The Annals of Mathematical Statistics
1968, Vol. 39, No. 6, 2098–2102

ON RECENT THEOREMS CONCERNING THE SUPERCRITICAL GALTON-WATSON PROCESS

BY E. SENETA

Australian National University, Canberra

1. Introduction. We consider a Galton-Watson process $\{Z_n\,;\, n = 0, 1, 2, \cdots\}$ initiated by a single ancestor, whose offspring distribution has probability generating function $F(s) = \sum_{j=0}^{\infty} s^j P[Z_1 = j]$, $s \varepsilon [0, 1]$, and $P[Z_1 = j] \neq 1$ for any $j = 0, 1, 2, \cdots$. In the present note, we are concerned only with the supercritical case, when $1 < m \equiv E[Z_1] < \infty$, in which case it is well known that the probability of extinction, q, is the unique real number in $[0, 1)$ satisfying $F(q) = q$. We recall that the generating function, $F_n(s)$, of Z_n is the nth functional iterate of $F(s)$ for the Galton-Watson process in general, and in the supercritical case $F_n(s) \to q$ as $n \to \infty$ for $s \varepsilon [0, 1)$. In particular $F_n(0) \uparrow q$.

Recently, a considerable amount of research has been devoted to refinements of the classical theorem concerning the convergence of the random variables (Z_n/m^n), $n = 0, 1, 2, \cdots$, (for a history of the theorem prior to these, see Harris [1]). In particular, an ultimate form of the theorem has been obtained by Kesten and Stigum [2], [6], who prove that these random variables converge a.e. to a random variable W, for which $P[W = 0] = q$ or 1, and which has a continuous density on the set of positive real numbers. Moreover $E[Z_1 \log Z_1] < \infty \Leftrightarrow P[W = 0] = q \Leftrightarrow E[W] = 1$.

It therefore follows that $E[Z_1 \log Z_1] = \infty \Leftrightarrow P[W = 0] = 1$.

Thus while Kesten and Stigum have provided a complete answer for the classical norming, by m^n, of the random variables Z_n, the limit r.v. may still be degenerate at the origin. This leads us to ask whether there exists a sequence of constants, c_n, such that (Z_n/c_n) *always* converge, in some sense, to a proper non-degenerate r.v.

We provide a partial answer to this question by producing such a sequence, for which the variables (Z_n/c_n) converge *in distribution* to such a limit variable W, for which $P[W = 0] = q$. Moreover $E[Z_1 \log Z_1] < \infty \Leftrightarrow E[W] < \infty \Leftrightarrow c_n \sim$ const m^n.

It is also shown that in this situation the random variables (Z_n/c_n) form a submartingale, although this does not appear sufficient to assert a.e. convergence.

2. Preliminary considerations. It turns out that it is relevant to use, instead of the generating function $F(s)$, the function

$$k(s) = -\log F(e^{-s}), \qquad s \geqq 0,$$

which we shall call the cumulant generating function (cgf) of Z_1. It is readily checked that the cgf of Z_n, $k_n(s)$ i.e.

Received 9 November 1967.

$$k_n(s) = -\log F_n(e^{-s}), \qquad s \geqq 0,$$

is in fact the nth functional iterate of $k(s)$.

Under the assumptions made about the process $\{Z_n\}$, it is readily seen that $k(s)$ is strictly monotone increasing and strictly concave for $s \geqq 0$. We also note in particular that $k(0) = 0$, $k'(0+) = m$, and $k(s) > s$ for $0 < s < r$, where $r = -\log q$; if $q > 0$ then r is the unique solution in $(0, \infty)$ of $k(s) = s$. Analogous properties hold for $k_n(s)$, $n \geqq 2$.

The continuity and strict monotonicity of $k(s)$ for $s \geqq 0$ imply that its inverse $h(\cdot) = k^{-1}(\cdot)$ exists for values of s in a right neighbourhood of the origin (in fact for $0 \leqq s < -\log F(0)$) and has properties "dual" to those of $k(s)$. Moreover the nth iterate of $h(s)$, $h_n(s)$ which is well defined for $0 \leqq s < -\log F_n(0)$, is the inverse of $k_n(s)$. (We note that $-\log F_n(0) \geqq r$ for $n \geqq 1$, since $F_n(0) \uparrow q$.) Some specific properties of $h(s)$ are emphasized in the proof of the following basic theorem.

THEOREM 2.1. *For $s \in [0, r)$ the sequence $h_n(s)/h_n(s_0)$ (where $s_0 \in (0, r)$ is fixed) approaches a finite limit $H(s)$ which is positive for $s \in (0, r)$, and is a solution of the Schröder functional equation*

$$H(h(s)) = m^{-1}H(s). \tag{2.1}$$

Further $H(s)$ is, up to constant factors, the unique solution of (2.1) such that $H(s)/s$ is monotonic in $(0, r)$.

PROOF. It is necessary only to check that the conditions in the note of Kuczma [3] are satisfied. These are that (i) $h(s)$ be continuous and strictly increasing in $[0, r)$, (ii) $h(0) = 0$ and $0 < h(s) < s$ for $s \in (0, r)$, (iii) $\lim_{s \to 0+} h(s)/s = m^{-1}$, and (iv) $h(s)/s$ be monotonic in $(0, r)$.

All these follow from the properties of $k(s)$; in particular (iv) is a consequence of the convexity of $h(s)$ (since $k(s)$ is concave) in conjunction with $h(0) = 0$.

The fact that $h(s)$ has properties stronger than those required for the applicability of Kuczma's result, leads to the following additional information on the limit function $H(s)$.

THEOREM 2.2. *The function $H(s)$ is continuous and strictly monotone increasing $[0, r)$. Moreover $H(s) \to \infty$ as $s \to r-$.*

PROOF. Since the function $H(s)$ is the limit of a sequence of functions, each of which is increasing and convex in $[0, r)$, it is itself increasing and convex in this interval.

Its convexity in $(0, r)$ implies its continuity in this interval; while its monotonicity in the same interval implies that $\lim_{s\to 0+} H(s) \equiv H(0+)$ exists. Further, since $H(0) = 0$ and $H(s)$ is convex in $[0, r)$ and positive in $(0, r)$, $H(0+) = 0 = H(0)$.

To prove *strict* monotonicity in $[0, r)$ assume to the contrary that $H(s) = c \equiv$ const., for $s \in [\alpha, \beta] \subset (0, r)$ where $\beta > \alpha$. Then $c > 0$, since $H(s) > 0$ in $(0, r)$, and $H(s)/s \equiv c/s$ in $[\alpha, \beta]$, is strictly decreasing for $s \in [\alpha, \beta]$, whereas by convexity of $H(s)$, it should be increasing. Thus we have arrived at a contradiction.

Finally from (2.1), letting $s \to r-$,

$$H(r-) = m^{-1}H(r-)$$

in the sense that both sides are finite or infinite. Clearly, since $0 < m^{-1} < 1$, $H(r-) = \infty$.

Finally, the sequel depends heavily on the following result, which we state as a lemma, without proof.

LEMMA 2.1. *Suppose* $\{f_{(n)}(s)\}$, $n \geqq 1$, *is a sequence of functions defined for a corresponding sequence of intervals* $[0, a_n)$, *and each is continuous and strictly monotonic increasing, with* $f_{(n)}(0) = 0$ *and* $f_{(n)}(s) \to \infty$ *as* $s \to a_n-$ *for each* n. *Suppose further that* $f_{(n)}(s) \to f(s)$ *as* $n \to \infty$ *for* $s \,\varepsilon\, [0, a)$ *where for each* n, $\infty \geqq a_n \geqq a > 0$, *and where* $f(s)$ *is continuous and strictly monotone increasing in* $[0, a)$ *such that* $f(s) \to \infty$ *as* $s \to a-$. *Then, for the inverse functions, as* $n \to \infty$

$$f_{(n)}^{-1}(s) \to f^{-1}(s), \qquad s \,\varepsilon\, [0, \infty).$$

3. Extension of the classical results. We shall prove the results announced in Section 1 in several stages.

THEOREM 3.1. *There exists a sequence of positive constants* c_n, $(c_n \to \infty)$ *such that the random variables* $W_n = Z_n/c_n$ *converge in distribution to a proper non-degenerate random variable* W *such that* $P[W = 0] = q$. *The* cgf *of* W, $K(s)$, *satisfies the Poincaré functional equation*

$$K(ms) = k(K(s)), \qquad s \geqq 0, \tag{3.1}$$

and is the unique strictly monotone increasing concave solution of it with $K(0+) = 0$, *apart from a scale factor (i.e. the only other such solutions are* $K(s/c)$, $0 < c =$ const).

PROOF. Consider the sequence of random variables W_n, where $W_n = h_n(s_0)Z_n$ where $s_0 \,\varepsilon\, (0, r)$ is fixed. Then the cgf of W_n is $k_n(h_n(s_0)s)$, and its inverse, well defined for a neighbourhood of the origin, is $h_n(s)/h_n(s_0)$.

Now, in fact $h_n(s)/h_n(s_0) \to \infty$ as $s \to a_n-$, where $a_n = -\log F_n(0)$, and $a_n \geqq r$. Moreover $h_n(s)/h_n(s_0) \to H(s)$ for $s \,\varepsilon\, [0, r)$ from Theorem 2.1, where $H(s)$ is continuous and strictly monotone increasing in this interval, with $H(s) \to \infty$ as $s \to r-$, from Theorem 2.2. Hence by Lemma 2.1,

$$k_n(h_n(s_0)s) \to K(s), \qquad s \geqq 0,$$

where $K(\cdot) \equiv H^{-1}(\cdot)$, and so is concave, continuous and strictly monotone increasing in $[0, \infty)$. The continuity theorem (for Laplace transforms) then yields the assertion of convergence in distribution to a proper rv, with $c_n \equiv 1/h_n(s_0)$, where $h_n(s_0) \to 0$ as $n \to \infty$.

Since $H(s)$ satisfies (2.1), it follows that $K(s)$ satisfies (3.1); the uniqueness assertion for (3.1) follows from that of (2.1) also, by a consideration of inverses. Moreover, the possibility of W being degenerate at a single point, i.e., $K(s) =$ const s(const $\geqq 0$) is excluded in the case const $= 0$ by the strict monotonicity of $K(s)$; and in the case const > 0 since then substitution of $K(s)$ in (3.1) yields that Z_1 has a linear cgf, which contradicts an initial assumption concerning the degeneracy of Z_1 at a single point.

Finally the concentration at 0 of the distribution of W may be determined by considering $K(s)$ as $s \to \infty$: let us call this limit $K(\infty)$ ($\leqq \infty$). Then (3.1) yields $K(\infty) = k(K(\infty))$ in the sense that both sides are finite or infinite. It is now easy to see that $K(\infty) = r$ ($r \leqq \infty$) so that $P[W = 0] = q$.

In conclusion to Theorem 3.1 we note that any sequence of positive constants c_n, for which the Z_n/c_n tend in distribution to a proper non-degenerate random variable, must be essentially unique. This is a direct consequence of Khintchine's theorem on convergence of (positive) types (see e.g. [4], p. 203). From the proof of Theorem 3.1 it then follows that, as $n \to \infty$,

$$c_n \sim \text{const}/h_n(s_0). \tag{3.2}$$

It is now relevant to explore briefly the connection of the above results with those of Kesten and Stigum. First we notice that if $E[Z_1 \log Z_1] < \infty$, the results of these authors, together with the theorem of Khintchine, imply that $h_n(s_0) \sim \text{const } m^{-n}$ as $n \to \infty$, and that $E[W] < \infty$, *where* W, *here and in the sequel, is the limit* rv *of Theorem* 3.1. On the other hand, $h_n(s_0) \sim \text{const } m^{-n}$ implies, from Theorem 3.1 and the Kesten-Stigum results, that $E[Z_1 \log Z_1] < \infty$. Hence to show that

$$E[W] < \infty \Leftrightarrow h_n(s_0) \sim \text{const } m^{-n} \Leftrightarrow E[Z_1 \log Z_1] < \infty \tag{3.3}$$

we need only show $E[W] < \infty \Rightarrow h_n(s_0) \sim \text{const } m^{-n}$. From (2.1) by iteration $H(s) = m^n H(h_n(s))$ for $s \in (0, r)$; and since $H(s_0) = 1$

$$1 = H(h_n(s_0))(h_n(s_0))^{-1} m^n h_n(s_0), \qquad n \geqq 1. \tag{3.4}$$

We note that since $h_n(s)$ is convex, and its slope at the origin is m^{-n}, $m^n h_n(s_0)/s_0 > 1$ for $n \geqq 1$; and recall that $h_n(s_0) \to 0$ as $n \to \infty$. We remark also that $H'(0+)$ exists and is in fact $1/K'(0+)$.

From the identity (3.4) it therefore follows that $H'(0+)$ is positive or zero, depending on whether $\lim_{n\to\infty} m^n h_n(s_0) < \infty$ or $= \infty$. Thus

$$K'(0+) \equiv E[W] < \infty \Leftrightarrow h_n(s_0) \sim \text{const } m^{-n}.$$

Finally, let us note that the sequence (W_n) where $W_n = h_n(s_0)Z_n$, is a submartingale, since W_n is Markovian and

$$E[W_{n+1} \mid W_n] = h_{n+1}(s_0) m Z_n = m h_{n+1}(s_0) W_n / h_n(s_0) > W_n$$

since $mh(h_n(s_0))/h_n(s_0) > 1$ as in the last part of the proof of (3.3).

However, in order to assert a.e. convergence of the W_n to W, it appears that some condition of the nature $E[W_n] < C \equiv \text{const}$ for all n, is required ([4], p. 393). Since

$$E[W_n] = h_n(s_0) m^n$$

it follows from the above discussion that $E[W_n]$ is bounded if and only if

$E[Z_1 \log Z_1] < \infty$, in which case we already know that a.e. convergence takes place since (Z_n/m^n) is a martingale, with mean one for every term, and hence converges a.e.

In conclusion, we remark that the recent paper [5] contains an assertion (Theorem c) closely related to our subject matter.

The author is grateful to the referee for some helpful suggestions.

REFERENCES

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

[2] KESTEN, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.* **37** 1211–1223.

[3] KUCZMA, M. (1964). Note on Schröder's functional equation. *J. Austral. Math. Soc.* **4** 149–151.

[4] LOÈVE, M. (1963). *Probability Theory*. Van Nostrand, Princeton.

[5] NAGAEV, A. V. and BADALBAEV, I. (1967). A refinement of certain theorems on branching random processes. (in Russian.) *Litovsk. Mat. Sb.* **7** 129–136.

[6] STIGUM, B. P. (1966). A theorem on the Galton-Watson process. *Ann. Math. Statist.* **37** 695–698.