The Annals of Statistics
1994, Vol. 22, No. 2, 1013–1023

INVALIDITY OF BOOTSTRAP FOR CRITICAL BRANCHING PROCESSES WITH IMMIGRATION[1]

BY T.N. SRIRAM

University of Georgia

This article considers a case of parametric bootstrap when the observations consist of generation sizes of the branching process with immigration together with the immigration component of each generation. Suppose we estimate the offspring mean m by the maximum likelihood estimator (m.l.e.). It is then shown that the bootstrap version of the standardized m.l.e. does not have the same limiting distribution as the standardized m.l.e., under the assumption that $m = 1$ (critical case). In other words, the asymptotic validity does not hold for the parametric bootstrap in the critical case. In fact, given the sample, the value of the conditional distribution function of the bootstrap version of standardized m.l.e. defines a sequence of random variables whose limit (in distribution) is also shown to be a random variable, when $m = 1$. The approach used here is via a sequence of branching processes for which a general weak convergence [in $D^+[0,\infty)$] result is established using operator semigroup convergence theorems.

1. Introduction. The branching process with immigration can be defined recursively by the following equation

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

The random variable Z_i can be interpreted as the size of the ith generation of a population, where $\xi_{i-1,k}$ is the offspring size of the kth individual belonging to the $(i-1)$th generation and Y_i is the number of immigrants contributing to the population's ith generation. Assume that $\{\xi_{i-1,k}\}$ and $\{Y_i\}$ are independent sequences of i.i.d. nonnegative, integer-valued random variables (r.v.'s) with finite means m and λ and finite variances σ^2 and b^2, respectively. Throughout the paper, we assume that the initial size $Z_0 = 1$ and the offspring and immigration r.v.'s ξ and Y have power series distributions with probability mass functions (p.m.f.'s) p_θ and q_ϕ given, respectively, by

$$\begin{aligned} p_\theta(u) &\equiv P[\xi = u] = a(u)\theta^u/A(\theta), \qquad u = 0, 1, \ldots \\ q_\phi(y) &\equiv P[Y = y] = b(y)\phi^y/B(\phi), \qquad y = 0, 1, \ldots. \end{aligned} \tag{1.2}$$

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 convergence of the two power series.

Received March 1992; revised August 1993.

[1]The revised version of this paper was prepared while the author was visiting the Department of Statistics, University of North Carolina at Chapel Hill.

AMS 1991 *subject classifications.* Primary 60J80, 62G09.

Key words and phrases. Critical branching processes, asymptotic invalidity, operator semigroup convergence theorems, diffusion process, generator, parametric bootstrap.

Suppose that a sample $\{(Z_i, Y_i),\ i = 1, 2, \dots, n\}$ is available. Then a natural estimator of the offspring mean m is given by

$$\hat{m}_n = \left[\sum_{i=1}^{n} Z_{i-1}\right]^{-1} \sum_{i=1}^{n} (Z_i - Y_i). \tag{1.3}$$

Note that for power series offspring and immigration distributions given in (1.2), the estimators $\hat{m}_n$ and $\hat{\lambda}_n = n^{-1}\sum_{i=1}^{n} Y_i$ are the maximum likelihood estimators of m and λ, respectively [see, e.g., Bhat and Adke (1981)]. In the literature, the estimation of m by $\hat{m}_n$, based on the full information on both generation sizes $\{Z_i\}$ and the immigration sizes $\{Y_i\}, i = 1, 2, \dots, n$, has been considered by Venkataraman and Nanthi (1982). Using only the partial information on $\{Z_i\}$ alone, it is possible to estimate m and study the properties of the estimators [see Heyde and Seneta (1972, 1974), Wei and Winnicki (1990) and the references therein].

It is well known that the limit distribution of $\hat{m}_n$ is subject to a threshold theorem, where m plays the crucial role of a threshold parameter. In particular, it can be shown that the limit distribution of the pivot

$$V_n = \left[\sum_{i=1}^{n} Z_{i-1}\right]^{1/2} (\hat{m}_n - m) \tag{1.4}$$

is normal if $m \neq 1$ and nonnormal if $m = 1$ (critical case). See Sriram, Basawa and Huggins (1991) for the cases $m < 1$ and $m = 1$.

In this paper, we introduce a parametric bootstrap for branching processes with immigration. We restrict our attention to the critical case, since the limit distribution of the pivot V_n is nonnormal in this case and is therefore of special interest in considering the bootstrap approximation for the distribution of V_n. The performance of the parametric bootstrap for the cases $m < 1$ and $m > 1$ is of interest as well, but these cases will be considered elewhere.

One way to assess the performance of our parametric bootstrap is to check whether the conditional limit distribution of the bootstrap pivot, say, V_n^* [see (2.4)], is the same as the limit distributions of V_n, when $m = 1$, as $n \to \infty$. In other words, does the asymptotic validity hold for the parametric bootstrap in the critical case? We shall show that the parametric bootstrap is asymptotically invalid, at $m = 1$. In fact, given the sample, the value of the conditional distribution function of V_n^* defines a sequence of r.v.'s whose limit (in distribution) is also shown to be a random variable, when $m = 1$. It is, however, conjectured that the parametric bootstrap proposed here is asymptotically valid for the cases $m < 1$ and $m > 1$.

One of the main causes of failure of the parametric bootstrap at $m = 1$ is the fact that the limit distributions of V_n are different for the two cases $m \neq 1$ and $m = 1$. The other reason, which seems to be the crucial one, concerns the rate of convergence of $\hat{m}_n$ to 1. The latter reason is explained further in a concluding remark (see Section 4).

In a similar situation, it has been shown recently by Basawa, Mallik, McCormick, Reeves and Taylor (1991) that for a first-order autoregressive process the standard bootstrap least squares estimator of the autoregressive parameter is asymptotically invalid, when the parameter is 1. This paper shows that a similar phenomenon occurs in our problem as well. Another instance of the invalidity of the naive bootstrap has been discussed by Athreya (1987) in the context of estimating the mean of a population when the variance is infinite.

In Section 2 we introduce a parametric bootstrap for the branching process with immigration and state the main theorem concerning the asymptotic invalidity of bootstrap in the critical case. In Section 3 we state a general result about convergence of a sequence of branching processes and prove it in the Appendix. In Section 4, we prove the main theorem using the results obtained in Section 3.

2. Parametric bootstrap. Recall the power series distributions given in (1.2). By the notation introduced above, we have the means $m = E(\xi) = \theta A'(\theta)/A(\theta)$ and $\lambda = E(Y) = \phi B'(\phi)/B(\phi)$ and the variances $\sigma^2 = \mathrm{var}(\xi) = \theta \partial m/\partial\theta$ and $b^2 = \mathrm{var}(Y) = \phi\, \partial\lambda/\partial\phi$. Here, prime denotes first derivative and ∂ denotes 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. Using the fact that $\theta = f^{-1}(m)$ and $\phi = g^{-1}(\lambda)$, we denote the distribution function of V_n defined in (1.4) by

$$H_n(m, \lambda, x) = P\{V_n \le x\}, \tag{2.1}$$

for $x \in (-\infty, \infty)$, $m > 0$ and $\lambda > 0$.

Now, collect a sample $\{(Z_i, Y_i),\ i = 1, 2, \dots, n\}$ and estimate m by $\widehat{m}_n$ in (1.3), and λ by $\widehat{\lambda}_n = n^{-1}\sum_{i=1}^n Y_i = \overline{Y}$. Replace θ and ϕ in (1.2) by their respective estimates $\widehat{\theta}_n = f^{-1}(\widehat{m}_n)$ and $\widehat{\phi}_n = g^{-1}(\widehat{\lambda}_n)$. Call the resulting estimated p.m.f.'s $p_{\hat{\theta}_n}$ and $q_{\hat{\phi}_n}$, respectively. Note that the values of $p_{\hat{\theta}_n}(u)$ and $q_{\hat{\phi}_n}(y)$ are known for each u and y since the data values are known. Conditional on $\{(Z_i, Y_i),\ i = 1, 2, \dots, n\}$, let $\{\xi^*_{i,j}\}$ be a sequence of i.i.d. r.v.'s having p.m.f. $p_{\hat{\theta}_n}$, and let $\{Y^*_i\}$ be a sequence of i.i.d. r.v.'s having p.m.f. $q_{\hat{\phi}_n}$. The parametric bootstrap sample $\{(Z^*_i, Y^*_i),\ i = 1, 2, \dots n\}$ is then obtained recursively from the relation

$$Z^*_i = \sum_{k=1}^{Z^*_{i-1}} \xi^*_{i-1,k} + Y^*_i, \qquad i = 1, 2, \dots, \tag{2.2}$$

with $Z^*_0 = 1$. The parametric bootstrap estimator of m is then given by

$$\widehat{m}^*_n = \left[\sum_{i=1}^n Z^*_{i-1}\right]^{-1} \sum_{i=1}^n (Z^*_i - Y^*_i). \tag{2.3}$$

The parametric bootstrap analogue of the pivot V_n defined in (1.4) is given by

$$V_n^* = \left[\sum_{i=1}^{n} Z_{i-1}^*\right]^{1/2} \left(\widehat{\widehat{m}}_n - \widehat{m}_n\right). \tag{2.4}$$

Now, observe that

$$P\{V_n^* \le x \,|\, (Z_i, Y_i),\ i = 1, 2, \dots, n\} = H_n\left(\widehat{m}_n, \widehat{\lambda}_n, x\right), \tag{2.5}$$

where H_n is defined in (2.1). Henceforth, suppose that the true model is (1.1) with $m = 1$. We are interested in the limit of $H_n(\widehat{m}_n, \widehat{\lambda}_n, x)$, for every fixed $x \in (-\infty, \infty)$.

For V_n defined in (1.4), it is shown in Sriram, Basawa and Huggins (1991) that, when $m = 1$,

$$V_n \to_{\mathfrak{D}} \frac{\{X(1) - \lambda\}}{\left\{\int_0^1 X(t)\,dt\right\}^{1/2}} = V \quad \text{as } n \to \infty, \tag{2.6}$$

where $\{X(t)\}$ is a nonnegative diffusion process with generator $Ah(x) = \lambda h'(x) + (1/2)x\sigma^2 h''(x)$, for $h \in C_c^\infty[0, \infty)$, and is obtained as a weak limit of the process $X_n(t) = Z_{[nt]}/n$, as $n \to \infty$. Here $C_c^\infty[0, \infty)$ is the space of infinitely differentiable functions on $[0, \infty)$ which have compact supports, and $''$ denotes second derivative. Arguing as in Sriram, Basawa and Huggins (1989), it is easy to show that

$$n(\widehat{m}_n - 1) \to_{\mathfrak{D}} \frac{\{X(1) - \lambda\}}{\left\{\int_0^1 X(t)\,dt\right\}} = V_0. \tag{2.7}$$

If the parametric bootstrap were to be asymptotically valid here, then we should have

$$\sup_{-\infty < x < \infty} \left|H_n(\widehat{m}_n, \widehat{\lambda}_n, x) - P(V \le x)\right| \to 0$$

almost surely, as $n \to \infty$, where $H_n(\widehat{m}_n, \widehat{\lambda}_n, x)$ is defined in (2.5) and V is defined in (2.6). However, we will show that this is not the case. There is a random limit. To describe this precisely, we proceed as follows. Let $\widetilde{\mathfrak{X}}_\alpha$ be a nonnegative diffusion process with (infinitesimal) generator

$$\mathcal{A}_\alpha h(x) = \alpha x h'(x) + \lambda_0 h'(x) + \tfrac{1}{2}x\sigma_0^2 h''(x), \qquad h \in C_c^\infty[0, \infty), \tag{2.8}$$

where λ_0 and σ_0^2 are positive and finite real numbers and α is a finite real number. For a definition of an infinitesimal generator, such as $\mathcal{A}_\alpha$ defined in (2.8), and an excellent presentation of the theory of weak convergence, the reader is referred to Ethier and Kurtz (1986).

Define

$$\nu\left(\alpha, \lambda_0, \sigma_0^2\right) = \left\{\int_0^1 \widetilde{\mathfrak{X}}_\alpha(t)\,dt\right\}^{-1/2} \left[\widetilde{\mathfrak{X}}_\alpha(1) - \lambda_0\right] - \alpha\left\{\int_0^1 \widetilde{\mathfrak{X}}_\alpha(t)\,dt\right\}^{1/2} \tag{2.9}$$

and let

$$H(\alpha,\lambda_0,\sigma_0^2,x)=P\Big[\nu(\alpha,\lambda_0,\sigma_0^2)\le x\Big]. \tag{2.10}$$

For $\widetilde{\mathcal{X}}_\alpha$ defined above, let $\widetilde{\mathcal{X}}_\alpha(0)=0$ for every real number α. The following arguments show that $H(\alpha,\lambda_0,\sigma_0^2,x)$ is continuous in α for every fixed λ_0, σ_0^2 and x. Suppose $\{\alpha_n\}$ is sequence of real numbers such that $\alpha_n\to\alpha$ as $n\to\infty$, for some real number α. Then, for $h\in C_c^\infty[0,\infty)$, it is easy to show that

$$\sup_{-\infty<x<\infty}|\mathcal{A}_{\alpha_n}h(x)-\mathcal{A}_\alpha h(x)|\to 0$$

as $n\to\infty$, where $\{\mathcal{A}_{\alpha_n}\}$ is a sequence of generators corresponding to $\{\widetilde{\mathcal{X}}_{\alpha_n}\}$ and defined as in (2.8) with α replaced by α_n. Therefore, from Ethier and Kurtz [(1986), Theorem 2.1 of Chapter 8, Theorem 6.1 of Chapter 1 and Theorem 2.5 of Chapter 4] and the fact that $\widetilde{\mathcal{X}}_{\alpha_n}(0)=\widetilde{\mathcal{X}}_\alpha(0)=0$, we conclude that

$$\widetilde{\mathcal{X}}_{\alpha_n}\to_{\mathcal{D}}\widetilde{\mathcal{X}}_\alpha$$

as $n\to\infty$. From this and the continuous mapping theorem, it follows that $H(\alpha,\ \lambda_0,\ \sigma_0^2,\ x)$ is continuous in α for each fixed λ_0, σ_0^2 and x. Therefore, $\omega\to H(V_0(\omega),\lambda_0,\ \sigma_0^2,\ x)$ defines a random variable for V_0 defined in (2.7). Our main result is the following theorem.

THEOREM 2.1. *For every real number x, the sequence of r.v.'s $H_n(\widehat{m}_n,\widehat{\lambda}_n,\ x)$ converges in distribution to the r.v. $H(V_0,\lambda,\sigma^2,x)$, where H is defined in* (2.10) *and V_0 is defined in* (2.7).

Before we prove the main result we obtain a general result about convergence of a sequence of branching processes and a corollary. The corollary is then used to prove the main theorem.

3. Sequence of branching processes. Consider a sequence of branching processes $\{Z_i^{(n)}\}$ given by

$$Z_i^{(n)}=\sum_{k=1}^{Z_{i-1}^{(n)}}\xi_{i-1,k}^{(n)}+Y_i^{(n)},\qquad i=1,2,\dots, \tag{3.1}$$

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.

Let $\mathcal{X}_n(t)=Z_{[nt]}^{(n)}/n, t\ge 0$. It is clear that $\{\mathcal{X}_n\}$ is a sequence of random elements that take values in $D^+[0,\infty)$, which is the space of nonnegative functions on $[0,\infty)$ that are right continuous and have left limits. Using operator semigroup convergence theorems [see Ethier and Kurtz (1986)], we prove below a weak convergence theorem for the sequence of random functions $\{\mathcal{X}_n\}$. The following assumptions are made, as $n \to \infty$:

(3.2)
(C1) $\mu_n = 1 + \alpha n^{-1} + o(1/n)$;
(C2) $\sigma_n^2 \to \sigma_0^2$;
(C3) $\lambda_n \to \lambda_0$ and $b_n^2 \to b_0^2$;
(C4) for any sequence $\{x_n\}$ such that $x_n \to x$, $0 < x < \infty$, $\lim_{n\to\infty} \sigma_n^{-2} E|\xi_{1,1}^{(n)} - \mu_n|^2 I_{\{|\xi_{1,1}^{(n)} - \mu_n| \geq \varepsilon\sigma_n\sqrt{nx_n}\}} = 0$, for all $\varepsilon > 0$, where λ_0, σ_0^2 and b_0^2 are positive and finite real numbers and α is a finite real number.

THEOREM 3.1. *Assume conditions* (C1)–(C4) *in* (3.2). *Then* $\mathcal{X}_n \to_{\mathcal{D}} \widetilde{\mathcal{X}}_\alpha$, *where* $\widetilde{\mathcal{X}}_\alpha$ *is the diffusion process with generator defined in* (2.8).

The arguments used to prove Theorem 3.1 are similar to that of Theorem 1.3 in Ethier and Kurtz [(1986), Chapter 9]. We only indicate the necessary steps in the Appendix.

REMARK 3.1. There is a rather large literature on the diffusion approximation for a sequence of branching processes with or without immigration. See, for instance, Grimvall (1974) and some of the relevant references listed under "notes" at the end of Chapter 9 in Ethier and Kurtz (1986).

For model (1.1) with $m = 1$, Wei and Winnicki (1989) showed that $X_n \to_{\mathcal{D}} X$, where $X_n(t) = Z_{[nt]}/n$ and X is a nonnegative diffusion defined in (2.6). The proof of Wei and Winnicki (1989) uses operator semigroup convergence theorems as done here.

Let

$$\widehat{\mu}_n = \left[\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{-1} \sum_{i=1}^{n} \left(Z_i^{(n)} - Y_i^{(n)}\right) \tag{3.3}$$

and

$$\nu_n = \left[\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{1/2} (\widehat{\mu}_n - \mu_n). \tag{3.4}$$

COROLLARY 3.1. *Assume conditions* (C1)–(C4) *in* (3.2). *Then for* ν_n *defined in* (3.4) *and* $\nu(\alpha, \lambda_0, \sigma_0^2)$ *defined in* (2.9) *we have*

$$\nu_n \to_{\mathcal{D}} \nu(\alpha, \lambda_0, \sigma_0^2) \quad as\ n \to \infty. \tag{3.5}$$

PROOF. By (3.3) and (3.4),

$$\begin{aligned}
\nu_n &= \left[\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{-1/2} \left\{\sum_{i=1}^{n} \left(Z_i^{(n)} - \mu_n Z_{i-1}^{(n)} - Y_i^{(n)}\right)\right\} \\
&= \left[\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{-1/2} \left\{\sum_{i=1}^{n} \left(Z_i^{(n)} - Z_{i-1}^{(n)} - Y_i^{(n)}\right)\right\} - (\mu_n - 1)\left[\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{-1/2} \\
&= \left[n^{-2}\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{-1/2} \left\{n^{-1}\left(Z_n^{(n)} - Z_0^{(n)} - \sum_{i=1}^{n} Y_i^{(n)}\right)\right\} - n(\mu_n - 1)\left[n^{-2}\sum_{i=1}^{n} Z_{i-1}^{(n)}\right]^{1/2} \\
&= \left[\int_0^1 \mathcal{X}_n(t)\,dt\right]^{-1/2} \left[\mathcal{X}_n(1) - \mathcal{X}_n(0) - n^{-1}\sum_{i=1}^{n} Y_i^{(n)}\right] - n(\mu_n - 1)\left[\int_0^1 \mathcal{X}_n(t)\,dt\right]^{1/2} \\
&\to_{\mathcal{D}} \nu(\alpha, \lambda_0, \sigma_0^2),
\end{aligned}$$

where we used Theorem 3.1, the continuous mapping theorem, (C1) and the fact that $n^{-1}\sum_{i=1}^{n} Y_i^{(n)} \to \lambda_0$ in probability as $n \to \infty$, by (C3). □

4. Proof of the main result.

PROOF OF THEOREM 2.1. By (2.7) and the fact that $\hat{\lambda}_n = \overline{Y}_n \to \lambda$ a.s., we have

$$(4.1) \qquad \left(n(\hat{m}_n - 1), \hat{\lambda}_n\right) \to_{\mathcal{D}} (V_0, \lambda) \quad \text{as } n \to \infty.$$

Therefore, by Skorohod's theorem [see Billingsley (1979), Theorem 29.6], there exist random vectors $(\tilde{m}_n, \tilde{\lambda}_n)$ and $(\tilde{V}_0, \lambda)$ on a common probability space $(\tilde{\Omega}, \mathcal{F}, Q)$ such that the following hold: $(\tilde{m}_n, \tilde{\lambda}_n)$ has the same distribution as $(\hat{m}_n, \hat{\lambda}_n)$; $(\tilde{V}_0, \lambda)$ has the same distribution as (V_0, λ); and, for each $\tilde{\omega} \in \tilde{\Omega}$,

$$(4.2) \qquad n\left(\tilde{m}_n(\tilde{\omega}) - 1\right) \to \tilde{V}_0(\tilde{\omega}) \quad \text{and} \quad \tilde{\lambda}_n(\tilde{\omega}) \to \lambda$$

as $n \to \infty$. Now, for each $\tilde{\omega} \in \tilde{\Omega}$, estimate θ and ϕ in (1.2) by $\tilde{\theta}_n(\tilde{\omega}) = f^{-1}(\tilde{m}_n(\tilde{\omega}))$ and $\tilde{\phi}_n(\tilde{\omega}) = g^{-1}(\tilde{\lambda}_n(\tilde{\omega}))$, respectively, as done in bootstrapping. As in Section 2, call the resulting estimated p.m.f.'s $p_{\tilde{\theta}_n(\tilde{\omega})}$ and $q_{\tilde{\theta}_n(\tilde{\omega})}$, respectively. For each $\tilde{\omega}$ or, equivalently, for each pair $(\tilde{m}_n(\tilde{\omega}), \tilde{\lambda}_n(\tilde{\omega}))$, a sequence of branching processes $\{\tilde{Z}_i^{(n)}\}$ can then be constructed recursively from the relation

$$(4.3) \qquad \tilde{Z}_i^{(n)} = \sum_{k=1}^{\tilde{Z}_{i-1}^{(n)}} \tilde{\xi}_{i-1,k}^{(n)} + \tilde{Y}_i^{(n)}, \qquad i = 1, 2, \ldots,$$

where (for each $\tilde{\omega} \in \tilde{\Omega}$ and n) $\{\tilde{\xi}_{i,j}^{(n)}\}$ is a sequence of i.i.d. r.v.'s having p.m.f.

$p_{\tilde\theta_n(\tilde\omega)}$ and $\{\widetilde{Y}_i^{(n)}\}$ is a sequence of i.i.d. r.v.'s. having p.m.f. $q_{\tilde\phi_n(\tilde\omega)}$, and $\widetilde{Z}_0^{(n)} = 1$. Set

$$\widehat{\widetilde{m}}_n = \left[\sum_{i=1}^{n} \widetilde{Z}_{i-1}^{(n)}\right]^{-1} \sum_{i=1}^{n} [\widetilde{Z}_i^{(n)} - \widetilde{Y}_i^{(n)}],$$
$$\widetilde{V}_n = \left[\sum_{i=1}^{n} \widetilde{Z}_{i-1}^{(n)}\right]^{1/2} \left(\widehat{\widetilde{m}}_n - \widetilde{m}_n(\widetilde{\omega})\right). \tag{4.4}$$

For each $\widetilde{\omega} \in \widetilde{\Omega}$, observe that

$$P\{\widetilde{V}_n \le x \mid \widetilde{m}_n(\widetilde{\omega}), \widetilde{\lambda}_n(\widetilde{\omega})\} = H_n\left(\widetilde{m}_n(\widetilde{\omega}), \widetilde{\lambda}_n(\widetilde{\omega}), x\right), \tag{4.5}$$

where H_n is defined in (2.1). Then, for each $\widetilde{\omega} \in \widetilde{\Omega}$, we can apply Corollary 3.1 to $\widetilde{V}_n$, provided conditions (C1)–(C4) are satisfied by the process $\{\widetilde{Z}_i^{(n)}\}$ defined in (4.3). To this end, recall from Section 2 that $m = f(\theta)$, $\sigma^2 = \theta f'(\theta)$, $\lambda = g(\phi)$ and $b^2 = \phi g'(\phi)$, where f and g are smooth functions. So, when $m = 1$ (which is the case we are interested in), $\sigma^2 = f^{-1}(1)f'(f^{-1}(1))$. By (4.2) it is clear that, for each $\widetilde{\omega}$, $\widetilde{m}_n(\widetilde{\omega})$ satisfies condition (C1) [with $\alpha = \widetilde{V}_0(\widetilde{\omega})$] and $\widetilde{\lambda}_n(\widetilde{\omega})$ satisfies the first part of (C3) [with $\lambda_0 = \lambda$] in (3.2). Also, from the above discussion and (4.2), we have for each $\widetilde{\omega}$ that

$$\widetilde{\sigma}_n^2(\widetilde{\omega}) = \operatorname{var}(\widetilde{\xi}_{1,1}^{(n)}) = f^{-1}(\widetilde{m}_n(\widetilde{\omega}))\, f'\left(f^{-1}(\widetilde{m}_n(\widetilde{\omega}))\right) \to \sigma^2 \quad \text{(when } m = 1\text{)},$$
$$\widetilde{b}_n^2(\widetilde{\omega}) = \operatorname{var}(\widetilde{Y}_1^{(n)}) = g^{-1}(\widetilde{\lambda}_n(\widetilde{\omega}))\, g'\left(g^{-1}(\widetilde{\lambda}_n(\widetilde{\omega}))\right) \to b^2 \tag{4.6}$$

as $n \to \infty$. Hence, $\widetilde{\sigma}_n^2(\widetilde{\omega})$ and $\widetilde{b}_n^2(\widetilde{\omega})$ satisfy condition (C2) (with $\sigma_0^2 = \sigma^2$) and the second part of (C3) (with $b_0^2 = b^2$), respectively. As for condition (C4), it suffices to show that, for each $\widetilde{\omega}$,

$$n^{-1/2} E\left|\widetilde{\xi}_{1,1}^{(n)} - \widetilde{m}_n(\widetilde{\omega})\right|^3 \to 0 \quad \text{as } n \to \infty. \tag{4.7}$$

Use the inequality $|x-y|^3 \le 4(|x|^3 + |y|^3)$, the exact value $E(\widetilde{\xi}_{1,1}^{(n)})^3 = \widetilde{\theta}_n^2(\widetilde{\omega})f''(\widetilde{\theta}_n(\widetilde{\omega})) + \widetilde{\theta}_n(\widetilde{\omega})f'(\widetilde{\theta}_n(\widetilde{\omega})) + 3\widetilde{\theta}_n(\widetilde{\omega})f(\widetilde{\theta}_n(\widetilde{\omega}))f'(\widetilde{\theta}_n(\widetilde{\omega})) + f^3(\widetilde{\theta}_n(\widetilde{\omega}))$, the fact that $\widetilde{\theta}_n(\widetilde{\omega}) \to f^{-1}(1)$ since $\widetilde{m}_n(\widetilde{\omega}) \to 1$ and the fact that f is smooth to conclude that (4.7) holds. Hence (C4) is also satisfied by $\{\widetilde{Z}_i^{(n)}\}$. Therefore, for each $\widetilde{\omega} \in \widetilde{\Omega}$, we have by Corollary 3.1 that

$$H_n\left(\widetilde{m}_n(\widetilde{\omega}), \widetilde{\lambda}_n(\widetilde{\omega}), x\right) \to H\left(\widetilde{V}_0(\widetilde{\omega}), \lambda, \sigma^2, x\right) \quad \text{as } n \to \infty, \tag{4.8}$$

where H is defined in (2.10). However, note that the constructions of $\{\widetilde{Z}_i^{(n)}\}$ in (4.3) and the bootstrap process $\{Z_i^*\}$ in (2.2) are similar and, furthermore, that $(\widetilde{m}_n, \widetilde{\lambda}_n)$ has the same distribution as $(\widehat{m}_n, \widehat{\lambda}_n)$ and that $(\widetilde{V}_0, \lambda)$ has the same distribution as (V_0, λ). Therefore, from (2.5), (2.10) and (4.5) we have, for each fixed x, that

$$
\begin{aligned}
(4.9)\qquad P\Big[\omega: H_n\big(\widehat{m}_n(\omega), \widehat{\lambda}_n(\omega), x\big) \le y\Big] &= Q\Big[\widetilde{\omega}: H_n\big(\widetilde{m}_n(\widetilde{\omega}), \widetilde{\lambda}_n(\widetilde{\omega}), x\big) \le y\Big],\\
P\Big[\omega: H\big(V_0(\omega), \lambda, \sigma^2, x\big) \le y\Big] &= Q\Big[\widetilde{\omega}: H\big(\widetilde{V}_0(\widetilde{\omega}), \lambda, \sigma^2, x\big) \le y\Big],
\end{aligned}
$$

for every y, where Q is defined below display (4.1). Clearly, (4.8) implies that, for each x,

$$\lim_{n\to\infty} Q\Big[\widetilde{\omega}: H_n\big(\widetilde{m}_n(\widetilde{\omega}), \widetilde{\lambda}_n(\widetilde{\omega}), x\big) \le y\Big] = Q\Big[\widetilde{\omega}: H\big(\widetilde{V}_0(\widetilde{\omega}), \lambda, \sigma^2, x\big) \le y\Big],$$

for every y. Hence, by (4.9), we have for each x that

$$\lim_{n\to\infty} P\Big[\omega: H_n\big(\widehat{m}_n(\omega), \widehat{\lambda}_n(\omega), x\big) \le y\Big] = P\Big[\omega: H\big(V_0(\omega), \lambda, \sigma^2, x\big) \le y\Big],$$

for every y. Hence the theorem. □

CONCLUDING REMARKS. Conditional on $\{(Z_i, Y_i), i = 1, 2, \dots, n\}$, one can view the bootstrap model in (2.2) as a sequence of branching processes defined in (3.1) with $\mu_n = \widehat{m}_n$, $\lambda_n = \widehat{\lambda}_n$, $\sigma_n^2 = \mathrm{var}(\xi_{1,1}^*)$ and $b_n^2 = \mathrm{var}(Y_1^*)$. Now, when $m = 1$ we have that $\widehat{m}_n \to 1$ almost surely (a.s.) as $n \to \infty$ [see, e.g., Sriram, Basawa and Huggins (1991), display (3.3)] and $\widehat{\lambda}_n = \overline{Y} \to \lambda$ a.s. Therefore, for each $\omega \in \{\widehat{m}_n \to 1 \text{ and } \widehat{\lambda}_n \to \lambda\}$, arguments similar to (4.6) and (4.7) imply that conditions (C2)–(C4) of Corollary 3.1 are satisfied by the bootstrap model with $\sigma_0^2 = \sigma^2$ (when $m = 1$), $\lambda_0 = \lambda$ and $b_0^2 = b^2$. Suppose for the moment that $\widehat{m}_n$ satisfies condition (C1) in (3.2) almost surely with $\alpha = 0$, that is, $n(\widehat{m}_n - 1) \to 0$ a.s. Then Corollary 3.1 could be directly applied to the bootstrap pivot V_n^* in (2.4), and it would imply that the parametric bootstrap is asymptotically valid, that is, $H_n(\widehat{m}_n, \widehat{\lambda}_n, x)$ converges a.s. to $H(0, \lambda, \sigma^2, x) = P[V \le x]$ for H in (2.10) and V in (2.6), since $\nu(0, \lambda, \sigma^2)$ in (3.5) has the same distribution as V. However, by (2.7) we have that $n(\widehat{m}_n - 1)$ *does not* converge to 0 a.s. Thus, the estimator $\widehat{m}_n$ not having the desired rate of convergence to 1 seems to be the main cause of failure of the parametric bootstrap.

APPENDIX

PROOF OF THEOREM 3.1. Observe that $\{Z_i^{(n)}/n, i = 0, 1, \dots\}$ is a Markov chain with values in $E_n = \{\ell/n: \ell = 0, 1, \dots\}$. Define for each $h \in C_c^\infty[0, \infty)$

$$(A.1)\qquad T_n h(x) = E\left[h\left\{n^{-1}\left[\sum_{k=1}^{nx} \xi_{1,k}^{(n)} + Y_1^{(n)}\right]\right\}\right], \qquad x \in E_n.$$

Since $\mathcal{X}_n(0) = Z_0^{(n)}/n = 1/n \to 0$ as $n \to \infty$, by Ethier and Kurtz [(1986), Theorem 6.5 of Chapter 1 and Corollary 8.9 of Chapter 4], it suffices to show that

$$(A.2)\qquad \lim_{n\to\infty} \sup_{x\in E_n} \left|n\left[T_n h(x) - h(x)\right] - \alpha x h'(x) - \lambda_0 h'(x) - \left(\sigma_0^2/2\right) x h''(x)\right| = 0,$$

for $h \in C_c^\infty[0, \infty)$. For $x \in E_n$, define

$$\varepsilon_n(x) = n\big[T_n h(x) - h(x)\big] - \alpha x h'(x) - \lambda_0 h'(x) - \big(\sigma_0^2/2\big) x h''(x) \tag{A.3}$$

and

$$S_{nx} = \left\{ \sum_{k=1}^{nx} \big(\xi_{1,k}^{(n)} - 1\big) + Y_1^{(n)} \right\} \Big/ \sqrt{nx}. \tag{A.4}$$

Note that $\sqrt{(x/n)} S_{nx} = \{\sum_{k=1}^{nx}(\xi_{1,k}^{(n)} - 1) + Y_1^{(n)}\}/n$. Then

$$\begin{aligned} T_n h(x) - h(x) &= h'(x) E\left\{ \sqrt{\frac{x}{n}} S_{nx} \right\} \\ &\quad + E\left[\frac{x}{n} S_{nx}^2 \int_0^1 (1 - v) h''\left[x + v\sqrt{\frac{x}{n}} S_{nx} \right] dv \right]. \end{aligned} \tag{A.5}$$

Since

$$E\left\{ \sqrt{\frac{x}{n}} S_{nx} \right\} = (\mu_n - 1)x + \frac{\lambda_n}{n}$$

and

$$E\{x S_{nx}^2\} = x\sigma_n^2 + \frac{b_n^2 + \lambda_n^2}{n} + n(\mu_n - 1)^2 x^2 + 2(\mu_n - 1) x \lambda_n,$$

we have from (A.3) that

$$\begin{aligned} \varepsilon_n(x) &= h'(x)\Big\{ \big[n(\mu_n - 1) - \alpha\big] x + (\lambda_n - \lambda_0) \Big\} \\ &\quad + E\left[x S_{nx}^2 \int_0^1 (1 - v) \Big[h''\big(x + v\sqrt{x/n}\, S_{nx}\big) - h''(x) \Big] dv \right] \\ &\quad + (x/2) h''(x) \big[\sigma_n^2 - \sigma_0^2\big] + \frac{1}{2} h''(x) \Big[(b_n^2 + \lambda_n^2)/n + n(\mu_n - 1)^2 x^2 + 2(\mu_n - 1) x \lambda_n \Big]. \end{aligned} \tag{A.6}$$

Suppose the support of h is contained in $[0, c]$. Now, argue as on page 389 of Ethier and Kurtz (1986) and use the fact that for any sequence $\{x_n\}$ such that $x_n \to x$, $0 < x < \infty$, conditions (C1)–(C4) in (3.2) imply

$$S_{nx_n} \to_{\mathcal{D}} N, \quad \text{as } n \to \infty, \tag{A.7}$$

where N is a $N(0, \sigma^2)$ r.v. The required result in (A.2) now follows from (A.7) and conditions (C1)–(C3). □

Acknowledgments. I would like to thank an Associate Editor and all three referees for their constructive suggestions on the original version of the paper. The first referee's comments and suggestions greatly improved the readability of the paper.

REFERENCES

ATHREYA, K. B. (1987). Bootstrap of the mean in the infinite variance case. *Ann. Statist.* **15** 724–731.

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. (1981). Maximum likelihood estimation for branching processes with immigration. *Adv. in Appl. Probab.* **13** 498–509.

BILLINGSLEY, P. (1979). *Probability and Measure.* Wiley, New York.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

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

HEYDE, C. C. and SENETA, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.* **9** 235–256.

HEYDE, C. C. and SENETA, E. (1974). Notes on "Estimation theory for growth and immigration rates in a multiplicative process." *J. Appl. Probab.* **11** 572–577.

SRIRAM, T. N., BASAWA, I. V. and HUGGINS, R. M. (1989). Sequential estimation for branching processes with immigration. Technical report, Dept. Statistics, Univ. Georgia.

SRIRAM, T. N., BASAWA, I. V. and HUGGINS, R. M. (1991). Sequential estimation for branching processes with immigration. *Ann. Statist.* **19** 2232–2243.

VENKATARAMAN, K. N. and NANTHI, K. (1982). A limit theorem on a subcritical Galton–Watson process with immigration. *Ann. Probab.* **10** 1069–1074.

WEI, C. Z. and WINNICKI, J. (1989). Some asymptotic results for the branching process with immigration. *Stochastic Process. 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.

DEPARTMENT OF STATISTICS
UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602-1952