The Annals of Statistics
1991, Vol. 19, No. 4, 2232–2243

SEQUENTIAL ESTIMATION FOR BRANCHING PROCESSES WITH IMMIGRATION

By T. N. Sriram[1], I. V. Basawa[2] and R. M. Huggins

University of Georgia, University of Georgia and La Trobe University

For the critical and subcritical Galton–Watson processes with immigration, it is shown that if the data were collected according to an appropriate stopping rule, the natural sequential estimator of the offspring mean m is asymptotically normally distributed for each fixed $m \in (0, 1]$. Furthermore, the sequential estimator is shown to be asymptotically normally distributed uniformly over a class of offspring distributions with $m \in (0, 1]$ bounded variance and satisfying a mild condition. These results are to be contrasted with the nonsequential approach where drastically different limit distributions are obtained for the two cases: (a) $m < 1$ (normal) and (b) $m = 1$ (nonnormal), thus leading to a singularity problem at $m = 1$. The sequential approach proposed here avoids this singularity and unifies the two cases. The proof of the uniformity result is based on a uniform version of the well-known Anscombe's theorem.

1. Introduction. Branching processes provide useful models in cell kinetics, population growth and other related areas. The estimation problem for Galton–Watson processes with immigration has been discussed extensively in the literature. The pioneering work in this area is due to Heyde and Seneta (1972, 1974). If m denotes the mean of the offspring distribution of the process and $\hat{m}_n$ its estimate, defined for instance in (2.2), it is 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 $\hat{m}_n$ is drastically different for the three cases $m < 1$ (subcritical), $m = 1$ (critical) and $m > 1$ (supercritical). The supercritical processes belong to the so-called regular nonergodic family studied by Basawa and Scott (1976, 1983) among others. The supercritical case will not be considered here. This paper is concerned with the two cases $m < 1$ and $m = 1$. It will be useful to develop a unified approach for estimation which does not require the prior information as to whether $m < 1$ or $m = 1$. We propose to use sequential approach to achieve this.

It will be shown that $\hat{m}_n$ has a limiting normal distribution for $m < 1$ and a nonnormal limit distribution for $m = 1$. If one therefore wishes to obtain a

Received September 1989; revised January 1991.
[1]Work partially supported by the University of Georgia Faculty Research Grant.
[2]Work partially supported by a grant from the Office of Naval Research.
AMS 1980 *subject classifications*. Primary 60J80; secondary 62L10.
Key words and phrases. Fixed-width confidence intervals, sequential estimation, branching processes with immigration, stopping time, uniform CLT, uniform Anscombe's theorem, uniform asymptotic normality.

confidence interval for $m, m \in (0, 1]$, one faces the problem of the singularity at the endpoint $m = 1$. In a similar situation dealing with the first-order nonexplosive autoregressive processes, Lai and Siegmund (1983) discussed a sequential approach which enables one to establish the asymptotic normality of the least squares estimator of the autoregressive parameter for stationary and unstable cases, thus avoiding the singularity at the endpoint in their problem. This suggests the question as to whether a similar sequential approach will resolve the singularity problem in the estimation of $m, m \in (0, 1]$ for the branching processes. We show in this paper that the answer is in the affirmative.

Our estimate $\hat{m}_n$ in (2.2) is based on the full information on both generation sizes $\{Z_j\}$ and the immigration process $\{Y_j\}$, $j = 1, 2, \dots, n$. Furthermore, $\hat{m}_n$ is the maximum likelihood estimate for a large class of offspring and immigration distributions, and consequently it is fully efficient; see, for instance, Hall and Heyde (1980). In the literature, the estimation of m by $\hat{m}_n$ based on the full information has been considered by Nanthi (1983) and Venkataraman and Nanthi (1982). Using only the partial information on $\{Z_j\}$ 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. The efficiency properties of the latter estimate are not known at this stage. Our main goal in this paper is to use a simple and fully efficient estimate $\hat{m}_n$ to develop a sequential approach which unifies the two cases $m < 1$ and $m = 1$. The stopping rule we use is related to the observed Fisher information as in Lai and Siegmund (1983). Extension of the sequential approach to the estimate of m based only on the $\{Z_j\}$ is a possibility, but it will not be considered in this paper.

Motivation for the sequential estimate of m is discussed in Section 2. Some basic, nonsequential limit results are derived in Section 3. Section 4 contains asymptotic normality of the sequential estimator for each fixed $m \leq 1$ and properties of the stopping time used here. Section 5 contains uniform asymptotic normality of the sequential estimate of m established uniformly over a class of offspring distributions. The results of Section 5 were obtained independently by the first author.

2. Motivation. Let Z_n denote the nth generation size of a Galton–Watson process with immigration. We then have the representation

$$Z_n = \sum_{k=1}^{Z_{n-1}} \xi_{n-1,k} + Y_n, \qquad n = 1, 2, \dots, \tag{2.1}$$

where $\xi_{n-1,k}$ is the number of offspring of the kth individual belonging to the $(n-1)$th generation and Y_n denotes the number of immigrants in the nth generation. Suppose that $\{\xi_{n-1,k}\}$, $n = 1, 2, \dots$, $k = 1, 2, \dots$, and $\{Y_n\}$, $n = 1, 2, \dots$ are two independent sequences of independent and identically distributed (i.i.d.) random variables. The initial state Z_0 is a random variable (not depending on m) which is independent of $\{\xi_{n,j}\}$ and $\{Y_n\}$ and has an arbitrary distribution. The offspring and the immigration distributions are assumed to be unspecified with means m and λ and variances $\sigma^2 \in (0,\infty)$ and $\sigma_Y^2 \in (0,\infty)$, respectively. It is assumed that a sample $\{Z_0, (Z_i, Y_i), i = 1,2,\dots,n\}$ is available. Let $\mathscr{F}_n$ be the σ-field generated by $\{Z_0, \xi_{i-1,j}, Y_i, 1 \le i \le n, j \ge 1\}$. Our primary goal is to make inferences regarding the offspring mean m. Our aim is to study the sequential approach and give a unified method of inference for the critical and subcritical cases $m \le 1$.

It is clear from (2.1) that a natural estimate of m is

$$\hat{m}_n = \left(\sum_{i=1}^{n} Z_{i-1}\right)^{-1} \sum_{i=1}^{n} (Z_i - Y_i). \tag{2.2}$$

Suppose we assume the power-series offspring and immigration distributions; we can then show that $\hat{m}_n$ in (2.2) and $\hat{\lambda}_n = n^{-1}\sum_{i=1}^{n} Y_i$ are maximum likelihood estimators of m and λ, respectively. Thus, for a large class of distributions, $\hat{m}_n$ in (2.2) is the maximum likelihood estimate of m. From now on, we shall not make any specific distributional assumptions regarding ξ and Y, but consider $\hat{m}_n$ in (2.2) as a reasonable estimate with which to work.

Consider the stopping rule N_c defined below by

$$N_c = \inf\left\{n \ge 1: \sum_{i=1}^{n} Z_{i-1} \ge c\sigma^2\right\}, \tag{2.3}$$

where $c > 0$ is chosen appropriately. One can motivate using (2.3) via the theory of fixed-width confidence intervals [see Chow and Robbins (1965)]. Note that (2.3) assumes that σ^2 is known. If σ^2 is unknown, we can replace it by a strongly consistent estimate $\hat{\sigma}_n^2$ of σ^2 for $m \le 1$ (to be shown later) defined by

$$\hat{\sigma}_n^2 = \left\{\sum_{i=1}^{n} \frac{Z_{i-1}}{(1 + Z_{i-1})^2}\right\}^{-1} \sum_{i=1}^{n} \frac{(Z_i - \hat{m}_n Z_{i-1} - Y_i)^2}{(1 + Z_{i-1})^2}. \tag{2.4}$$

We shall denote the resulting stopping time by $\tilde{N}_c$. We shall show that the limit distribution of $\hat{m}_{N_c}$ and $\hat{m}_{\tilde{N}_c}$ as $c \to \infty$ is normal for both the cases $m < 1$ and $m = 1$. Furthermore, we shall show that $\hat{m}_{N_c}$ is asymptotically normally distributed uniformly over a class of offspring distributions with $m \le 1$ and satisfying other conditions (see below).

3. Basic limit results. In this section we give some asymptotic results concerning the estimators $\hat{m}_n$ and $\hat{\sigma}_n^2$ defined in Section 2. The first theorem of this section (Theorem 3.1) concerns the strong consistency of $\hat{\sigma}_n^2$ for $0 < m \le 1$, which provides a foundation for consideration of the unknown σ case in Section 4. Theorem 3.2 gives the limit distribution of $\hat{m}_n$ for $0 < m < 1$ and $m = 1$, respectively.

Before we state the first theorem, let

$$(3.1)\qquad \begin{aligned} V_n &= \sum_{i=1}^{n} \frac{Z_{i-1}}{(1+Z_{i-1})^2}, \qquad T_n = \sum_{i=1}^{n} \frac{1}{Z_{i-1}+1}, \\ \varepsilon_i &= Z_i - mZ_{i-1} - Y_i \quad \text{and} \quad I_n = \sum_{i=1}^{n} Z_{i-1}. \end{aligned}$$

THEOREM 3.1. *Assume the model* (2.1). *If* $E\xi_{1,1}^4 < \infty$, *then for each* $m \leq 1$,

$$\hat{\sigma}_n^2 \to \sigma^2 \quad a.s. \text{ as } n \to \infty,$$

where $\hat{\sigma}_n^2$ *is as defined in* (2.4).

PROOF. By algebraic manipulations note that

$$(3.2)\qquad \begin{aligned} \hat{\sigma}_n^2 - \sigma^2 = V_n^{-1}\Bigg\{ &\sum_{i=1}^{n} \frac{\varepsilon_i^2 - \sigma^2 Z_{i-1}}{(1+Z_{i-1})^2} + (\hat{m}_n - m)^2 \sum_{i=1}^{n} \frac{Z_{i-1}^2}{(1+Z_{i-1})^2} \\ &- 2(\hat{m}_n - m) \sum_{i=1}^{n} \frac{\varepsilon_i Z_{i-1}}{(1+Z_{i-1})^2} \Bigg\}. \end{aligned}$$

Note that $\{\sum_{i=1}^{n}(\varepsilon_i^2 - \sigma^2 Z_{i-1})/(1+Z_{i-1})^2, \mathscr{F}_n\}$, $\{\sum_{i=1}^{n}\varepsilon_i Z_{i-1}/(1+Z_{i-1})^2, \mathscr{F}_n\}$ and $\{\sum_{i=1}^{n}\varepsilon_i, \mathscr{F}_n\}$ are martingales. Since $V_n \to \infty$ for each $m \leq 1$, by the strong law of large numbers for martingales [e.g., Hall and Heyde (1980), Theorem 2.18], for each $m \leq 1$,

$$V_n^{-1} \sum_{i=1}^{n} \frac{\varepsilon_i^2 - \sigma^2 Z_{i-1}}{(1+Z_{i-1})^2} \to 0 \quad \text{a.s.},$$

$$V_n^{-1} \sum_{i=1}^{n} \frac{\varepsilon_i Z_{i-1}}{(1+Z_{i-1})^2} \to 0 \quad \text{a.s.}$$

and

$$(3.3)\qquad (\hat{m}_n - m) = \frac{\sum_{i=1}^{n}\varepsilon_i}{I_n} \to 0 \quad \text{a.s. as } n \to \infty.$$

In view of (3.3), it suffices to show for each $m \leq 1$ that

$$(3.4)\qquad V_n^{-1}(\hat{m}_n - m)^2 \sum_{i=1}^{n} \frac{Z_{i-1}^2}{(1+Z_{i-1})^2} \to 0 \quad \text{a.s. as } n \to \infty.$$

For $m < 1$, (3.4) follows easily from (3.3) and the ergodicity of $\{Z_i\}$. As for

$m = 1$, for some suitable $\varepsilon \in (0, \frac{1}{2})$ write

$$V_n^{-1}(\hat{m}_n - m)^2 \sum_{i=1}^{n} \frac{Z_{i-1}^2}{(1 + Z_{i-1})^2} = \left(1 + \frac{n}{I_n}\right)^2 \left[\frac{(\sum_{i=1}^{n} \varepsilon_i)^2}{(I_n + n)^{1+\varepsilon}}\right] \times \left[\sum_{i=1}^{n} \frac{Z_{i-1}^2}{(1 + Z_{i-1})^2}\right] V_n^{-1}(I_n + n)^{-(1-\varepsilon)}. \tag{3.5}$$

Since $Z_i \geq Y_i$, $i \geq 1$, $(1 + n/I_n)^2 \leq (1 + n/\sum_{i=1}^{n-1} Y_i)^2 \to (1 + \lambda^{-1})^2$ a.s. as $n \to \infty$ by strong law of large numbers (SLLN). Moreover,

$$\left[\sum_{i=1}^{n} \frac{Z_{i-1}^2}{(1 + Z_{i-1})^2}\right] V_n^{-1}(I_n + n)^{-(1-\varepsilon)} \leq n(I_n + n)^{-(1-\varepsilon)} V_n^{-1} \to 0 \quad \text{a.s. as } n \to \infty,$$

since $V_n \to \infty$ a.s. and by Cauchy–Schwarz inequality $n^2(I_n + n)^{-1} \leq T_n$ which implies $n^{1/(1-\varepsilon)}[I_n + n]^{-1} \leq n^{-[2-1/(1-\varepsilon)]} T_n \to 0$ a.s. as $n \to \infty$ by Theorems 2.16, 2.22 and Corollary 2.21 of Wei and Winnicki (1989). Also, by SLLN for martingales $(\sum_1^n \varepsilon_i)^2/(I_n + n)^{1+\varepsilon} \to 0$ a.s. Hence, (3.4) follows from (3.5) and the above arguments. Hence the theorem. □

The next theorem gives the limit distribution of $\hat{m}_n$ for the cases $m < 1$ and $m = 1$. The proof is omitted; see Sriram, Basawa and Huggins (1989) for details.

THEOREM 3.2. *Assume the model* (2.1). *For the estimator $\hat{m}_n$ defined in* (2.2),

$$\left(\sum_{i=1}^{n} Z_{i-1}\right)^{1/2} (\hat{m}_n - m) \to_D \begin{cases} N(0, \sigma^2), & \text{if } m < 1, \\ \{Y(1) - \lambda\} \Big/ \left\{\int_0^1 Y(t)\, dt\right\}^{1/2}, & \text{if } m = 1, \end{cases}$$

as $n \to \infty$, where $Y(t)$ is a nonnegative diffusion process with a generator, which is obtained as a weak limit of the process $Y_n(t) = Z_{[nt]}/n$ as $n \to \infty$.

REMARK 3.1. The result in Theorem 3.2 shows that drastically different limit results obtain for the cases $m < 1$ and $m = 1$. This phenomenon is analogous to a similar result for autoregressive processes corresponding to the stationary and unstable cases discussed by Lai and Siegmund (1983).

In the next section we show that if we replace n in Theorem 3.2 by N_c given by (2.3) for the case of known σ, then the limit distribution as $c \to \infty$ turns out to be $N(0, \sigma^2)$ for both $m < 1$ and $m = 1$, thus avoiding the singularity at

$m = 1$. It is also shown that a similar result holds if we replace n by $\tilde{N}_c$ (defined below) for the case of unknown σ.

4. Asymptotic normality of $\hat{m}_{N_c}$ and $\hat{m}_{\tilde{N}_c}$. In this section we will be concerned with the limit distribution of $\hat{m}_{N_c}$ and $\hat{m}_{\tilde{N}_c}$ for $m \le 1$. Now, since $Z_i \ge Y_i$ for $i \ge 1$, we have $\Sigma_1^n Z_{i-1} \ge \Sigma_1^{n-1} Y_i \to \infty$ a.s. as $n \to \infty$. From this we have that $P_m\{N_c < \infty\} = 1$ for all $m \le 1$. Obviously, $P_m\{\lim_{c\to\infty} N_c = \infty\} = 1$ for all $m \le 1$. Similar results hold for $\tilde{N}_c$ as well, since $\hat{\sigma}_n^2 \to \sigma^2$ a.s. as $n \to \infty$. The following theorems give the limit distribution of $\hat{m}_{N_c}$ and $\hat{m}_{\tilde{N}_c}$.

THEOREM 4.1. *For each fixed $m \le 1$ and N_c defined in* (2.3), *we have*

$$\left(\sum_{i=1}^{N_c} Z_{i-1}\right)^{1/2} \left(\hat{m}_{N_c} - m\right) \to_D N(0, \sigma^2) \quad \text{as } c \to \infty. \tag{4.1}$$

THEOREM 4.2. *Assume that $E\xi_{1,1}^4 < \infty$. Suppose the stopping time $\tilde{N}_c$ is defined by $\tilde{N}_c = \inf\{n \ge 2: \Sigma_{i=1}^n Z_{i-1} \ge c\hat{\sigma}_n^2\}$, where $\hat{\sigma}_n^2$ is as in* (2.4). *Then for each fixed $m \le 1$,*

$$\left(\sum_{i=1}^{\tilde{N}_c} Z_{i-1}\right)^{1/2} \frac{\hat{m}_{\tilde{N}_c} - m}{\hat{\sigma}_{\tilde{N}_c}} \to_D N(0, 1) \quad \text{as } c \to \infty. \tag{4.2}$$

The proof of Theorems 4.1 and 4.2 depend on a lemma which is similar to condition (2.6) of Proposition 2.1 of Lai and Siegmund (1983). It is worth pointing out that the conclusion of Lemma A holds uniformly over a whole class of offspring distributions with $m \in (0, 1]$ and bounded variance, where as in Lai and Siegmund (1983) the corresponding result holds when the distribution of the error terms is fixed and only the autoregressive parameter varies. Lemma A will be used heavily to establish the asymptotic normality of $\hat{m}_{N_c}$ uniformly over a whole class of offspring distributions. The proof of Lemma A is given in the Appendix.

LEMMA A. *Assume the model* (2.1). *Then for each $\delta > 0$,*

$$\lim_{k\to\infty} \sup_{F\in G_{\sigma_1}} P_F\left\{Z_n \ge \delta \sum_{i=1}^{n} Z_{i-1} \text{ for some } n \ge k\right\} = 0, \tag{4.3}$$

where supremum is taken over $F \in G_{\sigma_1} = \{F\colon E_F(\xi_{1,1}) = m \in (0,1]$ and $\mathrm{Var}_F(\xi_{1,1}) = \sigma^2 \in (0, \sigma_1^2]\}$ with $\sigma_1^2 < \infty$ and known.

PROOF OF THEOREM 4.1. Note that

$$I_{N_c}^{1/2}\left(\hat{m}_{N_c} - m\right) = \sum_{i=1}^{I_{N_c}} (\xi_i - m)/I_{N_c}^{1/2}. \tag{4.4}$$

The desired result would follow from an application of the random sum central limit theorem [see Billingsley (1968), Theorem 17.1, page 146] once we show that

$$I_{N_c}/c\sigma^2 \to 1 \quad \text{a.s. as } c \to \infty \text{ for each } m \le 1. \tag{4.5}$$

From (2.3) it follows that

$$I_{N_c - 1} < c\sigma^2 \le I_{N_c},$$

which in turn implies that $I_{N_c-1}^{-1} c\sigma^2 \to 1$ a.s. as $c \to \infty$, provided we show that for each $m \le 1$,

$$Z_{N_c - 1} \Big/ \sum_{i=1}^{N_c - 1} Z_{i-1} \to 0 \quad \text{a.s. as } c \to \infty. \tag{4.6}$$

Since $N_c \to \infty$ a.s. for all $m \le 1$, the result (4.6) follows from Lemma A. Now, (4.5) follows easily. □

Proof of Theorem 4.2. Replace N_c in (4.4) by $\tilde{N}_c$. Use the definition of $\tilde{N}_c$, the fact that $\sigma_{\tilde{N}_c}^2 \to \sigma^2$ a.s. for each $m \le 1$, the random sum central limit theorem, Lemma A and argue as in Theorem 4.1 to get the desired result. □

We now state some properties of the stopping time N_c defined by (2.3). The proofs are omitted; see Sriram, Basawa and Huggins (1989) for details.

Theorem 4.3. *For the model* (2.1) *and stopping time* N_c *defined by* (2.3), *the following hold as* $c \to \infty$:

(i) *for each* $m < 1$, $c^{-1} N_c \to (1 - m)\sigma^2/\lambda$ *a.s.*

and

(ii) *for* $m = 1$, $c^{-1/2} N_c \to_D \inf\{t: \int_0^t Y(s)\,ds = 1\}$,

where $Y(s)$ *is as defined in Theorem* 3.2.

5. Uniform asymptotic normality of $\hat{m}_{N_c}$. The main result of this section is the uniform asymptotic normality of $\hat{m}_{N_c}$ (Theorem 5.2 below). Here the uniformity is with respect to a class of distribution functions $\mathscr{F}$, where

$$\mathscr{F} \subset \left\{ F: \int_{-\infty}^{\infty} (x - m)\,dF(x) = 0 \text{ and } \int_{-\infty}^{\infty} (x - m)^2\,dF(x) = \sigma_F^2 \in (0, \infty) \text{ for some } m \in (0, 1] \right\}$$

and satisfies the conditions

$$\sup_{F \in \mathscr{F}} \int_{\{|(x-m)| > a\}} (x - m)^2\,dF(x) = o(1) \quad \text{as } a \to \infty$$
$$\text{and } \inf_{F \in F} \sigma_F^2 > 0. \tag{5.1}$$

The method of proof of the main result is a very natural one. Our approach is motivated by the observation that

$$\hat{m}_n - m = \sum_{j=1}^{\sum_{i=1}^{n} Z_{i-1}} (\xi_j - m) \Big/ \sum_{i=1}^{n} Z_{i-1} \tag{5.2}$$

(randomly stopped average of i.i.d. random variables) where $\{\xi_j\}$ denotes the number of offspring of the jth individual disregarding the generations. We first state a version of the uniform central limit theorem (CLT) for i.i.d. mean zero r.v.'s which was originally due to Parzen (1954) [see also Datta (1990)], and use it to obtain a uniform version of the well-known Anscombe's theorem. The main result is then obtained as an application of the uniform Anscombe's theorem. Incidentally, it was pointed out by the referee that a uniform version of Anscombe's theorem is implicitly used in Siegmund (1982) in a way that somewhat resembles Theorem 5.1 below. We begin with the statement of the uniform CLT and a definition of the uniform version of uniform continuity in probability (u.c.i.p.), often referred to as Anscombe's condition.

LEMMA 5.1 (Uniform CLT). *Let*

$$\mathscr{G} \subset \left\{G\colon \int_{-\infty}^{\infty} x\,dG(x) = 0 \text{ and } \int_{-\infty}^{\infty} x^2\,dG(x) = \sigma_G^2 = \sigma^2 \in (0,\infty)\right\}.$$

Assume further that $\sup_{G\in\mathscr{G}}\int_{\{|x|>\sigma_G a\}}(x^2/\sigma_G^2)\,dG(x) = o(1)$ *as* $a \to \infty$. *Define* $S_n = \sum_{i=1}^{n} X_i$, *where* $X_1, X_2, \ldots$ *are i.i.d. random variables with distribution function* $G \in \mathscr{G}$. *Then for each real number* x,

$$\lim_{n\to\infty} \sup_{G\in\mathscr{G}} \left| P_G\left\{(\sigma^2 n)^{-1/2} S_n \le x\right\} - \Phi(x)\right| = 0, \tag{5.3}$$

where Φ *is the standard normal distribution function.*

NOTE. For the family $\mathscr{G}$ in Lemma 5.1, the condition (5.1) (with $m = 0$) implies the extra condition assumed in Lemma 5.1. This fact is used in our main theorem.

DEFINITION 5.1. A sequence W_n, $n \ge 1$, of random variables is said to be u.c.i.p. uniformly over a class $\mathscr{G}$ of distribution functions if and only if for every $\varepsilon > 0$, there is a $\delta > 0$ for which

$$\sup_{G\in\mathscr{G}} P_G\left\{\max_{0\le k\le n\delta} |W_{n+k} - W_n| > \varepsilon\right\} < \varepsilon \quad \text{for all } n \ge 1. \tag{5.4}$$

The proof of the next lemma and Theorem 5.1 below follow along the same lines as that of Example 1.8 and Theorem 1.4 of Woodroofe (1982). Hence, we omit the proofs.

LEMMA 5.2. *Consider the class* $\mathscr{G}$ *and* S_n, $n \ge 1$, *as defined in Lemma* 5.1. *Then* $S_n^* = S_n/(\sqrt{n}\,\sigma)$, $n \ge 1$, *is u.c.i.p. uniformly over the class* $\mathscr{G}$.

THEOREM 5.1 (Uniform Anscombe's theorem). *Assume the conditions of Lemma* 5.1. *Let* τ_c, $c > 0$, *be integer-valued random variables for which*

$$\lim_{c\to\infty} \sup_{G\in\mathscr{G}} P_G\{|\tau_c/[c\theta] - 1| > \varepsilon\} = 0 \tag{5.5}$$

for $\theta > 0$. *Then for* S_n^* *defined in Lemma* 5.2,

$$\lim_{c\to\infty} \sup_{G\in\mathscr{G}} \left|P_G\{S_{\tau_c}^* \le x\} - \Phi(x)\right| = 0. \tag{5.6}$$

We now return to the model (2.1). The principal result is:

THEOREM 5.2. *Define* $\hat{m}_n$, $n \ge 1$, *by* (2.2) *and* N_c *by* (2.3). *For the class* $\mathscr{F}$ *satisfying condition* (5.1), *we have*

$$\lim_{c\to\infty} \sup_{F\in\mathscr{F}} \left| P_F\left\{ \left(\sum_{i=1}^{N_c} Z_{i-1} \right)^{1/2} \frac{\hat{m}_{N_c} - m}{\sigma} \le x \right\} - \Phi(x) \right| = 0.$$

PROOF. Let $\tau_c = \sum_{i=1}^{N_c} Z_{i-1}$, $X_i = (\xi_i - m)$, $S_n = \sum_{i=1}^{n} X_i$ and use (5.2) to write

$$\left(\sum_{i=1}^{N_c} Z_{i-1} \right)^{1/2} \frac{\hat{m}_{N_c} - m}{\sigma} = S_{\tau_c}^*, \tag{5.7}$$

where S_n^* is as defined in Lemma 5.2. The theorem follows immediately from Theorem 5.1 once we have verified condition (5.5). To this end, let $\theta = \sigma^2$ in Theorem 5.1. By definition (2.3), $\tau_c > c\sigma^2$. Therefore

$$\begin{aligned} P_F\left\{\left|(c\sigma^2)^{-1}\tau_c - 1\right| > \varepsilon\right\} &\le P_F\left\{(c\sigma^2)^{-1}\tau_c - 1 > \varepsilon\right\} \\ &\le P_F\left\{ \left(Z_{N_c - 1} \Big/ \sum_{i=1}^{N_c - 1} Z_{i-1} \right) > \varepsilon \right\} \\ &\le P_F\left\{ Z_{N_c - 1} > \varepsilon \left(\sum_{i=1}^{N_c - 1} Z_{i-1} \right), N_c \ge k + 1 \right\} \\ &\quad + P\{N_c < k + 1\} \\ &\le P_F\left\{ Z_n > \varepsilon \sum_{i=1}^{n} Z_{i-1} \text{ for some } n \ge k \right\} \\ &\quad + P\{N_c < k + 1\}. \end{aligned} \tag{5.8}$$

Use Lemma A to choose a large k (and fix it) so that

$$\sup_{F\in\mathscr{F}} P_F\left\{ Z_n > \varepsilon \sum_{i=1}^{n} Z_{i-1} \text{ for some } n \ge k \right\} < \varepsilon/2. \tag{5.9}$$

For the fixed k and using $E_F(Z_i) = O(i)$ for all $F \in \mathscr{F}$, we have

$$\begin{aligned} P_F(N_c < k+1) &\le P_F\left(\sum_{i=1}^{n} Z_{i-1} \ge c\sigma^2 \text{ for some } n \le k\right) \\ &\le P_F\left(\sum_{i=1}^{k} Z_{i-1} \ge c\sigma^2\right) \\ &\le (c\sigma^2)^{-1} E_F \sum_{i=1}^{k} Z_{i-1} \\ &= O(k^2)/(c\sigma^2) \to 0 \quad \text{as } c \to \infty \qquad (5.10) \end{aligned}$$

uniformly over $\mathscr{F}$. The required result follows easily. Hence the theorem. □

A remaining question is whether the uniformity of Theorem 5.2 holds when σ^2 is unknown, as in Theorem 4.2. It is conjectured that the uniformity continues to hold for $\tilde{N}_c$ defined in Theorem 4.2 for the case of unknown σ^2.

APPENDIX

PROOF OF LEMMA A. Recall that we set $I_n = \sum_{i=1}^{n} Z_{i-1}$. Write

$$\frac{Z_n^2}{I_n^2} = \frac{Z_n^2}{(I_n+n)^2}\left[1 + \frac{n}{I_n}\right]^2 \le \frac{Z_n^2}{(I_n+n)^2}\left[1 + \left(\frac{n}{\sum_{i=1}^{n-1} Y_i}\right)\right]^2. \qquad (A.1)$$

Since $1 + (n/\sum_{i=1}^{n-1} Y_i) \to 1 + \lambda^{-1}$ a.s. as $n \to \infty$, uniformly over $F \in G_{\sigma_1}$ (defined in Section 4), it suffices to show that

$$\lim_{k\to\infty} \sup_{F \in G_{\sigma_1}} P_F\left\{Z_n^2 \ge \delta^2\left[\sum_{i=1}^{n} (Z_{i-1}+1)\right]^2 \text{ for some } n \ge k\right\} = 0. \qquad (A.2)$$

Let $M_n = Z_n^2/(I_n + n)^2$. Then

$$E\{M_{n+1} | \mathscr{F}_n\} = \frac{m^2 Z_n^2 + (\sigma^2 + 2m\lambda) Z_n + EY_1^2}{[I_{n+1} + (n+1)]^2} \le M_n + \zeta_n, \qquad (A.3)$$

which satisfies condition (1) of Robbins and Siegmund (1971) for $\zeta_n = \{(\sigma^2 + 2\lambda) Z_n + EY_1^2\}/[I_{n+1} + (n+1)]^2$. Now apply Proposition 2 of Robbins and Siegmund (1971) to get

$$P\left\{\max_{n \ge k} M_n \ge \delta^2\right\} \le \delta^{-2}\left\{EM_k + E\sum_{n=k}^{\infty} \zeta_n\right\} \to 0 \quad \text{as } k \to \infty \qquad (A.4)$$

uniformly over $F \in G_{\sigma_1}$ since

$$E \sum_{n=k}^{\infty} \zeta_n \le (\sigma_1^2 + 2\lambda) E \sum_{n=k}^{\infty} \frac{I_{n+1} - I_n}{[I_{n+1} + (k+1)]^2} + EY_1^2 \sum_{n=k}^{\infty} (n+1)^{-2}$$
$$\le (\sigma_1^2 + 2\lambda) \int_0^{\infty} [(k+1) + x]^{-2}\, dx + EY_1^2 \int_k^{\infty} x^{-2}\, dx$$
$$\to 0 \quad \text{as } k \to \infty, \text{ uniformly over } F \in G_{\sigma_1}$$

and

$$EM_k \le EZ_{k-1}^2/[I_{k-1} + k]^2 + (\sigma_1^2 + 2\lambda) EZ_{k-1}/[I_k + k]^2 + k^{-2} EY_1^2$$
$$\le \cdots \le E(\sigma_1^2 + 2\lambda) \sum_{i=1}^{k} Z_{i-1}/[k + I_i]^2 + k^{-1} EY_1^2$$
$$\to 0 \quad \text{as } k \to \infty \text{ uniformly over } F \in G_{\sigma_1}. \qquad \square$$

Acknowledgments. The first author would like to express his thanks to Professor W. P. McCormick for many useful discussions on the issue of uniform asymptotic normality given in Section 5. The first author would also like to thank Professor S. Datta for pointing out references to the uniform CLT stated in Lemma 5.1. We thank the referees and an Associate Editor for a careful reading of the paper and their detailed constructive suggestions.

REFERENCES

BASAWA, I. V. and SCOTT, D. J. (1976). Efficient tests for branching processes. *Biometrika* **63** 531–536.

BASAWA, I. V. and SCOTT, D. J. (1983). *Asymptotic Optimal Inference for Non-Ergodic Models. Lecture Notes in Statist.* **17**. Springer, New York.

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

CHOW, Y. S. and ROBBINS, H. E. (1965). On the asymptotic theory of fixed width confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

DATTA, S. (1990). A note on continuous Edgeworth expansions and the bootstrap. *Sankhyā Ser. A*. To appear.

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

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.

LAI, T. L. and SIEGMUND, D. (1983). Fixed accuracy estimation of an autoregressive parameter. *Ann. Statist.* **11** 478–485.

NANTHI, K. (1983). *Statistical Estimation for Stochastic Process. Queen's Papers in Pure Appl. Math.* **62**. Queen's Univ.

PAKES, A. G. (1971). Branching process with immigration. *J. Appl. Probab.* **8** 32–42.

PARZEN, E. (1954). On uniform convergence of families of sequences of random variables. *University of California Publications in Statistics* **2** 23–54.

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

SIEGMUND, D. (1982). A sequential confidence interval for the odds ratio. *Probab. Math. Statist.* 149–156.

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

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.

WOODROOFE, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*. SIAM, Philadelphia.

DEPARTMENT OF STATISTICS
204 STATISTICS BUILDING
UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602

DEPARTMENT OF STATISTICS
LA TROBE UNIVERSITY
BUNDOORA, VICTORIA 3083
AUSTRALIA