Norman Kaplan

A Note on the Distribution of Generations in a Branching Process

Preprint
October
1974

13

Institute of Mathematical Statistics
University of Copenhagen

Norman Kaplan*

A NOTE ON THE DISTRIBUTION OF GENERATIONS IN A BRANCHING PROCESS

Preprint 1974 No. 13

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

October 1974

*Department of Statistics, University of California, Berkeley.

ABSTRACT

Some limit theorems relating to the number of partials in the n^{th} generation alive at time t are reexamined. A single method of proof is given and shown to work equally as well for certain generalizations.

Key words: Age-dependent branching process, generation number.

Written while on leave at the Institute of Mathematical Statistics, University of Copenhagen. Research partially supported by N.S.F. Grant, 31091 X.

SECTION 1

1. Introduction

Let $\{X(t)\}_{t \geq 0}$ be a supercritical Bellman-Harris branching process with age distribution F, and offspring p.g.f. f(s). Some attention has been given to the study of the limit behavior of the following random quantities:

$U_k(t)$ = number of particles of the k^{th} generation born before t

and

$V_k(t)$ = number of particles of the k^{th} generation alive at time t.

Results can be found in [1], [2], [4], [7], [8], [9]. In each of these papers the technique of proof is highly analytic. It turns out that some of the existing theorems can be gotten very quickly as a consequence of the structure of the process and the Berry-Esséen Theorem. This technique we feel is the natural one since it is simple and applies as well to a generalization of the process where the lifetime distribution of each generation is varible. This case was recently studied by Fildes [4].

We now introduce the necessary notation and state our result for the classical case. Put:

$$\mu_i = \int_0^\infty v^i dF(v), \quad m_i = \frac{\partial^i f(1-)}{\partial^i s} \qquad i = 1,2,\ldots$$

$$\sigma^2 = \mu_2 - (\mu_1)^2, \quad m = m_1$$

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \text{ and } \Phi(x) = \int_{-\infty}^{x} \phi(v)dv, -\infty < x < \infty$$

If X_k = number of particles in the k^{th} generation, then it is well known [6] that $\lim_{k\to\infty} m^{-k}X_k = W$ w.p.1 and if $m_2 < \infty$, the limit exists in mean square.

Theorem. Assume $\mu_i < \infty$, $m_i < \infty$ $i = 1,2$, $m_1 > 1$ and $m_3 < \infty$. Let $t_k = k\mu_1 + \sigma\sqrt{k}t$. Then,

$$\lim_{k\to\infty} E\left([m^{-k}U_k(t_k) - \Phi(t)W]^2\right) = 0 \tag{1.1}$$

Suppose in addition F is non lattice. Then,

$$\lim_{k\to\infty} E\left([\sqrt{k}\sigma m^{-k}V(t_k) - m\phi(t)W]^2\right) = 0 \tag{1.2}$$

SECTION 2

Proof of the Theorem

We first prove (1.1). Since $m^{-k}X_k \overset{L_2}{\to} W$, it suffices to show,

$$\lim_{k\to\infty} E\left([m^{-k}U(t_k) - \Phi(t)X_k m^{-k}]^2\right) = 0$$

Following the notation of [6, Chapter 5], it is not difficult to check that,

$$U_k(t_k) = \sum_{i_1=1}^{n_o} \sum_{i_2=1}^{n_{i_1}} \cdots \sum_{i_{k-1}=1}^{n_{i_1 i_2 \cdots i_{k-2}}} n_{i_1 \cdots i_{k-1}}$$

$$\times\left(I_{\{T_o + T_{i_1} + T_{i_1 i_2} + \cdots + T_{i_1 i_2 \cdots i_{k-1}} < t\}}\right)$$

where

$T_{i_1 i_2 \cdots i_{k-1}}$ is the length of life of $\langle \underset{\sim}{i}_k \rangle = \langle i_1 i_2 \cdots i_{k-1} \rangle$

$n_{i_1 i_2 \cdots i_{k-1}}$ is the number of offspring produced by $\langle \underset{\sim}{i}_k \rangle$

I_A is the indicator function of the set A.

We introduce some notation. Put for $k \geq 1$,

$$n_{\underset{\sim}{i}_k} = n_{i_1 i_2 \ldots i_{k-1}}, \quad T_{\underset{\sim}{i}_k} = T_{i_1 i_2 \ldots i_{k-1}}$$

$$\sum_{\underset{\sim}{i}_k} = \sum_{i_1=1}^{n_o} \sum_{i_2=1}^{n_{i_1}} \ldots \sum_{i_{k-1}=1}^{n_{i_1 i_2 \ldots i_{k-2}}}$$

and

$$S_{\underset{\sim}{i}_k} = T_0 + T_{i_1} + T_{i_1 i_2} + \ldots + T_{i_1 i_2 \ldots i_{k-1}}$$

Then,

$$U_k(t_k) - \Phi(t)X_k = \sum_{\underset{\sim}{i}_k} \left(I_{\{S_{\underset{\sim}{i}_k} < t_k\}} - \Phi(t) \right) n_{\underset{\sim}{i}_k}$$

For any collection of random variables G, let $\sigma(G)$ be the σ-field generated by G. Define for $k \geq 1$,

$$G_k = \{(n_{\underset{\sim}{i}_j}, T_{\underset{\sim}{i}_j}) \text{ for all } \langle \underset{\sim}{i}_j \rangle \text{ such that } j \leq k\}.$$

Then,

$$E\left([U_k(t_k) - \Phi(t)X_k]^2 \right) = E\left(E\left([U_k(t_k) - \Phi(t)X_k]^2 \,\middle|\, G_{k-1} \right) \right)$$

$$= E\left(\mathrm{Var}\left(U_k(t_k) \,\middle|\, G_{k-1} \right) \right) + E\left([E(U_k(t_k) \,|\, G_{k-1}) - \Phi(t)X_k]^2 \right)$$

Conditioned on G_{k-1}, $U_k(t_k)$ is just the sum of independent random variables with the typical term having variance equal to

$$n^2_{\underset{\sim}{i}_k} [F(t_k - S_{\underset{\sim}{i}_{k-1}}) - F^2(t_k - S_{\underset{\sim}{i}_{k-1}})] I_{\{S_{\underset{\sim}{i}_{k-1}} < t_k\}}$$

Hence,

$$E\left(\mathrm{Var}\left(U_k(t_k) \,\middle|\, G_{k-1} \right) \right) = m^{k-1} E(X_1^2) [F_{(k)}(t_k) - F^2 * F_{(k-1)}(t_k)]$$

where * represents convolution and $F_{(k)}(t) = F * F_{(k-1)}(t)$, $k \geq 1$. Also observe that

$$E\left([E\left(U_k(t_k)\middle|G_{k-1}\right) - \Phi(t)X_k]^2\right)$$

$$= E\left([\sum_{\underset{\sim}{i}_{k-1}} (\sum_{i_{k-1}=1}^{n_{\underset{\sim}{i}_{k-1}}} n_{\underset{\sim}{i}_k})\left(F(t_k - S_{\underset{\sim}{i}_{k-1}})I_{\{S_{\underset{\sim}{i}_{k-1}}<t\}} - \Phi(t)\right)]^2\right)$$

The last expression is of exactly the same form as (2.1) except that the summation is over one less index. Thus we can repeat the calculation k-1 times to obtain

$$\Theta_k = E\left([U(t_k) - \Phi(t)X_k]^2\right) = \sum_{j=1}^{k} m^{k-j}E(X_j^2)B_j(k) \qquad (2.2)$$

where

$$B_j(k) = F^2_{(j-1)} * F_{(k-j+1)}(t_k) - F^2_{(j)} * F_{(k-j)}(t_k)$$

and $\qquad 1 \le j \le k-1$

$$B_k(k) = (F_{(k-1)} - \Phi(t))^2 * F(t_k)$$

It is easy to show that $E(X^2_j) = O(m^{2j})$. Therefore, there exists some constant A independent of k such that

$$m^{-2k}\Theta_k \le A \sum_{j=1}^{k} m^{-(k-j)} B_j(k)$$

Let $\varepsilon > 0$. Choose N_o such that $\sum_{j \ge N_o} m^{-j} < \varepsilon$

Then for $k > N_o$,

$$m^{-2k}\Theta_k \le 2\varepsilon A + \sum_{j=k-N_o+1}^{k} m^{(k-j)}B_j(k)$$

We now apply the Berry Esséen Theorem [5, pg. 201] to conclude that

Esseen

$$\lim_{k\to\infty} B_j(k) = 0, \quad k - N_o + 1 \le j \le k$$

This proves (1.1).

To prove (1.2) we first observe that

$$\sqrt{k}\sigma V_k(t_k) - m\phi(t)X_k = \sum_{\underset{\sim}{i}_{k+1}} \left(\sqrt{k}\sigma I_{\{S_{\underset{\sim}{i}_k} < t_k < S_{\underset{\sim}{i}_{k+1}}\}} - m\phi(t)\right)$$

Thus if we condition exactly as before we obtain,

$$\Lambda_k = E\left([\sqrt{k}\sigma V_k(t_k) - m\phi(t)X_k]^2\right)$$

$$= \sum_{j=0}^{k} m^{k-j} E(X_j^2) C_j(k) \tag{2.3}$$

where

$$C_j(k) = k\sigma^2(F_{(j-1)} - F_{(j)})^2 * F_{(k-j+1)}(t_k)$$

$$- k\sigma^2(F_{(j)} - F_{(j+1)})^2 * F_{(k-j)}(t_k), 1 \le j \le k-1$$

$$C_o(k) = k\sigma^2(1-F) * F_{(k)}(t_k) - k\sigma^2(1-F)^2 * F_{(k)}(t_k)$$

and

$$C_k(k) = \left(\sqrt{k}\sigma(F_{(k-1)} - F_{(k)}) - m\phi(t)\right)^2 * F(t_k)$$

Let $N_o > 0$. Then using the extended Berry Esséen Theorem [5, pg 210] it is not difficult to show that

$$\lim_{k\to\infty} C_j(k) = 0, k - N_o \le j \le k$$

and

$$\sup_{1 \le j \le k-N_o} |C_j(k)| \le D_1 \frac{k}{j}$$

where D_1 is a constant independent of k.

Hence,

$$\limsup_{k\to\infty}(m^{-2k}\Lambda_k) = \limsup_{k\to\infty}\left(D_2 \sum_{j=1}^{k-N_o} m^{-(k-j)} \frac{k}{j}\right)$$

where D_2 is a suitable constant. Let $\varepsilon > 0$ such that

$m^{-1}(1+\varepsilon) = r < 1$. Then for $k > [\varepsilon^{-1}]$,

$$\sum_{j=N_o}^{[k-\varepsilon^{-1}]-1} m^{-j} \frac{k}{k-j} \leq m^{-N_o} \frac{k}{k-N_o} \frac{1}{1-r}$$

and

$$\sum_{j=[k-\varepsilon^{-1}]-1}^{k-1} m^{-j} \frac{k}{k-j} \leq D_3\, k\, m^{-[k-\varepsilon^{-1}]}$$

for D_3 a suitable constant. Since N_o is arbitrary we are done.

SECTION 3

A Generalization.

Suppose a particle in generation n lives a random length of time governed by F_n which depends on n. One can easily check that expressions similar to (2.2) and (2.3) hold in this more general setup. In fact, the only thing that one needs to carry through the arguments of the previous section is a Berry-Esséen type theorem for non-identically distributed random variables. Fortunately such results exist under suitable assumptions. [3, pgg 78 and 84].

We now introduce some notation and state the result. Let

$$\mu_i = \int v\, dF_i(v) \qquad \sigma_i^2 = \int v^2\, dF_i(v) - \mu_i^2.$$

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

$$t_k = m_k + S_k t$$

Theorem. Assume there exist constants $0 < A < B < \infty$ such that

$$A < \inf_i \sigma_i^2 \text{ and } \sup_i \int v^3\, dF_i(v) < B.$$

Then

$$\lim_{k\to\infty} E\left([\frac{U_k(t_k)}{m_k} - W\Phi(t)]^2 \right) = 0$$

Assume in addition $\sup_i \int v^4 \, dF_i(v) < \infty, F_n \underset{w}{\rightarrow} F, \limsup_{|\omega| \to \infty} |f(\omega)| < 1$

and for some L,

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

where

$$f_k(\omega) = \int e^{iv\omega} \, dF_k(v), \quad f(\omega) = \int e^{iv\omega} \, dF(v).$$

Then,

$$\lim_{k \to \infty} E\left(\left[\frac{S_k V_k(t_k)}{m^k} - \mu W \, \phi(t) \right]^2 \right) = 0$$

where

$$\mu = \int v \, dF(v).$$

Proof. The conditions of the theorem are sufficient to imply the required versions of the Berry Esséen Theorem given in [3]. Q.E.D.

Remark. The assumptions of Fildes imply ours.

Conjecture. I believe that this method can be applied to the generalized age-dependent model to obtain analogous results.

BIBLIOGRAPHY

[1] Bühler, W.J. (1970). The distribution of generations and other aspects of the family structure of a branching process. Proc. of 6th Berkely Sy. Math.Stat.Prob. 3, 463-480.

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

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

[4] Fildes, R. (1974). An age-dependent branching process with varible lifetime distribution: the generation size. Adv. in App. Prob. 291-308.

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

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

[7] Kharlamov, B.P. (1969). On the generation numbers of particles in a branching process with overlapping generations. Theor.Prob.Appl. 14, 44-50.

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

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