of Bootstrap,* R. LePage and L. Billard, Eds, 77-97, Wiley.

Datta, S. and McCormick, W.P. (1993). Regeneration based bootstrap for Markov chains. *Canadian J. Statist.* (to appear).

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7, 1-26.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer, New York.

Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application.* Academic, New York.

Heyde, C.C. (1970). Extension of a result of Seneta for the super-critical Galton-Watson process. *Ann. Math. Statist.* 41, 739-742.

Heyde, C.C. and Seneta, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Prob.* 9, 235-256.

Heyde, C.C. and Seneta, E. (1974). Notes on "Estimation theory for growth and immigration rates in a multiplicative process." *Appl. Prob.* 11, 572-577.

Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* 17, 1217-1241.

Lahiri, S.N. (1991). Second order optimality of stationary bootstrap. *Statist. Probab. Let.* 11, 335-341.

Lai, T.L. and Wei, C.Z. (1983). Lacunary and generalized linear processes. *Stoch. Processes Appl.*, 14, 187-199.

LeCam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. of California Publications in Statist.* 1, 277-330.

Liu, R.Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap*, R. LePage and L. Billard, Eds, 225-248, Wiley.

Politis, D. N. and Romano, J. P. (1992). A general resampling scheme for triangular arrays of α-mixing random variables with application to the problem of spectral density estimation. *Ann. Statist.* 20, 1985-2007.

Seneta, E. (1970). A note on the supercritical Galton-Watson process with immigration, *Math. Biosc.* 6, 305-311.

Sriram, T.N. (1992). Invalidity of bootstrap for critical branching processes with immigration. Under revision.

Sriram, T.N., Basawa, I.V., and Huggins, R.M. (1991). Sequential estimation for branching processes with immigration. *Ann. Statist.* 19, 2232-2243.

Wei, C.Z. (1985). Asymptotic properties of least-squares estimates in stochastic regression models. *Ann. Statist.*, 13, 1498-1508.

Wei, C.Z. and Winnicki, J. (1989). Some asymptotic results for the branching process with immigration. *Stochastic Processes. 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.