The Annals of Probability
1982, Vol. 10, No. 4, 1075–1078

SOME LIMIT THEOREMS ON A SUPERCRITICAL SIMPLE GALTON-WATSON PROCESS

BY K. N. VENKATARAMAN AND K. NANTHI

University of Madras

Let $X = (X_n;\ n \geq 0;\ X_0 = 1)$ be a supercritical Galton-Watson process possessing an offspring mean $1 < m < \infty$, and variance $0 < \sigma^2 < \infty$. The limiting distribution of $\{X_n^{-1/2}(X_{n+r} - \hat{m}^r X_n);\ r = 2, \cdots, T\}$ where $\hat{m} = X_{n+1}/X_n$, is obtained. As a consequence of this result a Quenouille-Bartlett type of asymptotic goodness of fit test is also proposed for the process X.

1. Introduction. In this paper we are concerned with a supercritical Galton-Watson process $X = (X_n;\ n \geq 0;\ X_0 = 1)$ defined on a Tree probability space (Harris, 1963). Let $1 < m < \infty$, and $0 < \sigma^2 < \infty$ be the offspring mean and variance of X respectively. It is well known (Athreya and Ney, 1972, page 19) that, as $n \to \infty$, $m^{-n}X_n$ converges almost surely to a nonnegative, non-degenerate random variable ω, say, and $P(\omega > 0) = 1 - P(\text{extinction of } X)$.

The results on X, presented in this paper, are motivated by the following theorem stated without proof.

THEOREM A. (Jagers, 1975, page 38). *Let us define that*

$$\begin{aligned} Y_n(r) &= \sigma^{-1} m^{1-r} X_n^{1/2}(X_{n+r} - m^r X_n) && \text{when} \quad X_n > 0 \\ &= 0 && \text{when} \quad X_n = 0. \end{aligned}$$

Then, under $P(\cdot \mid \omega > 0)$, the following statements hold.
(a) $(Y_n(r);\ r = 1, \cdots, T)$ *converges in law, as $n \to \infty$, to a normal random vector with mean zero, and a non-singular covariance matrix $[C_{rs}]$, where, for $r, s, \geq 1$,*

$$C_{rs} = m^{1-v}(m^v - 1)(m - 1)^{-1}; \qquad v = \min(r, s).$$

(b) $(Y_n(r + 1) - Y_n(r);\ r = 1, \cdots, T)$ *converges in law, as $n \to \infty$, to a normal vector with mean zero, and, the covariance matrix* $\operatorname{diag}\{m^{-r};\ r = 1, \cdots, T\}$.

In place of $Y_n(r)$, we study here the sample functions

$$\begin{aligned} \hat{Y}_n(r) &= X_n^{-1/2}(X_{n+r} - \hat{m}^r X_n) && \text{when} \quad X_n > 0 \\ &= 0 && \text{when} \quad X_n = 0 \end{aligned} \tag{1.1}$$

where

$$\begin{aligned} \hat{m} &= X_{n+1}/X_n && \text{when} \quad X_n > 0 \\ &= 0 && \text{when} \quad X_n = 0 \end{aligned} \tag{1.2}$$

is recognized as the Lotka-Nagaev estimator (Nagaev, 1967) for m. As a basic result we obtain the limiting distribution of $(\hat{Y}_n(r);\ r = 2, \cdots, T)$ (Theorem 2.1), which leads to a parameter-free result (Theorem 2.4). To highlight the relevance of Theorem 2.4, it is observed that for a classical supercritical Galton-Watson process, no goodness of fit tests have so far been proposed with/without specifying an alternative. Although a subcritical

Received August 1980; revised December 1981.
AMS 1980 *subject classifications.* Primary, 62M99; secondary 60F99.
Key words and phrases. Supercritical Galton-Watson process, asymptotic goodness of fit test.

Galton-Watson process may be considered as an alternative, the standard likelihood ratio type approach does not lend to any satisfactory treatment since under the alternative, the unknown parameters defy estimation. In view of these difficulties it will be prudent to look for a Quenouille-type of goodness of fit test, that is advocated by Bartlett (Bartlett, 1955) in time series analysis, on the considerations that this type of test is formulated purely in terms of the null hypothesis, making no reference to the alternative. In our opinion Theorem 3.3 certainly provides such a goodness of fit test.

2. A Basic results on $\hat{Y}_n(r)$.

The fundamental result of this paper is:

THEOREM 2.1. *Under* $P(\cdot \mid \omega > 0)$, $(\hat{Y}_n(r);\ r = 2, \cdots, T)$ *converges in law, as* $n \to \infty$, *to a normal random vector* $(Y(r);\ r = 2, \cdots, T)$, *say, having zero mean and covariance matrix* $[h_{rs}]$, *where, for* $r, s \geq 2$,

(i) $$h_{rs} = (m-1)^{-1}\sigma^2[(r-1)(s-1)(m-1)m^{r+s-2} + m^{u-1}(m^{v-1} - 1)]$$

(ii) $$u = \max(r, s), \quad \text{and} \quad v = \min(r, s).$$

We split the proof of this theorem into a few lemmas.
Let, for $r \geq 2$

$$\hat{\theta}_r = (\hat{m} - m)^{-1}(\hat{m}^r - m^r) \quad \text{when} \quad \hat{m} \neq m$$
$$= r\hat{m}^{r-1} \quad \text{when} \quad \hat{m} = m \tag{2.1}$$

and $\theta_r = rm^{r-1}$.

LEMMA 2.1. *Under* $P(\cdot \mid \omega > 0)$, (a) $\hat{\theta}_r \to_P \theta_r$ *and* (b) $X_n^{1/2}(\hat{m} - m)(\hat{\theta}_r - \theta_r) \to_P 0$.

NOTE. The simple proof of (a) is omitted. An appeal to Dion's Theorem (Dion, 1974), together with (a) yields (b).

Next we define that

$$\eta_{i,r} = (Z_{n,r}^{(i)} - m^r) - \theta_r(Z_{n,1}^{(i)} - m); \qquad r \geq 1 \tag{2.2}$$

where $Z_{n,r}^{(i)}$ is interpreted as the size of the rth generation of offspring flowing from the ith individual among the X_n individuals in the nth generation of X. These are random variables since X is defined on the Harris tree probability space.

The routine manipulative proof of the following lemma concerning (2.2) is omitted.

LEMMA 2.2. $E(\eta_{i,r}\eta_{i,s}) = h_{rs}; \qquad r, s \geq 2.$

Our next lemma relates to

$$\delta_i = \textstyle\sum_{r=2}^{T} \alpha_r \eta_{i,r}; \qquad i = 1, \cdots, X_n (> 0) \tag{2.3}$$

where α_r are real numbers, not all of them being zero, but otherwise arbitrary.

LEMMA 2.3. *Under* $P(\cdot \mid \omega > 0)$, $X_n^{-1/2}\sum_{i=1}^{X_n} \delta_i$ *converges in law, as* $n \to \infty$, *to a normal random variable with mean zero and variance* $\sum_{r,s=2}^{T} \alpha_r \alpha_s h_{rs}$.

PROOF. Use Theorem 2 of the appendix in Jagers (1975) and properties of ω.

We are now settled to prove Theorem 2.1. By definition, on $(\omega > 0)$,

$$\hat{Y}_n(r) = \textstyle\sum_{i=1}^{X_n} \eta_{i,r} - X_n^{1/2}(\hat{m} - m)(\hat{\theta}_r - \theta_r); \qquad r \geq 2 \tag{2.4}$$

so that

$$\hat{Y}_n = \textstyle\sum_{r=2}^{T} \alpha_r \hat{Y}_n(r) = X_n^{-1/2} \sum_{i=1}^{X_n} \delta_i - \sum_{r=2}^{T} \alpha_r X_n^{1/2}(\hat{m} - m)(\hat{\theta}_r - \theta_r). \tag{2.5}$$

Based on (2.5) the proof of Theorem 2.1 follows from application of Lemmas 2.1, 2.3, a Slutsky type theorem (See Loeve, 1963, page 168), and Cramer-Wold device (See Billingsley, 1968, page 48). Details are omitted.

3. Applications of Theorem 2.1. The following application of Theorem 2.1, stated without proof relates to

$$V_n(r) = \hat{Y}_n(r+2) - 2m\hat{Y}_n(r+1) + m^2\hat{Y}_n(r); \qquad r \geq 2. \tag{3.1}$$

THEOREM 3.1. *Under $P(\cdot \mid \omega > 0)$, $(V_n(r); r = 2, \cdots, T)$ converges in law, as $n \to \infty$, to a normal vector $(U(r), r = 2, \cdots, T)$, say, with mean zero, and a non-singular covariance matrix implying that, for $r, s \geq 2$,*

(i) $$E\,U^2(r) = \sigma^2(m+1)m^{r+1}$$

(ii) $$E[U(r)U(s)] = 0 \quad \text{for} \quad |r-s| \geq 2.$$

NOTE. An easy way to arrive at Theorem 3.1 is to derive the limiting distribution of $(\hat{Y}_n(r+1) - m\hat{Y}_n(r); r = 2, \cdots, T)$, and to note that

$$V_n(r) = (\hat{Y}_n(r+2) - m\hat{Y}_n(r+1)) - m(\hat{Y}_n(r+1) - m\hat{Y}_n(r)). \tag{3.2}$$

Next, let us define that

$$\hat{V}_n(r) = \hat{Y}_n(r+2) - 2\hat{m}\hat{Y}_n(r+1) + \hat{m}^2\hat{Y}_n(r), \qquad r \geq 2. \tag{3.3}$$

We note that, under $P(\cdot \mid \omega > 0)$, $\hat{V}_n(r) - V_n(r) \to_P 0$, as $n \to \infty$ on account of the fact that (a) $\hat{Y}_n(r)$ converges in law, and (b) $(\hat{m} - m) \to_P 0$, as $n \to \infty$. This remark together with Theorem 3.1, yields:

THEOREM 3.2. *Under $P(\cdot \mid \omega > 0)$, $(\hat{V}_n(r); r = 2, \cdots, T)$ converges in law, as $n \to \infty$, to the normal random vector $(U(r); r = 2, \cdots, T)$, defined in Theorem 3.1.*

For computational purposes, it is relevant to observe that

$$\hat{V}_n(r) = X_n^{-1/2}(X_{n+r+2} - 2\hat{m}X_{n+r+1} + \hat{m}^2X_{n+r}) \quad \text{when} \quad X_n > 0. \tag{3.4}$$

Let $\hat{\sigma}^2$ be the estimator proposed for σ^2, either by Heyde (1974), or Dion (1975), and,

$$\hat{\rho}^2(r) = \hat{\sigma}^2(\hat{m}+1)\hat{m}^{r+1}; \; r \geq 2. \tag{3.5}$$

An application of Theorem 3.2 together with standard results in probability yields the following parameter-free result on X.

THEOREM 3.3. *Let $r_1, \cdots, r_T$ be fixed positive integers such that, for $i, j = 1, \cdots, T$, $r_i \geq 2$, and, $|r_i - r_j| \geq 2$ when $i \neq j$. Then, under $P(\cdot \mid \omega > 0)$*

$$((\hat{\rho}(r_1))^{-2}\hat{V}_n^2(r_1) + \cdots + (\hat{\rho}(r_T))^{-2}\hat{V}_n^2(r_T))$$

converges in law, as $n \to \infty$, to a Chi squared variable with T degrees of freedom.

4. Concluding remarks. Making use of the facts that (a) $(\omega > 0) \subset (X_n > 0)$, and (b) as $n \to \infty$, $P(X_n > 0) \to P(\omega > 0)$, it can be shown that as $n \to \infty$, the statements

$$\begin{aligned} &P(\xi_n \leq \lambda \mid \omega > 0) \to \text{p.d.f.} \quad F(\lambda) \\ &P(\xi_n \leq \lambda \mid X_n > 0) \to \text{p.d.f.} \quad F(\lambda) \end{aligned} \tag{4.1}$$

are equivalent. Thus Theorems 2.1, 3.1, 3.2, and 3.3 are valid under $P(\cdot \mid X_n > 0)$ in place of $P(\cdot \mid \omega > 0)$.

Acknowledgment. The authors place on record their sincere thanks to a referee for his very useful remarks on an earlier version of this paper.

REFERENCES

ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes.* Springer, Berlin.

BARTLETT, M. S. (1955). *An Introduction to Stochastic Processes with Special Reference to Methods and Applications.* Cambridge University Press.

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

DION, J. P. (1974). Estimation of mean and the initial probabilities of a branching process. *J. Appl. Probability* **11** 687–694.

DION, J. P. (1975). Estimation of the variance of a branching process. *Ann. Statist* **3** 1183–1187.

HARRIS, T. E. (1963). *The Theory of Branching Processes.* Springer, Berlin.

HEYDE, C. C. (1974). On estimating the variance of the offspring distribution in a simple branching process. *Adv. Appl. Probability* **6** 421–433.

JAGERS, P. (1975). *Branching Processes with Biological Applications.* Wiley, New York.

LOÈVE, M. (1968). *Probability Theory.* Affiliated East-West Press Pvt. Ltd., New Delhi.

NAGAEV, A. V. (1967). On estimating the expected number of direct descendents of a particle in a branching process. *Theor. Probability Appl.* **12** 314–320.

DEPARTMENT OF STATISTICS
UNIVERSITY OF MADRAS
MADRAS, INDIA