SOME CENTRAL LIMIT ANALOGUES FOR SUPERCRITICAL GALTON-WATSON PROCESSES

BY

C. C. HEYDE

Reprinted from the

JOURNAL OF APPLIED PROBABILITY

Vol. 8, No. 1, pp. 52–59

March 1971

SOME CENTRAL LIMIT ANALOGUES FOR SUPER-CRITICAL GALTON-WATSON PROCESSES

C. C. HEYDE, *Australian National University*

1. Introduction

It is possible to interpret the classical central limit theorem for sums of independent random variables as a convergence rate result for the law of large numbers. For example, if X_i , $i = 1, 2, 3, \dots$ are independent and identically distributed random variables with $EX_i = \mu$, $\text{var } X_i = \sigma^2 < \infty$ and $S_n = \sum_{i=1}^n X_i$, then the central limit theorem can be written in the form

$$\lim_{n \rightarrow \infty} P(\sigma^{-1}n^{\frac{1}{2}}(n^{-1}S_n - \mu) \leq x) = \Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^x e^{-\frac{1}{2}u^2} du.$$

This provides information on the rate of convergence in the strong law $n^{-1}S_n \xrightarrow{\text{a.s.}} \mu$ as $n \rightarrow \infty$. (“a.s.” denotes almost sure convergence.) It is our object in this paper to discuss analogues for the super-critical Galton-Watson process.

Let $Z_0 = 1, Z_1, Z_2, \dots$ denote a super-critical Galton-Watson process with a non-degenerate offspring distribution having probability generating function $F(s) = \sum_{j=0}^{\infty} s^j P(Z_1 = j)$, $0 \leq s \leq 1$, where $1 < m = EZ_1 < \infty$. The process evolves in such a way that the generating function $F_n(s)$ of Z_n is the n th functional iterate of $F(s)$. The convergence problem for Z_n , when appropriately normed, has been studied by quite a number of authors; for an ultimate form see Heyde [2]. However, we shall herein suppose that $EZ_1 \log Z_1 < \infty$ in which case $W_n = m^{-n}Z_n$ converges almost surely to a non-degenerate random variable W as $n \rightarrow \infty$ (e.g., Stigum [8], Kesten and Stigum [4]). We shall be dealing with the case where Z_1 belongs to the domain of attraction of a stable law of index α , $1 < \alpha \leq 2$, so that $EZ_1^a < \infty$ for any $a < \alpha$ and in particular $EZ_1 \log Z_1 < \infty$.

Now the result $m^{-n}Z_n \xrightarrow{\text{a.s.}} W$ is an analogue of the strong law of large numbers. Furthermore, the existence of central limit analogues in this context has been shown by Heyde [3] where the following theorem is established.

Theorem A. Let $\text{var } Z_1 = \sigma^2 < \infty$. As $n \rightarrow \infty$, $m^{\frac{1}{2}n}(W_n - W)$ converges in distribution. The limit law is given by the characteristic function relation

$$\lim_{n \rightarrow \infty} E[\exp\{itm^{\frac{1}{2}n}(W - W_n)\}] = E[\exp\{-\frac{1}{2}t^2\sigma^2(m^2 - m)^{-1}W\}].$$

We shall extend this result in Theorem 2 and then go on to obtain a different central limit analogue by modifying the norming (Theorem 3).

Received in revised form 8 June 1970.

2. Results

As a preliminary to setting up the central limit analogues we shall firstly obtain the following result of independent interest.

Theorem 1. W belongs to the domain of attraction of a stable law of index α , $1 < \alpha \leq 2$, if and only if Z_1 belongs to the same domain of attraction. More specifically, if U_i , $i = 1, 2, 3, \dots$ are independent and identically distributed, each with the distribution of $W - 1$ and if $Z^{(i)}$, $i = 1, 2, 3, \dots$ are independent and identically distributed, each with the distribution of Z_1 , then $c_n^{-1} \sum_{i=1}^n (Z^{(i)} - m)$ converges in distribution to a stable law Y if and only if $(m^\alpha - m)^{1/\alpha} c_n^{-1} \sum_{i=1}^n U_i$ converges in distribution to Y . Here $\{c_n\}$ is a monotone sequence of positive constants.

Next, making use of Theorem 1, we shall obtain the following extension of Theorem A.

Theorem 2. Let Z_1 belong to the domain of attraction of a stable law of index α , $1 < \alpha \leq 2$. More specifically, suppose that if $Z^{(i)}$, $i = 1, 2, 3, \dots$ are independent and identically distributed, each with the distribution of Z_1 , then there is a monotone sequence $\{c_n\}$ of constants such that $c_n^{-1} \sum_{i=1}^n (Z^{(i)} - m)$ converges in distribution to a stable law with characteristic function $f(t)$. Then, we have

$$\lim_{n \rightarrow \infty} E[\exp\{it(m^\alpha - m)^{1/\alpha} m^n c_n^{-1} (W - W_n)\}] = E[f(t)]^W.$$

We remark that since Z_1 is non-negative, $f(t)$ must be of the form $\exp\{-c|t|^\alpha(1 - it \tan \frac{1}{2}\pi\alpha \operatorname{sgn} t)\}$ where c is a positive constant (Gnedenko and Kolmogorov [1], Theorem 2, page 175).

Now the dependence of the limit result of Theorem 2 on the distribution of W is unfortunate as comparatively little can be said about W in general. The distribution function corresponding to the characteristic function $E[f(t)]^W$ can rarely be written down in a form which is suitably explicit. It would be much more pleasing to have a limit which did not involve W .

The clue to the problem of obtaining a limit independent of W comes from noting that $E[f(t)]^W$ is the characteristic function of a product $VW^{1/\alpha}$ where V and W are independent, V is stable with characteristic function $f(t)$ and α is the index of the stable law. This indicates an appropriate change to make in the norming which is done in the following theorem.

Theorem 3. Under the same conditions on the distribution of Z_1 as in Theorem 2,

$$\lim_{n \rightarrow \infty} E[\exp\{it(m^\alpha - m)^{-1/\alpha} m^n c_n^{-1} (W - W_n)\} | Z_n > 0] = f(t).$$

The relation between Theorems 2 and 3 is further elucidated in Section 6.

3. Proof of Theorem 1

The finite variance case of this result is well-known. In fact $\text{var } W = (m^2 - m)^{-1} \text{var } Z_1$. Furthermore, Lamperti [5] has shown that W belongs to the domain of normal attraction of a stable law of index α , $1 < \alpha \leq 2$, if and only if Z_1 belongs to the same domain of normal attraction. The result of Theorem 1 is obtained by a straightforward extension of Lamperti's methods.

Let $\phi(\lambda)$ be the Laplace transform $E[\exp\{-\lambda W\}]$ and recall that F is the probability generating function of Z_1 . The connection between F and ϕ is well-known to be

$$(1) \quad \phi(m\lambda) = F(\phi(\lambda)),$$

which, together with the requirement that they be a probability generating function and a Laplace transform respectively, determines either uniquely in terms of the other (e.g., Levinson [6]). Now W , which has mean one, belongs to the domain of attraction of a stable law of index α , $1 < \alpha \leq 2$, if and only if

$$(2) \quad \phi(\lambda) = 1 - \lambda + \lambda^\alpha L(\lambda^{-1})(1 + o(1))$$

as $\lambda \rightarrow 0$ where L is non-negative and slowly varying at infinity. Also, Z_1 belongs to the domain of attraction of a stable law of index α , $1 < \alpha \leq 2$, if and only if

$$(3) \quad 1 - F(x) = m(1-x) - (1-x)^\alpha M((1-x)^{-1})(1 + o(1))$$

as $x \rightarrow 1$ where M is non-negative and slowly varying at infinity. Consequently, we must derive the form (3) from (1) and (2) and conversely the form (2) from (1) and (3).

Firstly, assume (2). Then, using (1) and writing $\phi(\lambda) = x$, we have as $\lambda \rightarrow 0$,

$$(4) \quad \begin{aligned} 1 - F(x) &= m\lambda - m^\alpha \lambda^\alpha L((m\lambda)^{-1})(1 + o(1)) \\ &= m\lambda - m^\alpha \lambda^\alpha L(\lambda^{-1})(1 + o(1)), \end{aligned}$$

since L is slowly varying. But

$$(5) \quad x = 1 - \lambda + \lambda^\alpha L(\lambda^{-1})(1 + o(1)),$$

so that, using (5) in (4),

$$(6) \quad 1 - F(x) = m(1-x) - (m^\alpha - m)\lambda^\alpha L(\lambda^{-1})(1 + o(1)).$$

However, from (5),

$$\lambda = (1-x)(1 + o(1))$$

as $x \rightarrow 1$, $\lambda \rightarrow 0$ and using this result in (4),

$$1 - F(x) = m(1-x) - (m^\alpha - m)(1-x)^\alpha L((1-x)^{-1})(1 + o(1)),$$

which is of the form (3).

Next, let

$$(7) \quad 1 = F(x) - m(1-x) - (1-x)^\alpha M((1-x)^{-1})(1 + A(1-x)),$$

where $A(1-x) = o(1)$ as $x \rightarrow 1$ and write

$$(8) \quad \phi(\lambda) = 1 - \lambda + a(\lambda)$$

where $a(\lambda) = o(\lambda)$ as $\lambda \rightarrow 0$ since W has mean one. We then obtain from (1) and (7) that

$$(9) \quad a(m\lambda) = ma(\lambda) + \lambda^\alpha M(\lambda^{-1})(1 + B(\lambda))$$

for a certain $B(\lambda)$ which is a function of $a(\lambda)$ and satisfies $B(\lambda) = o(1)$ as $\lambda \rightarrow 0$.

Now consider the more general functional equation

$$(10) \quad b(m\lambda) = mb(\lambda) + \lambda^\alpha M(\lambda^{-1})(1 + B(\lambda))$$

with $B(\lambda)$ as in (9). The homogeneous part of (10) has the general solution $\lambda P(\log \lambda)$ where P is any function with period $\log m$. Thus, (10) has at most one and hence exactly one solution which is $o(\lambda)$ as $\lambda \rightarrow 0$, for if there were two, their difference would be of the form $\lambda P(\log \lambda)$ which is not $o(\lambda)$ unless $P(\log \lambda) \equiv 0$. This solution is easily seen to be of the form

$$(11) \quad b(\lambda) = (m^\alpha - m)^{-1} \lambda^\alpha M(\lambda^{-1})(1 + o(1))$$

and this is the form of $a(\lambda)$ since $a(\lambda)$ is uniquely specified. The required form (2) then follows from (8) and (11).

Now we note from the above analysis that the slowly varying functions L and M of (2) and (3) are related by

$$(12) \quad L(\lambda^{-1}) = (m^\alpha - m)^{-1} M(\lambda^{-1})(1 + o(1))$$

as $\lambda \rightarrow 0$. We shall exclude from consideration the case $\alpha = 2$ and $L(\lambda^{-1}), M(\lambda^{-1})$ asymptotically constant for then the variances of W and Z_1 are finite and are related by $\text{var } W = (m^2 - m)^{-1} \text{var } Z_1$ and the result of the theorem is obvious. We also note that when $\alpha = 2$ and the variances of W and Z_1 are not finite, $L(\lambda^{-1})$ and $M(\lambda^{-1})$ must be unbounded above as $\lambda \rightarrow 0$. Then, since

$$E[\exp\{-\lambda Z_1\}] = F(e^{-\lambda}),$$

we have from (3) that

$$\begin{aligned} E[\exp\{-\lambda c_n^{-1} \sum_{i=1}^n (Z^{(i)} - m)\}] &= [F(\exp\{-c_n^{-1} \lambda\}) \exp\{m c_n^{-1} \lambda\}]^n \\ &\rightarrow \exp\{c \lambda^\alpha\}, c > 0, \end{aligned}$$

as $n \rightarrow \infty$ if and only if

$$(13) \quad n c_n^{-\alpha} M(c_n) \rightarrow c$$

as $n \rightarrow \infty$. Using (2) and (12), we see that the condition (13) is also necessary and sufficient for

$$\begin{aligned} E[\exp\{-\lambda(m^\alpha - m)^{1/\alpha} c_n^{-1} \sum_{i=1}^n U_i\}] &= [\phi(\lambda c_n^{-1}(m^\alpha - m)^{1/\alpha}) \exp\{\lambda c_n^{-1}(m^\alpha - m)^{1/\alpha}\}]^n \\ &\rightarrow \exp\{c\lambda^\alpha\} \end{aligned}$$

as $n \rightarrow \infty$. This completes the proof of the theorem.

4. Proof of Theorem 2

This follows the pattern of that of the theorem of Heyde [3] (Theorem A). Write $b_n = c_n(m^\alpha - m)^{-1/\alpha}$. Then, exactly as in [3] we find that

$$\begin{aligned} (14) \quad E[\exp\{itm^n b_{m^n}^{-1}(W - W_n)\}] &= E[E[\exp\{itb_{m^n}^{-1}(W - 1)\}]]^{Z_n} \\ &= \int_0^\infty (E[\exp\{itb_{m^n}^{-1}(W - 1)\}])^{xm^n} dP(m^{-n}Z_n \leq x) \\ &= 1 + m^n \log E[\exp\{itb_{m^n}^{-1}(W - 1)\}] \\ &\quad \times \int_0^\infty P(m^{-n}Z_n > x) (E[\exp\{itb_{m^n}^{-1}(W - 1)\}])^{xm^n} dx \end{aligned}$$

using integration by parts. (In the right hand side of Equation 3 of [3], the term $P(Z_n = 0)$ should be replaced by 1.) Now, we have

$$E[\exp\{itc_n^{-1} \sum_{i=1}^n (Z^{(i)} - m)\}] = (E[\exp\{itc_n^{-1}(Z_1 - m)\}])^n \rightarrow f(t)$$

as $n \rightarrow \infty$ and consequently, using Theorem 1,

$$(E[\exp\{itb_n^{-1}(W - 1)\}])^n \rightarrow f(t)$$

as $n \rightarrow \infty$. Thus,

$$(15) \quad (E[\exp\{itb_{m^n}^{-1}(W - 1)\}])^{m^n} \rightarrow f(t)$$

and

$$(16) \quad m^n \log E[\exp\{itb_{m^n}^{-1}(W - 1)\}] \rightarrow \log f(t)$$

as $n \rightarrow \infty$. Also,

$$|f(t)| = \exp\{-c|t|^\alpha\}$$

for some constant $c > 0$ so that for suitably large n ,

$$|E[\exp\{itb_{m^n}^{-1}(W - 1)\}]|^{m^n} < \exp\{-\frac{1}{2}c|t|^\alpha\},$$

and therefore the integrand in the rightmost term of (14) is bounded in absolute

value by $\exp\{-\frac{1}{2}c|t^\alpha|x\}$. It then follows from the dominated convergence theorem together with (15), (16) and since $P(m^{-n}Z_n > x) \rightarrow P(W > x)$ as $n \rightarrow \infty$, that

$$\begin{aligned} \lim_{n \rightarrow \infty} E[\exp\{itm^n b_{m^n}^{-1}(W - W_n)\}] &= 1 + \log f(t) \int_0^\infty P(W > x) [f(t)]^x dx \\ &= \int_0^\infty [f(t)]^x dP(W \leq x) \\ &= E[f(t)]^W \end{aligned}$$

again using integration by parts. This completes the proof.

5. Proof of Theorem 3

Take $r > n$ and again write $b_n = c_n(m^\alpha - m)^{-1/\alpha}$. We have

$$\begin{aligned} &E[\exp\{itm^n b_{Z_n}^{-1}(W_r - W_n)\} | Z_n > 0] \\ &= [P(Z_n > 0)]^{-1} \sum_{j=1}^{\infty} E[\exp\{itm^n b_{Z_n}^{-1}(m^{-r}Z_r - m^{-n}Z_n)\} | Z_n = j] P(Z_n = j) \\ (17) \quad &= [P(Z_n > 0)]^{-1} \sum_{j=1}^{\infty} \exp\{-itj b_j^{-1}\} (E[\exp\{itb_j^{-1}W_{r-n}\}])^j P(Z_n = j) \\ &\rightarrow [P(Z_n > 0)]^{-1} \sum_{j=1}^{\infty} \exp\{-itj b_j^{-1}\} (E[\exp\{itb_j^{-1}W\}])^j P(Z_n = j) \end{aligned}$$

as $r \rightarrow \infty$, the justification for taking the limit inside the summation following exactly as in [3]. Now define Z_n^* , $n \geq 1$, such that

$$P(Z_n^* = j) = P(Z_n = j | Z_n > 0) = [P(Z_n > 0)]^{-1} P(Z_n = j), \quad j = 1, 2, 3, \dots$$

Then, from (17),

$$\begin{aligned} &E[\exp\{itm^n b_{Z_n}^{-1}(W - W_n)\} | Z_n > 0] \\ &= \lim_{r \rightarrow \infty} E[\exp\{itm^n b_{Z_n}^{-1}(W_r - W_n)\} | Z_n > 0] \\ &= \sum_{j=1}^{\infty} (E[\exp\{itb_j^{-1}(W - 1)\}])^j P(Z_n^* = j) \\ (18) \quad &= E[E[\exp\{itb_{Z_n^*}^{-1}(W - 1)\}]]^{Z_n^*}. \end{aligned}$$

Now, $E[E[\exp\{itb_{Z_n^*}^{-1}(W - 1)\}]]^{Z_n^*}$ is the characteristic function of $b_{Z_n^*}^{-1}(U_1 + \dots + U_{Z_n^*})$ where the U_i , $i = 1, 2, 3, \dots$ are independent of Z_n^* and are independent and identically distributed, each with the distribution of $W - 1$. Also, as $n \rightarrow \infty$, $m^{-n}Z_n^* \xrightarrow{\text{a.s.}} W^*$ where W^* is positive and has an absolutely continuous distribution function given by

$$P(W \leq x) = P(W = 0)D(x) + P(W > 0)P(W^* \leq x)$$

where

$$D(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

(see Stigum [8]).

Then, since $b_n^{-1}(U_1 + \dots + U_n)$ converges in distribution to the stable law with characteristic function $f(t)$ (from Theorem 1) we can, given $\varepsilon > 0$, choose N so large that

$$|(E[\exp\{itb_j^{-1}U_i\}])^j - f(t)| < \varepsilon$$

for $j \geq N$ uniformly in any fixed finite t interval. Thus, for any fixed δ , $0 < \delta < 1$, we can choose n so large that $[m^n\delta] \geq N$ ($[x]$ denotes the integer part of x) and then

$$\begin{aligned} & \left| \sum_{j=1}^{\infty} [(E[\exp\{itb_j^{-1}U_i\}])^j - f(t)]P(Z_n^* = j) \right| \\ & \leq \sum_{j=1}^{[m^n\delta]-1} |(E[\exp\{itb_j^{-1}U_i\}])^j - f(t)|P(Z_n^* = j) \\ & \quad + \sum_{j=[m^n\delta]}^{\infty} |(E[\exp\{itb_j^{-1}U_i\}])^j - f(t)|P(Z_n^* = j) \\ & \leq 2P(Z_n^* < [m^n\delta]) + \varepsilon P(Z_n^* \geq [m^n\delta]) \\ & \rightarrow 2P(W^* < \delta) + \varepsilon P(W^* > \delta) \end{aligned}$$

as $n \rightarrow \infty$. Thus, since W^* is absolutely continuous and δ and ε can be arbitrarily small, we have

$$(19) \quad \lim_{n \rightarrow \infty} \sum_{j=1}^{\infty} (E[\exp\{itb_j^{-1}U_i\}])^j P(Z_n^* = j) = f(t)$$

and using (19) in (18)

$$\lim_{n \rightarrow \infty} E[\exp\{itm^n b_{Z_n}^{-1}(W - W_n)\} | Z_n > 0] = f(t)$$

as required.

6. Concluding remarks

To illuminate the connection between Theorems 2 and 3, we note from (14) that $m^n b_{m^n}^{-1}(W - W_n)$ has the same distribution as X_n which is given by

$$X_n = \begin{cases} 0, & \text{if } Z_n = 0 \\ b_{m^n}^{-1}(U_1 + \dots + U_{Z_n}), & \text{if } Z_n > 0, \end{cases}$$

where the U_i , $i = 1, 2, 3, \dots$ are independent of Z_n and are independent and identically distributed each with the distribution of $W - 1$. Take Z_n^* and W^* as in the proof of Theorem 3. Then, conditional on $Z_n > 0$, $m^n b_{m^n}^{-1}(W - W_n)$ has the distribution of $b_{m^n}^{-1}(U_1 + \dots + U_{Z_n^*})$ which converges in distribution to a product $V(W^*)^{1/\alpha}$ of independent random variables (using Theorem 2) where V has characteristic function $f(t)$. On the other hand, from the proof of Theorem 3 we see that, conditional on $Z_n > 0$, $m^n b_{Z_n}^{-1}(W - W_n)$ has the distribution of $b_{Z_n}^{-1}(U_1 + \dots + U_{Z_n^*})$ which converges in distribution to V .

Finally, it is worth remarking that a central limit analogue is not obtainable for the Galton-Watson process with $EZ_1 = \infty$. Clearly it is possible to find a monotone sequence $\{c_n\}$ of positive constants such that $c_n^{-1}Z_n \xrightarrow{\text{a.s.}} 0$ as $n \rightarrow \infty$. However, there is no norming sequence $\{d_n\}$ for which $d_n^{-1}Z_n$ converges in distribution to a non-degenerate limit law ([7], Theorem 4.4).

References

- [1] GNEDENKO, B. V. AND KOLMOGOROV, A. N. (1954) *Limit Distributions for Sums of Independent Random Variables*. Addison Wesley, Reading, Mass.
- [2] HEYDE, C. C. (1970) Extension of a result of Seneta for the super-critical Galton-Watson process. *Ann. Math. Statist.* **41**, 739–742.
- [3] HEYDE, C. C. (1970) A rate of convergence result for the super-critical Galton-Watson process. *J. Appl. Prob.* **7**, 451–454.
- [4] KESTEN, H. AND STIGUM, B. P. (1966) A limit theorem for multi-dimensional Galton-Watson processes. *Ann. Math. Statist.* **37**, 1211–1223.
- [5] LAMPERTI, J. (1967) Limiting distributions for branching processes. *Proc. 5th Berkeley Symposium on Math. Statist. and Prob.* **II**, 225–241.
- [6] LEVINSON, N. (1959) Limiting theorems for Galton-Watson branching process. *Illinois J. Math.* **3**, 554–565.
- [7] SENETA, E. (1969) Functional equations and the Galton-Watson process. *Adv. Appl. Prob.* **1**, 1–42.
- [8] STIGUM, B. P. (1966) A theorem on the Galton-Watson process. *Ann. Math. Statist.* **37**, 695–698.

Note added in proof

A central limit analogue for the finite variance case of a continuous time Markov branching process has been obtained by K. B. Athreya (1968) Some results on multitype continuous time Markov branching processes. *Ann. Math. Statist.* **39**, 347–357.