Søren Asmussen
Norman Kaplan

Branching Random Walks I

Preprint
Februar
1975

4

Institute of Mathematical Statistics
University of Copenhagen

Søren Asmussen and Norman Kaplan*

BRANCHING RANDOM WALKS I

Preprint 1975 No. 4

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

*On leave from: Dept. of Statistics University of California
Berkeley, California 94701

KEYWORDS

Branching random walks, age-dependent branching processes, mean square convergence, central limit analogues, decomposable multitype Galton-Watson processes.

ABSTRACT

A general method is developed with which various theorems on the mean square convergence of functionals of branching random walks are proven. The results cover extensions and generalizations of classical central limit analogues as well as a result of a different type.

1. INTRODUCTION

Consider a branching random walk, i.e. a Galton-Watson process $\{Z_n\}$ with offspring distribution $\{p_j\}$, on which is superimposed the additional structure of random walk on the line. A particle whose parent is at x, moves to x + y and the y's of different particles are i.i.d. with common distribution function G. This model covers as an important special case the age-dependent branching process, where the position of a particle corresponds to the time of its death. Some attention has been given to study the distribution of the particles of the n^{th} generation over the line. As usual in branching processes, this problem is closely related to the behaviour of the mean, which in the present context, requires a study of the convolution powers of G. A number of results closely related to the central limit theorem have already been proven ([2], [3, Ch.6], [4], [5], [7], [11], [12], [13], [14], [15], [16]) and the case where G is in the domain of attraction of a stable law, has also been considered ([2]). We give in Section 3 some further results and generalizations along these lines, while Section 4 is devoted to a different type of limiting behaviour. We consider there the situation $G(0-) = 0$

$G(0) > 0$ and obtain a result, which for G lattice is closely related to the limit theory of decomposable multitype Galton-Watson processes ([10]).

While the method of proof in some of the references cited is highly analytic, we present in Section 2 a simple and general conditioning argument, which reduces the problem to a study of the convolution powers of G. We feel, that this is the natural approach, because it is based on the structure of the process and it is of sufficient generality to give unified and simple proofs of known and new results as well as to deal with certain generalizations.

2. PRELIMINARIES

Following the notation of [9, Ch. 5], we denote any particle of the n^{th} generation by $<\underset{\sim}{i}_n> = <i_1 i_2 \ldots i_n>$ and its position on the line by $X_{\underset{\sim}{i}_n}$. Then

$$X_{i_1 i_2 \ldots i_n} = X_{i_1 i_2 \ldots i_{n-1}} + Y_{i_1 i_2 \ldots i_n}$$

where the $Y_{\underset{\sim}{i}_n}$'s are i.i.d. with law governed by G. We introduce some more notation: Let $n \geq 0$

$Z_n(A) = \sum_{\underset{\sim}{i}_n} 1_{\{X_{\underset{\sim}{i}_n} \in A\}}$ = the number of particles of the n^{th} generation in A (A a Borel subset of the line)

$\mathcal{F}_n = \sigma(X_{\underset{\sim}{i}_k}, k \leq n)$

$Z_n(\underset{\sim}{i}_k)$: The number of offsprings of $\underset{\sim}{i}_k$ at time $n \geq k$

G_n: The n^{th} convolution of G

$G_0(x) = 0$ for $x < 0$, $G_0(x) = 1$ for $x \geq 0$

$f_n(x) = \int_{-\infty}^{\infty} f(x+y)\,dG_n(y)$ where f is any bounded measurable function.

The assumptions on the position $X_{\underset{\sim}{i}_0}$ of the original particle will be of minor importance. It is usually assumed, that $X_{\underset{\sim}{i}_0} = 0$ for the random walk situation, while $X_{\underset{\sim}{i}_0}$ is taken to be randomly distributed according to G in the age-dependent case. Thus the distribution of $X_{\underset{\sim}{i}_n}$ is G_n, respectively G_{n+1} in the two cases. We treat for convenience only the case $X_{\underset{\sim}{i}_0} = 0$. Also, we assume throughout that $Z_0 = 1$, $1 < m = \Sigma j p_j < \infty$ and $\Sigma j^2 p_j < \infty$.

For any fixed n, the distribution of the $\underset{\sim}{i}_n$'s on the line may be described by functionals of the form $U_{n,n} = \sum_{\underset{\sim}{i}_n} f(X_{\underset{\sim}{i}_n})$, where f may depend on n. More generally, define for k = 0,1,...,n

$$U_{k,n} = \sum_{\underset{\sim}{i}_k} Z_n(\underset{\sim}{i}_k) f_{n-k}(X_{\underset{\sim}{i}_n})$$

It will be convenient to center f, i.e. to assume $Ef(X_{\underset{\sim}{i}_n}) = f_n(0) = 0$. We obtain by conditioning the following basic expression for the mean square of $U_{n,n}$:

$$EU^2_{n,n} = E[E(U^2_{n,n} \mid \acute{F}_{n-1})] = E[\mathrm{Var}(U_{n,n} \mid \acute{F}_{n-1})] + EU^2_{n-1,n} = \ldots = \tag{2.1}$$

$$\sum_{k=1}^{n} E[\mathrm{Var}(U_{k,n} \mid \acute{F}_{k-1})] =$$

$$\sum_{k=1}^{n} E\Big[\sum_{\underset{\sim}{i}_{k-1}} Z_n^2(\underset{\sim}{i}_{k-1})\{\int f^2_{n-k}(X_{\underset{\sim}{i}_{k-1}}+y)\,dG(y) - f^2_{n-k+1}(X_{\underset{\sim}{i}_{k-1}})\}\Big] =$$

$$\sum_{k=1}^{n} m^{k-1} EZ^2_{n-k} B_k(n) \leq \text{constant} \cdot m^{2n} \sum_{k=1}^{n} m^{-k} B_k(n)$$

where $B_k(n) = \int f^2_{n-k}\,dG_k - \int f^2_{n-k+1}\,dG_{k-1}$. Thus in the various examples and generalizations to be investigated in the rest of the paper we are left with a closer examination of the $B_k(n)$'s.

3. CENTRAL LIMIT ANALOGUES

Let $\mu = \int_{-\infty}^{\infty} x dG(x)$, $\sigma^2 = \int_{-\infty}^{\infty} x^2 dG(x) - \mu^2$, $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$, $\Phi(x) = \int_{-\infty}^{x}\phi(t)dt$. If $\sigma^2 < \infty$ and $t_n \simeq n\mu_1 + \sigma\sqrt{n}\,t$, then $G_n(t_n) \simeq \Phi(t)$ and one would expect $Z_n]-\infty,t_n]/m^n$ to be close to $\Phi(t)Z_n/m^n$ and thus to $\Phi(t)W$, where as usual $W = \lim_n Z_n/m^n$. In fact, this has long been known. We state the result for the sake of completeness and give the proof in order to demonstrate our method:

<u>Theorem 1.</u> <u>If $\sigma^2 < \infty$ and</u> $t_n = n\mu_1 + \sigma\sqrt{n}\,t + O(n^{-1/2})$, <u>then</u>

$$\lim_{n\to\infty} E((Z_n]-\infty,t_n]/m^n - \Phi(t)W)^2) = 0$$

<u>Proof</u>: For any fixed n let $f(x) = 1_{\{x<t_n\}} - G_n(t_n)$ and note that $Z_n]-\infty,t_n] = U_{n,n} + G_n(t_n)Z_n$. The conditions of the theorem ensure that $G_n(t_n) \to \Phi(t)$ (more generally, $G_{n-k}(t_n+x) \to \Phi(t)$ for any fixed k and x). Thus it suffices to show, that $E([U_{n,n}/m^n]^2) \to 0$. Obviously

$$0 \le B_k(n) \le \int f_{n-k}^2 dG_k = \int (G_{n-k}(t_n-x) - G_n(t_n))^2 dG_k(x) \le 1.$$

From this relation it is clear, that for each k, $\lim_{n\to\infty} B_k(n) = 0$ and the assertion follows from (2.1) and the dominated convergence theorem.

A slight modification of the argument yields a local limit theorem.

<u>Theorem 2.</u> <u>Suppose in addition to the assumptions of Theorem 1, that G is non-lattice and that</u> $\int_{\infty}^{\infty}|x|^3 dG(x) < \infty$. <u>Then for</u> $a < b$, $\lim_{n\to\infty} E([\sigma\sqrt{n}\, Z_n[t_n+a,t_n+b]/m^n - (b-a)\phi(t)W]^2) = 0$.

<u>Proof</u>. Let $f(x) = 1_{\{t_n+a \le x \le t_n+b\}} - (G_n(t_n+b)-G_n(t_n+a))$ Then $Z_n[t_n+a,t_n+b] = U_{n,n} + (G_n(t_n+b) - G_n(t_n+a))Z_n$ and so we are to prove, that $E([\sqrt{n}\, U_{n,n}/m^n]^2) \to 0$ and that

$\sigma\sqrt{n}(G_n(t_n+b) - G_n(t_n+a)) \to (b-a)\phi(t)$. The latter assertion follows from an extension of the Berry-Esseen theorem ([8], pg. 210). The same result yields the estimates $nB_k(n) \to 0$ for each $k \geq 1$ and $\sup_{1\leq k\leq n} (n-k)B_k(n) < C$ for some constant C. This follows since

$$B_k(n) \leq \int f_{n-k}^2 dG_k =$$

$$\int (G_{n-k}(t_n+b-x) - G_{n-k}(t_n+a-x) - (G_n(t_n+b) - G_n(t_n+a)))^2 dG_k(x)$$

Thus, for any n_0

$$\overline{\lim_{n\to\infty}} E([\sqrt{n}\ U_{n,n}/m_n])^2 \leq \overline{\lim_{n\to\infty}}\ n \sum_{k=n_0}^{n} m^{-k} B_k(n) \leq$$

$$\overline{\lim_{n\to\infty}} \sum_{k=n_0}^{n/2} m^{-k} \cdot 2C + \overline{\lim_{n\to\infty}}\ n \sum_{k=n/2}^{n} m^{-k} = 2C \sum_{k=n_0}^{\infty} m^{-k}$$

and since n_0 is arbitrary, the proof is complete.

Several generalizations of these results are possible. We well indicate two.

1) Suppose that the displacement of the n^{th} generation $Y_{\underset{\sim}{i}_n}$ has distribution F_n depending on n. The only thing that one needs to carry out the previous arguments is a Berry-Esseen type theorem for non-identically distributed independent random variables. Such results exist under suitable assumptions ([6], pg.78 and 81).

We introduce some notation and state the result. Let

$$\mu_i = \int_{-\infty}^{\infty} x\,dF_i(x),\quad \sigma_i^2 = \int_{\infty}^{\infty} x^2 dF_i(x) - \mu_i^2$$

$$m_k = \sum_{i=1}^{k} \mu_i \qquad S_k^2 = \sum_{i=1}^{k} \sigma_i^2$$

$$t_k = m_k + S_k t + O(n^{-1/2})$$

<u>Theorem 3</u>. Assume there exist constants $0 < A < B < \infty$ such that

$$A < \inf_i \sigma_i^2 < \sup_i \int_{-\infty}^{\infty} |x|^3 dF_i(x) < B$$

Then

$$\lim_{k\to\infty} E([m^{-n}Z_n]-\infty,t_n] - \Phi(t)W]^2) = 0$$

Assume in addition:

$$\sup_i \int_{-\infty}^{\infty} x^4 dF_i(x) < \infty, \qquad F_n \underset{w}{\to} F,$$

$\lim\sup_{|\omega|\to\infty} |f(\omega)| < 1$, and for some $L > 0$,

$$\lim_{k\to\infty}\{\sup_{|\omega|>L} |f_k(\omega) - f(\omega)|\} = 0$$

where

$$f_k(\omega) = \int e^{i\omega x} dF_k(x), \quad f(\omega) = \int e^{i\omega x} dF(x)$$

Then for $a < b$

$$\lim_{n\to\infty} E([S_n m^{-n} Z_n[t_n+a, t_n+b] - (b-a)\phi(t)W])^2 = 0$$

A result similar to Theorem 3 was proven in [7] under more restrictive assumptions.

It would be reasonable to believe that more general motions of the particles could be allowed providing analogs of the Berry-Esséen Theorem exist.

2) The underlying branching mechanism could also be generalized. For example we could allow the process to grow as a supercritical branching process with random environments [1]. The statement of the result is analogous and hence omitted.

<u>Remarks</u>. The technique described here can also be applied to Bellman-Harris processes where particle motion is permitted [14]. The details are as expected more complicated.

4. THE CASE $G(0-) = 0$, $G(0) > 0$

If we assume G to be concentrated on $[0,\infty[$, no particle on $]a,\infty[$ $(a > 0)$ can produce offspring in $[0,a]$ and thus the particles in $[0,a]$ form a continuous-type Galton-Watson process with $[0,a]$ as the set of types. A closer study of this process might be of some interest, especially because, as we shall see, it is not positively regular, but rather analogous to the decomposable processes considered in [10], for which no generalization of the limit results to the case of an infinite number of types has so far been obtained. In most cases, however, extinction will occur a.s.. For example if $G(0) = 0$, a standard assumption in the theory of age-dependent branching processes, extinction follows since $\Sigma EZ_n[0,a] = \Sigma m^n G_n(a) < \infty$ ([3], pg. 144). Also for $G(0) > 0$ but $mG(0) < 1$ we have extinction a.s. by the same reason. However, if $mG(0) > 1$, then $EZ_n[0,a] \to \infty$ and we shall give a limit result for this case.

Let $Z_n(0)$ be the number of particles of the n^{th} generation located at 0 and note that the sequence $Z_0(0)$, $Z_1(0),\ldots$ is an ordinary supercritical Galton-Watson process with mean $mG(0)$. Consequently $W = \lim_n Z_n(0)/m^n G_n(0)$ exists a.s. and in mean square.

<u>Theorem 4. If $mG(0) > 1$, then</u>

$$\lim_n E([Z_n[0,a]/m^n G_n(a) - W]^2) = 0$$

To motivate the result, consider the simplest case, where G is lattice, e.g. a distribution $g_0, g_1, g_2, \ldots, g_a$ on $\{0,1,2,\ldots,a\}$ and a is an integer. Then the particles in $\{0,1,2,\ldots,a\}$ form an $(a+1)$-type Galton-Watson process with mean matrix

$$m \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_a \\ 0 & g_0 & g_1 & \cdots & g_{a-1} \\ 0 & 0 & g_0 & \cdots & g_{a-2} \\ \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & \cdots & g_0 \end{pmatrix}$$

Limit theorems for such a situation can be found in [10]. It turns out, that W governs the behaviour of the process, so that the asserted convergence in Theorem 4 in the lattice case may be deduced from the results of [10].

The proof of Theorem 4 follows the lines of Section 3 by introducing a suitable f and estimating the $B_k(n)$'s. While in Section 3 central limit theorems were applicable, we need here a lemma on the behaviour of the $G_n(a)$'s, which is slightly stronger than the estimate $G_n(a) = O(\delta^n)$ for all $\delta > G(0)$ ([3], pg. 144) useful in standard renewal theory.

The assumption $G(0) > 0$ may here be relaxed to $G(\varepsilon) > G(0)$ for all $\varepsilon > 0$ so that $G_n(a) > 0$ for all n.

<u>Lemma 1</u>. $G_{n+1}(a)/G_n(a) \to G(0)$

<u>Proof</u>. For any $\varepsilon > 0$, write $G(x) = G(\varepsilon)F(x) + (1-G(\varepsilon))H(x)$, where F is concentrated on $[0,\varepsilon]$ and H on $]\varepsilon,\infty[$. Then $H_k(a) = 0$ for $k > N = a/\varepsilon$ and expanding by the binomial formula we get

$$G_{n+1}(a) = G(\varepsilon)^{n+1} \sum_{k=0}^{N} \binom{n+1}{k} \alpha^k F_{n+1-k} * H_k(a),$$

where $\alpha = (1-G(\varepsilon))/G(\varepsilon)$. Thus

$$G_{n+1}(a) \le G(\varepsilon)^{n+1} \sum_{k=0}^{N} \binom{n}{k} \frac{n+1}{n+1-k} \alpha^k F_n * H_k(a)$$

$$\le G(\varepsilon) \frac{n+1}{n+1-N} G_n(a)$$

and we need only to let first n tend to infinity and then ε to zero.

<u>Proof</u>. of Theorem 4: Let for any fixed n, $f(x) = 1_{\{x \le a\}}G_n(0) - 1_{\{x=0\}}G_n(a)$ and note that

$$Z_n[0,a]/m^nG_n(0) - Z_n(0)/m^nG_n(0) = U_{n,n}/m^nG_n(0)G_n(a) \qquad (4.1)$$

Also,

$$B_k(n) \le \int f_{n-k}^2 dG_k = \qquad (4.2)$$

$$\int_0^a [G_{n-k}(a-x)G_n(0) - 1_{\{x=0\}}G_{n-k}(0)G_n(a)]^2 dG_k(x) =$$

$$[G_{n-k}(a)G_n(0)-G_{n-k}(0)G_n(a)]^2 G_k(0)+G_n^2(0)\int_{0+}^a G_{n-k}^2(a-x)dG_k(x) =$$

$$C_k(n) + D_k(n) \quad \text{(say)}$$

We shall show the existence of constants $A < \infty$, $\rho > 1$, $0 \leq I_k(n) \leq 1$ with $\lim\limits_n I_k(n) = 0$ for all k such that

$$m^{-k}B_k(n) \leq AG_n^2(0)G_n^2(a)\rho^{-k}I_n(k) \text{ for all } n,k \tag{4.3}$$

Using (4.1), (2.1) and the dominated convergence theorem, this will complete the proof. It suffices to establish (4.3) for the $C_k(n)$'s and the $D_k(n)$'s separately. The treatment of these two cases are very similar and we consider only the latter which is more complicated. We define A,ρ such that $G_k(a)/m^kG_k^2(0) \leq A\rho^{-k}$ for all k. This is possible because of Lemma 1 and $mG(0) > 1$. If

$$I_k(n) = \int_{0+}^a \frac{G_{n-k}^2(a-x)}{G_{n-k}^2(a)} \frac{dG_k(x)}{G_k(a)},$$

then obviously $0 \leq I_k(n) \leq 1$ and

$$m^{-k}D_k(n) \leq m^{-k}G_n^2(0)G_{n-k}^2(a)G_k(a)I_k(n)$$

$$\leq AG_n^2(0)G_n^2(a)\rho^{-k}I_k(n)$$

It only remains to prove that $\lim\limits_n I_k(n) = 0$. But

$$\overline{\lim_n} \int_{0+}^a \frac{G_{n-k}^2(a-x)}{G_{n-k}^2(a)} dG_k(x) \leq \overline{\lim_n} \int_{0+}^a \frac{G_{n-k}(a-x)}{G_{n-k}(a)} dG_k(x) =$$

$$\overline{\lim_n} \, (G_n(a)/G_{n-k}(a) - G_k(0)) = 0 \text{ by Lemma 1.}$$

Also, the distribution of the particles of the n^{th} generation over $[0,a]$ may be described. We shall show, that this in an appropriate sense tends to a probability measure G^a on $[0,a]$, which has the interpretation as the limiting distribution of the position $X_{\underset{\sim}{i}_n}$ of a particle of the n^{th} generation conditioned upon $\{X_{\underset{\sim}{i}_n} \leq a\}$. If $G(\varepsilon) > G(0)$ for all $\varepsilon > 0$, we define G^a to be degenerate at a. Otherwise, let $F(x) = (G(x) - G(0))/(1-G(0))$. Since $F(\varepsilon) = 0$ for some $\varepsilon > 0$, there is a greatest integer N such that $F_N(a) > 0$ and we let $G^a(x) = F_N(x)/F_N(a)$, $x \leq a$. Then

<u>Lemma 2.</u> For $0 \leq x \leq a$, $G_n(x)/G_n(a) \to G^a(x)$

From this lemma and Theorem 4, we immediately get

<u>Corollary</u>: For $0 \leq x \leq a$, $Z_n[0,x]/Z_n[0,a] \xrightarrow{P} G^a(x)$

<u>Proof</u> of Lemma 2: We have to distinguish between the two cases in the definition of G^a. In the latter, the assertion is clear from

$$\frac{G_n(x)}{G_n(a)} = \frac{\sum_{k=1}^{N}\binom{n}{k}\alpha^k F_k(x)}{\sum_{k=1}^{N}\binom{n}{k}\alpha^k F_k(a)}, \text{ where } \alpha = (1-G(0))/G(0).$$

In the former, let $A_n = \{x \in]0,\varepsilon] \mid G_n(a-x)/G_n(a) > \delta\}$ where $0 < \varepsilon \leq a$, $\delta > 0$. Then

$$\int_{A_n} dG \leq \frac{1}{\delta}\int_{0+}^{a} G_n(a-x)/G_n(a)\,dG(x) = \frac{1}{\delta}(G_{n+1}(a)-G_n(a)G(0))/G_n(a)$$

Thus by Lemma 1, $\lim_n \int_{A_n} dG = 0$. Since $G(\varepsilon) > G(0)$, it follows that $\int_{A_n^c \cap]0,\varepsilon]} dG > 0$ for n large so that in particular, there is a $x_n \in A_n^c \cap]0,\varepsilon]$. Then

$$G_n(a-\varepsilon)/G_n(a) \leq G_n(a-x_n)/G_n(a) \leq \delta$$

Thus by letting δ tend to zero, we get $G_n(a-\varepsilon)/G_n(a) \to 0$ and since ε is arbitrary we are done.

Note that again Lemma 2 holds only for $G(\varepsilon) > 0$ for all $\varepsilon > 0$.

ACKNOWLEDGEMENTS

We are indebted to Martin Jacobsen and Peter Jagers for stimulating discussions on the subject of Lemmas 1 and 2.

BIBLIOGRAPHY

[1] Athreya, K.B. and Karlin, S. (1972): On branching processes in random environments: I. Extinction probability. Ann.Math.Stat 42, 1843-1858.

[2] Athreya, K.B. and Ney, P.E. (1971): Limit theorems for the means of branching random walks Proc.Sixth Prague Conf.Inf.Theory, 63-72.

[3] Athreya, K.B. and Ney, P.E. (1972): Branching Processes. Springer Verlag, Berlin Heidelberg New York.

[4] Bühler, W.J. (1970): The distribution of generations and other aspects of the family structure of a branching process. Proc. 6^{th} Berk.Symp.Math.Stat. Prob. 3, 463-480.

[5] Bühler, W.J. (1971): Generations and the degree of relationship in a supercritical Markov branching process. Z.Wahrscheinlichkeittheorie Verw.Geb. 8, 141-152.

[6] Cramer, H. (1937): Random Variables and Probability Distributions. 3^{rd} Edition Cambridge University Press, Cambridge, England.

[7] Fildes, R. (1974): An age-dependent branching process with variable lifetime distribution: the generation size. Adv.Appl.Prob. 291-308.

[8] Gnedenko, B.V. and Kolmogorov, A.N. (1954): Limit Distributions for Sums of Independent Random Variables. Addison-Wesley. Reading Mass.

[9] Harris, T.E. (1963): The Theory of Branching Processes. Springer Verlag, Berlin.

[10] Kesten, H. and Stigum, B.P. (1967): Limit theorems for decomposable multi-dimensional Galton-Watson processes. J.Math.An.Appl. 17, 309-338.

[11] Kharlamov, B.P. (1969): On the generation number of particles in a branching process Th.Prob.Appl. 14, 44-50.

[12] Martin-Löf, A. (1966): A limit theorem for the size of the n^{th} generation of an age-dependent branching process. J.Math.An.Appl. 12, 273-279.

[13] Ney, P.E. (1965): The limit distribution of a binary cascade process. J.Math.An.Appl. 10, 30-36.

[14] Ney, P.E. (1965): The convergence of a random distribution function associated with a branching process. J.Math.An.Appl. 12, 316-327.

[15] Samuels, M.L. (1971): Distribution of the branching process population among generations. J.Appl. Prob. 8, 655-667.

[16] Stam, A.J. (1966): On a conjecture of Harris. Z. Wahrscheinlichkeitstheorie verw. Geb. 5, 202-206.