The Annals of Probability
1982, Vol. 10, No. 4, 1069–1074

A LIMIT THEOREM ON A SUBCRITICAL GALTON-WATSON PROCESS WITH IMMIGRATION

BY K. N. VENKATARAMAN AND K. NANTHI

University of Madras, Madras

For a sub-critical Galton-Watson process **X** with immigration, estimators have been studied by Pakes (1971) for the mean of the stationary distribution of **X**, and by Nanthi (1979) for the offspring mean and the immigration mean of **X**. These estimators have been shown by them to be asymptotically normal. In this paper it is shown that they have a joint limiting distribution which is singularly normal.

1. Introduction. Let $Z(n, r)$, $Y(s)$; $r, s, n = 1, 2, \cdots$; be independent non-negative, integer-valued random variables such that (i) $Z(n, r)$ are identically distributed like Z (say), and $0 < E(Z) = m < 1$, and $0 < \operatorname{Var}(Z) = \sigma_1^2 < \infty$; and (ii) $Y(s)$ are identically distributed like Y (say), and $E(Y) = \lambda$, and $\operatorname{Var}(Y) = \sigma_2^2$ are both positive and finite.

In this paper we study the subcritical Galton-Watson process $X = (X(n); n \geq 0)$ with immigration which has the conventional specification that $X(0) = 1$, and

$$X(n) = Z(n) + Y(n); \qquad n \geq 1 \tag{1.1}$$

where

$$\begin{aligned} Z(n) &= Z(n, 1) + \cdots + Z(n, X(n-1)) \quad &\text{if} \quad X(n-1) > 0 \\ &= 0 &\text{if} \quad X(n-1) = 0. \end{aligned} \tag{1.2}$$

We recall (Heathcote, 1966; Senata 1969; Pakes, 1971) that X has a stationary probability distribution with mean $\mu = (1 - m)^{-1}\lambda$. Pakes (1971) has shown that

$$\hat{T} = n^{-1}(X(0) + \cdots + X(n)) \tag{1.3}$$

is strongly consistent for μ, and $n^{1/2}(\hat{T} - \mu)$ converges in law ($\rightarrow_{\mathscr{L}}$), as $n \rightarrow \infty$, to a normal random variable with mean zero, and variance $A_0^2\sigma_0^2$, where $A_0 = (1 - m)^{-1}$, and $\sigma_0^2 = \mu\sigma_1^2 + \sigma_2^2$, using a classical triangular array approach. An alternative proof of this result based on a time series approach, has been recently provided by Venkataraman (1982).

On the assumption that $Z(r)$ and $Y(r)$, $r = 1, \cdots, n$ are observed, Nanthi (1979) has identified

$$\hat{m} = (X(0) + \cdots + X(n-1))^{-1} \textstyle\sum_{r=1}^{n} Z(r)$$

and

$$\hat{\lambda} = n^{-1}(Y(1) + \cdots + Y(n)) \tag{1.4}$$

as maximum likelihood estimators of m and λ, and has shown that $n^{1/2}(\hat{m} - m)$ converges in law, as $n \rightarrow \infty$, to a normal random variable with mean zero and variance $\mu^{-1}\cdot\sigma_1^2$. It is elementary to note that $n^{1/2}(\hat{\lambda} - \lambda)$ converges in law, as $n \rightarrow \infty$, to a normal random variable with mean zero and variance σ_2^2.

The principal aim of this paper is to derive the joint limiting distribution of $\hat{T}$, $\hat{m}$, and $\hat{\lambda}$. To be precise the following basic result is proved.

Received May 1980; revised December 1981.

AMS 1980 *subject classification.* Primary 62M99; secondary 60F99.

Key words and phrases. Subcritical Galton-Watson process with immigration, limit distribution.

THEOREM 1. *Under the basic assumptions on X,*

$$(n^{1/2}(\hat{T} - \mu), \quad n^{1/2}(\hat{m} - m), \quad n^{1/2}(\hat{\lambda} - \lambda))$$

converges in law, as $n \to \infty$, *to a normal random vector* (ξ_1, ξ_2, ξ_3) *say which has zero mean, and the properties that* (a) $\xi_1 = A_0(\mu\, \xi_2 + \xi_3)$; (b) ξ_2 *and* ξ_3 *are independent; and* (c) $\mathrm{Var}(\xi_2) = \mu^{-1}\sigma_1^2$, *and* $\mathrm{Var}(\xi_3) = \sigma_2^2$.

The singularity of the joint limiting distribution of $\hat{T}$, $\hat{m}$, and $\hat{\lambda}$, as well as the asymptotic independence of $\hat{m}$ and $\hat{\lambda}$ are perhaps anticipated, but are not readily established. The basic lemmas that are needed to prove Theorem 1 are presented in Section 2. The final argument in the proof of Theorem 1 is developed in Section 3. Some concluding remarks are made in Section 4. We adopt a direct characteristic function-approach to prove Theorem 1, relating it to the probability generating functions $f(u)$ and $h(u)$ of Z and Y respectively, which are well defined on the compact set $D_0 = (u;\, |u| \le 1)$ of the complex plane. To list the basic features of $f(u)$ and $h(u)$ that are exploited in the sequel, we observe that these functions have derivates up to second order on D_0 by virtue of the moment conditions satisfied by Z and Y. Next it is obvious that

$$\begin{aligned} f(u) - f(v) &= \textstyle\sum_{r=0}^{\infty} P(Z = r)(u^r - v^r) \\ f'(u) - f'(v) &= \textstyle\sum_{r=0}^{\infty} r\, P(Z = r)(u^{r-1} - v^{r-1}). \end{aligned} \tag{1.5}$$

Using the identity that $u^r - v^r = (u - v)(u^{r-1} + u^{r-2}\, v + \cdots + v^{r-1})$ we infer from (1.5) that, for $u, v \in D_0$

$$\begin{aligned} |f(u) - f(v)| &\le m\, |u - v| \\ |f'(u) - f'(v)| &\le f''(1)\, |u - v|. \end{aligned} \tag{1.6}$$

A similar proof shows that, for $u, v \in D_0$

$$\begin{aligned} |h(u) - h(v)| &\le \lambda\, |u - v| \\ |h'(u) - h'(v)| &\le h''(1)\, |u - v|. \end{aligned} \tag{1.7}$$

The moment conditions imposed on Z and Y do not guarantee such relations for $f''(u)$ and $h''(u)$. However these functions are uniformly continuous on D_0.

2. Preliminary ground work. Following Pakes (1975) we introduce, for $u \in D_0$

$$g_0(u) = u, \quad \text{and} \quad g_n(u) = u\, f(g_{n-1}(u)); \qquad n \ge 1. \tag{2.1}$$

It is easy to check that, as observed by Pakes (1975), $g_n(u)$ is the probability generating function of the total progeny $\eta(n)$, say of a simple Galton-Watson process, originating from a single ancestor, and being governed by $f(u)$ as its offspring probability generating function.

LEMMA 1. *Under the moment conditions on Y and Z, the following statements hold for* $u, v \in D_0$, *and, for* $n \ge 0$.

(a) $|g_n'(u)| \le A_0$; $\quad |g_n''(u)| \le A_0 B_0$; $\quad B_0 = f''(1)A_0^2 + 2mA_0$

(b) $|g_n(u) - g_n(v)| \le A_0\, |u - v|$; $\quad |g_n'(u) - g_n(v)| \le A_0 B_0\, |u - v|$

(c) *Given* $\varepsilon > 0$, *there exists for all n, a* $\delta_0(\varepsilon) > 0$, *depending only on* ε, *such that*

$$|g_n(u) - g_n(v)|, \qquad |g_n'(u) - g_n'(v)|, \qquad g_n''(u) - g_n''(v)| < \varepsilon$$

whenever $|u - v| < \delta_0$.

(d)
$$\begin{aligned} g_n'(1) &= A_0(1 - m^{n+1}) \\ g_n''(1) &= A_0 B_0 + C_1(n-1)m^n + C_2(n-1)m^{2n} + C_3 m^n \end{aligned}$$

where C_1, C_2 and C_3 are constants of appropriate determination.

PROOF. By definition $g'_n(1) = E\eta(n) = A_0(1 - m^{n+1})$. An easy way to work out $g''_n(1)$ is to note that

$$g''_n(u) = uf'(g_{n-1}(u))g''_{n-1}(u) + uf''(g_{n-1}(u))(g'_{n-1}(u)) + 2f'(g_{n-1}(u))g'_{n-1}(u). \tag{2.2}$$

Letting $u = 1$ in (2.2), and solving the resulting difference equation, on substituting for $g'_n(1)$, the expression for $g''_n(1)$ in (d) is obtained. Another implication of (2.2) is that

$$g''_n(1) \leq m\, g''_{n-1}(1) + B_0 \tag{2.3}$$

which recursively implies that $g''_n(1) \leq A_0B_0$. The proof of (a) is completed on noting that $g'_n(1) \leq A_0$, and $|g'_n(u)| \leq g'_n(1)$, and $|g''_n(u)| \leq g''_n(1)$. The proof of (b) needs only the additional remarks that, as in the case of (1.7),

$$|g_n(u) - g_n(v)| \leq g'_n(1)\,|u - v|, \qquad |g'_n(u) - g'_n(v)| \leq g''_n(1)\,|u - v|. \tag{2.4}$$

To prove (c), we choose $|u - v| < \min(A_0^{-1}\varepsilon, (A_0B_0)^{-1}\varepsilon) = \delta_1(\varepsilon)$(say) to render, by virtue of (b), each of $|g_n(u) - g_n(v)|$, and $|g'_n(u) - |g'_n(v)|$ less than ε for $|u - v| < \delta_1$. We make use of (1.6), (2.2), (a) and (b), and the uniform continuity of $f''(u)$ on D_0 to derive, on some manipulation, that there exists a $\delta_2(\varepsilon) > 0$ such that, for $|u - v| < \delta_2$,

$$|g''_n(u) - g''_n(v)| \leq A_0^{-1}\varepsilon + m\,|g''_{n-1}(u) - g''_{n-1}(v)| \tag{2.5}$$

and thus $|g''_n(u) - g''_n(v)| < \varepsilon$. The proof of (c) is completed on choosing $\delta_0 = \min(\delta_1, \delta_2)$.

Next, for fixed real a, b, and θ, we let

$$\phi_r(\theta) = e^{i\theta a}g_r(e^{i\theta b}); \qquad r \geq 0 \tag{2.6}$$

where i is the complex root of -1. It is obvious that $\phi_r(\theta)$ is the characteristic function of $a + b\,\eta(n)$. Further, on direct evaluation, for $r \geq 0$.

$$\phi'_r(\theta) = iae^{i\theta a}g_r(e^{i\theta b}) + ibe^{i\theta(a+b)}g'_r(e^{i\theta b}) \tag{2.7}$$

$$\phi''_r(\theta) = -[a^2e^{i\theta a}g_r(e^{i\theta b}) + (2ab + b^2)e^{i\theta(a+b)}g'_r(e^{i\theta b}) + b^2e^{i\theta(a+2b)}g''_r(e^{i\theta b})]. \tag{2.8}$$

These remarks on $\phi_r(\theta)$, together with Lemma 1, and the fact that for real α, and θ, $|\exp(i\,\alpha\,\theta) - 1| \leq 2\,|\alpha|\,\theta$ yield the following lemma stated without proof.

LEMMA 2. *Under the moment conditions on Y and Z, the following statements hold for $r \geq 0$, and $|a| + |b| > 0$.*

(a) $$|\phi_r(\theta)| < 1; \qquad |\phi'_r(\theta)| \leq |a| + |b|A_0 = M_1 \quad (say), \text{ and}$$

$$|\phi''_r(\theta)| \leq (a^2 + 2\,A_0\,|a||b| + A_0(1 + B_0)b^2) - M_2 \quad (say)$$

(b) $$\phi_r(0) = 1; \qquad \phi'_r(0) = ik_1 + L_1(r); \qquad \phi''_r(0) = -K_2 + L_2(r)$$

where (i) $K_1 = a + b\,A_0$; (ii) $K_2 = (a^2 + 2ab\,A_0 + A_0(1 + B_0)b^2)$,

$$\textstyle\sum_{r=0}^{\infty} L_t(r) \leq M_3 \quad (constant) \text{ for } \quad t = 1, 2$$

(c) *Given $\varepsilon > 0$, there exists, for all r, a $\tau_0(\varepsilon) > 0$, depending only on ε, such that,*

$$|\phi_r(\theta) - \phi_r(0)|, \qquad \phi'_r(\theta) - \phi'_r(0)|, \qquad |\phi''_r((\theta) - \phi''_r(0)| < \varepsilon$$

whenever $|\theta| < \tau_0$.

Let us define that, for $r \geq 0$

$$H_r(\theta) = h(\phi_r(\theta)). \tag{2.9}$$

We list below the basic properties of $H_r(\theta)$ of interest.

LEMMA 3. *Under the moment conditions on Y and Z, the following statements hold for $r \geq 0$, and $|a| + |b| > 0$.*

(a) $$H_r(\theta) = H_r(0) + \theta H_r'(0) + \tfrac{1}{2}\theta^2 H_r''(0) + \theta^2 J_r(\theta)$$

where

(i) $$H_r(0) = 1, \text{ and } H_r'(0) = \lambda\, \phi_r'(0)$$

(ii) $$H_r''(0) = h''(1)\phi_r'^2(0) + \lambda\phi_r''(0)$$

(b) *Given $\varepsilon > 0$, there exists for all r, a $\tau_1(\varepsilon) > 0$, depending only on ε, such that $|J_r(\theta)| < \varepsilon$, whenever $|\theta| < \tau_1$.*

PROOF. Maclaurin expansions of the real and imaginary parts of $H_r(\theta) = H_{1r}(\theta) + iH_{2r}(\theta)$ yield (a) on identifying

$$2\, J_r(\theta) = (H_{1r}''(C_1\theta) - H_{1r}''(0)) + i(H_{2r}''(C_2\,\theta) - H_{2r}''(0)) \tag{2.10}$$

where $C_1(r, \theta)$, and $C_2(r, \theta)$ are some real numbers in the open interval (0, 1). A direct evaluation of $H_r''(\theta)$ implies that

$$\begin{aligned} |H_r''(\theta) - H_r''(0)| \leq 2M_1 h''(1)\,|\phi_r'(\theta) - \theta_r'(0)| + \lambda\,|\phi_r''(\theta) - \phi_r''(0)| \\ + M_2\,|h'(\phi_r(\theta)) - h'(\phi_r(0))| \\ + M_1^2\,|h''(\phi_r(\theta)) - h''(\phi_r(0))| \end{aligned} \tag{2.11}$$

M_1 and M_2 being specif ed in Lemma 2. A joint appeal to Lemma 2(c), (2.4), (2.11) and the uniform continuity of $h''(u)$ on D_0 yields, on some manipulation that given $\varepsilon > 0$, there exists a $\tau_1(\varepsilon) > 0$ such that $|H_r''(\theta) - H_r''(0)| < \varepsilon$ whenever $|\theta| < \tau$, and hence, by virtue of (2.10) $|J_r(\theta)| < \varepsilon$ whenever $|\theta| < \tau_1$.

Next we observe that $H_r(\theta)$ being a characteristic function, and $|H_r'(\theta)| \leq M_4$(a constant), it is possible to choose $\rho_0 > 0$ such that $|H_r(\theta) - 1| < \tfrac{1}{2}$ if $|\theta| < \rho_0$. As a consequence, an appropriate expansion of $\log(1 + z)$ yields that, for $|\theta| < \rho_0$,

$$\log H_r(\theta) = \psi_r(\theta) - \tfrac{1}{2}\psi_r^2(\theta) + d(r, \theta)\psi^3(r, \theta) \tag{2.12}$$

where $\psi_r(\theta) = (H_r(\theta) - 1)$, and $|d(r, \theta)| < 1$ for $|\theta| < \rho_0$. We invoke Lemma 2(b), and Lemma 3 to derive from (2.12) the following lemma stated without proof.

LEMMA 4. *Under the moment conditions on X,*

$$-n^{1/2}i\lambda K_1 + \textstyle\sum_{r=0}^{n-1} \log H_r(n^{-1/2}) \to -\tfrac{1}{2}(h''(1)K_1^2 + \lambda K_2 - \lambda^2 K_1^2)$$

as $n \to \infty$, for fixed real a, and, b with $|a| + |b| > 0$.

3. Final argument in the proof of Theorem 1. We invoke the Markov property of X to derive from (1.1) that $E\,X_n = \lambda + (1 - \mu)m^n$, and thus

$$n^{-1/2}X_n \to_p 0 \quad \text{as} \quad n \to \infty. \tag{3.1}$$

Let us define that

$$\hat{T}_1 = n^{1/2}(\hat{T} - \mu);\ \hat{T}_2 = n^{-1/2}\textstyle\sum_{r=1}^{n}(Z(r) - mX(r-1));\ \hat{T}_3 = n^{1/2}(\hat{\lambda} - \lambda). \tag{3.2}$$

An appeal to (3.1) yields that, for real α_1, α_2, and α_3

$$\begin{aligned} \alpha_1\hat{T}_1 + \alpha_2\hat{T}_2 + \alpha_3\hat{T}_3 = -n^{1/2}(\alpha_1\mu + \alpha_3\lambda) + n^{-1/2}(\alpha_1 - m\alpha_2 + \alpha_3)(X(1) + \cdots + X(n)) \\ + n^{-1/2}(\alpha_2 - \alpha_3)(Z(1) + \cdots + Z(n)) + O_p(1) \end{aligned}$$

A recursive use of (1.1) together with the Markov property of X implies that the characteristic function $I_n(\alpha, \beta)$ of

(3.4) $$(n^{-1/2}(X(1) + \cdots + X(n)), \quad n^{-1/2}(Z(1) + \cdots + Z(n)))$$

is given by

(3.5) $$f[g_n(\exp(in^{-1/2}(\alpha + \beta)))]\prod_{r=0}^{n-1} H_r(n^{-1/2})$$

on identifying that $a = -\beta$, and $b = \alpha + \beta$. The first factor on (3.5) can be shown to converge to unity, as $n \to \infty$, on invoking (1.6) and Lemma 1(b). Thus an application of Lemma 4 together with (3.5) yields that, as $n \to \infty$

(3.6) $$\log I_n - in^{-1/2}\lambda K_1 \to -\tfrac{1}{2}(h''(1)K_1^2 + \lambda K_2 - \lambda^2 K_1^2).$$

These observations as applied to (3.3) enable us to infer that, as $n \to \infty$, the characteristic function of $(\hat{T}_1, \hat{T}_2, \hat{T}_3)$ converges to the normal characteristic function $\exp(-\tfrac{1}{2}Q(\alpha_1, \alpha_2, \alpha_3))$, where

(3.7) $$Q(\alpha_1, \alpha_2, \alpha_3) = A_0^2\sigma_0^2\alpha_1^2 + \mu\sigma_1^2\alpha_2^2 + \sigma_2^2\alpha_3^2 + 2\mu A_0\sigma_1^2\alpha_1\alpha_2 + 2A_0\sigma_2^2\alpha_1\alpha_3.$$

It can be easily checked that $Q(\alpha_1, \alpha_2, \alpha_3) = 0$ when $-A_0\alpha_1 = \alpha_2 = \alpha_3$. Theorem 1 follows from these results and the additional observations:

(i) $$n^{-1/2}(\hat{m} - m) = n(X(0) + \cdots + X(n-1))^{-1}\hat{T}_2$$

(3.8) (ii) $$n^{-1}(X(0) + \cdots + X(n-1)) \to_p \mu \quad \text{as} \quad n \to \infty.$$

4. Concluding remarks. Based on $\hat{m}$ and $\hat{\lambda}$, the maximum likelihood estimator of μ is identified to be $\hat{\mu} = (1 - \hat{m})^{-1}\hat{\lambda}$, which satisfies the relation that

(4.1) $$(\hat{\mu} - \mu) = (1 - \hat{m})^{-1}(\hat{\lambda} - \lambda) + \lambda(1 - \hat{m})^{-1}(1 - m)^{-1}(\hat{m} - m).$$

This relation together with Theorem 1 yields:

THEOREM 2. *Under the basic assumptions on X, $n^{1/2}(\hat{\mu} - \mu)$ converges in law, as n $\to \infty$, to the normal random variable ξ_1 (of Theorem 1).*

Thus $\hat{T}$ and $\hat{\mu}$ turn out to be equally asymptotically efficient estimators of μ, although the determination of $\hat{\mu}$ calls for the values of the components $Z(r)$ and $Y(r)$, $r = 1, \cdots, n$.

Let $(\tilde{m}, \tilde{\lambda})$ and (m^*, λ^*) be the estimators proposed for (m, λ) by Quine (1976) and Klimko and Nelson (1978), based on $X(t)$, $t = 1, \cdots, n$. Venkataraman (1982) has shown that each of

$$(n^{1/2}(\hat{T} - \mu), \quad n^{1/2}(\tilde{m} - m), n^{1/2}(\hat{\lambda} - \lambda)), \quad \text{and}$$

$$(n^{1/2}(\hat{T} - \mu), \quad n^{1/2}(m^* - m), \quad n^{1/2}(\lambda^* - \lambda))$$

converges in law, as $n \to \infty$, to the same normal vector $(\xi(1), \xi(2), \xi(3))$ of mean zero, exhibiting a singularity determined by the relation $\xi(1) = A_0(\mu\xi(2) + \xi(3))$, in spite of $\xi(2)$ and $\xi(3)$ being correlated. This pattern of singularity looks interesting.

Acknowledgments. The authors wish to place on record their sincere thanks to Prof. Pakes, and the referee for their helpful remarks on an earlier version of this paper.

REFERENCES

[1] HEATHCOTE, C. R. (1966). Corrections and comments on the paper "A branching process allowing immigration." *J. Roy. Statist. Soc. B* **28** 213–217.

[2] KLIMKO, L. A. and NELSON, P. I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* **6** 629–642.

[3] NANTHI, K. (1979). Some limit theorems of statistical relevance on branching processes. Ph.D.

thesis submitted to the University of Madras.
[4] PAKES, A. G. (1971). Branching process with immigration. *J. Appl. Probability* **8** 32–42.
[5] PAKES, A. G. (1975). Some results for non-super-critical Galton-Watson processes with immigration. *Math. Bio-Sciences* **24** 71–92.
[6] QUINE, M. P. (1976). Asymptotic results for estimators in a subcritical branching process with immigration. *Ann. Probability* **4** 319–325.
[7] SENETA, E. (1969). Functional equations and the Galton-Watson Process. *Adv. Appl. Probability.* **1** 1–42.
[8] VENKATARAMAN, K. N. (1982). A time series approach to the study of the simple subcritical Galton-Watson process with immigration. *Adv. Appl. Probability* **14** 1–20.

DEPARTMENT OF STATISTICS
UNIVERSITY OF MADRAS
CHAPAUK, MADRAS 600005
INDIA