J. Appl. Prob. **42,** 581–586 (2005)
Printed in Israel

B-STABILITY

COLIN MALLOWS,* *Avaya Labs*
LARRY SHEPP,** *Rutgers University*

Abstract

A random variable Y is *branching stable* (*B-stable*) for a nonnegative integer-valued random variable J with $\mathrm{E}(J) > 1$ if $Y^{*J} \sim cY$ for some scalar c, where Y^{*J} is the sum of J independent copies of Y. We explore some aspects of this notion of stability and show that, for any Y_0 with finite nonzero mean, if we define $Y_{n+1} = {Y_n}^{*J} / \mathrm{E}(J)$ then the sequence Y_n converges in law to a random variable Y_∞ that is B-stable for J. Also Y_∞ is the unique B-stable law with mean $\mathrm{E}(Y_0)$. We also present results relating to random variables Y_0 with zero means and infinite means. The notion of B-stability arose in a scheme for cataloguing a large network of computers.

Keywords: Stable law; branching process; cataloguing large networks of computers

2000 Mathematics Subject Classification: Primary 60J80; 60E07; 60E10; 60F17

1. Introduction

Let Y be a random variable (RV). Denote by Y^{*j} the sum of j independent copies of Y. According to Feller (1971), Y is *strictly stable* if, for all j, there is a constant c_j such that $Y^{*j} \sim c_j Y$, where $X \sim Y$ means that X and Y have the same distribution. Suppose that J is an RV taking values in $\{1, 2, \dots\}$. We denote by Y^{*J} the RV that takes the value Y^{*j} with probability $p_j = \mathrm{P}(J = j)$. We define Y to be *B-stable* (i.e. *branching stable*) if and only if, for some scalar c (depending on the distributions of Y and J), we have $Y^{*J} \sim cY$. Then $((Y^{*J})^{*J}) \sim (cY)^{*J} \sim c^2 Y$, and so on. Let the probability generating function (PGF) of J be

$$f(u) = \sum_{j=1}^{\infty} u^j \, \mathrm{P}(J = j),$$

and let the characteristic function (CF) of Y be $\phi(t)$. Then the CF of Y^{*J} is $f(\phi(t))$, and the condition of B-stability is

$$f(\phi(t)) = \phi(ct), \tag{1}$$

for some c. Clearly, if $\mathrm{E}(Y)$ and $\mathrm{E}(J)$ are finite and $\mathrm{E}(Y) \neq 0$ then we must have $c = \mathrm{E}(J)$. Our object in this paper is to initiate the study of B-stable distributions. Clearly, if $\phi(t)$ satisfies (1) then so does $\phi(at)$ for any real a. We will ignore this trivial nonuniqueness. If $\mathrm{P}(J = j) = 1$, for some $j > 1$, then (1) reduces to $\phi(t)^j = \phi(c_j t)$, which (if it holds for all j, or even for any two relatively prime js; see Feller (1971)) requires Y to be (strictly) stable in the usual sense.

Received 8 September 2004; revision received 11 November 2004.
* Postal address: Avaya Labs, 233 Mount Airy Road, Basking Ridge, NJ 07920, USA.
Email address: colinm@research.avayalabs.com
** Postal address: Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA.
Email address: shepp@stat.rutgers.edu

https://doi.org/10.1239/jap/1118777189 Published online by Cambridge University Press

The reason for using the name 'B-stable' is that this new theory is related to the theory of branching processes, for which the classic reference is Harris (1963). A branching process $\{X_n\}$ is defined by setting $X_0 = 1$ and

$$X_n = J^{*X_{n-1}}, \qquad n = 1, 2, \dots,$$

i.e. X_n is the sum of X_{n-1} independent copies of J.

We can define another process $\{Y_n\}$ by taking Y_0 to be an arbitrary RV, and setting

$$Y_n = (Y_{n-1})^{*J}, \qquad n = 1, 2, \dots,$$

i.e. Y_n is the sum of J independent copies of Y_{n-1}. Note that the sequence $\{Y_n\}$ is a sequence of independent RVs.

2. Preliminary results

We begin with two simple lemmas.

Lemma 1. *If $Y_0 = 1$ then $Y_n \sim X_n$.*

Proof. As above, let the PGF of J be $f(u) = \sum_{j=1}^{\infty} u^j \, \mathrm{P}(J = j)$. The PGF of both X_0 and Y_0 is simply u. Suppose we have proved that both X_{n-1} and Y_{n-1} have the PGF $f_{n-1}(u)$. Then the PGF of X_n is $f_{n-1}(f(u))$, and the PGF of Y_n is $f(f_{n-1}(u))$. So both X_n and Y_n have the same PGF, namely the nth iterate of f.

Lemma 2. *For any Y_0, Y_n is distributed in the same way as the sum of X_n independent and identically distributed copies of Y_0.*

Proof. If the CF of Y_0 is $\phi(t)$ then we can easily show by induction that the CF of Y_n is $f_n(\phi(t))$.

Note that we have not shown that, for a single realization of $\{X_n\}$, we can define Y_n to be the sum of X_n copies of Y_0. This is not true because, with that definition, it is not the case that Y_n is distributed in the same way as the sum of J copies of Y_{n-1}. For example, suppose that $f(u) = (u + u^2)/2$. A possible realization of (X_0, X_1, X_2) is $(1, 2, 3)$, which occurs with probability $\frac{1}{2} \cdot \frac{2}{4} = \frac{1}{4}$. But, if Y_1 is the sum of two copies of Y_0 then Y_2 cannot be the sum of three copies of Y_0; it would have to be either the sum of two copies or the sum of four copies (with a probability of $\frac{1}{2}$ each).

In the classical theory, it is convenient to work with the process $\{X_n\}$ rather than $\{Y_n\}$, because if $\mathrm{E}(J)$ is finite and equal to μ, say, then the process

$$Z_n = \frac{X_n}{\mu^n}, \qquad n = 0, 1, 2, \dots,$$

is a martingale, and so converges almost surely to a random variable Z, say. (Clearly the distribution of Z depends on the distribution of J.) Thus, when $Y_0 = 1$, Y_n/μ^n must converge, in distribution, to Z. Proving this directly is not straightforward because Y_n/μ^n is not a martingale. In the following we will show that, for any Y_0 with nonzero finite mean μ, Y_n/μ^n does converge, in distribution, to a B-stable RV. We will always assume that $\mathrm{E}(J) = \mu$ is finite and greater than 1. Except for in Section 8, below, we assume that $\mathrm{P}(J = 0) = 0$.

3. Motivation

The motivation for this work was the following problem. In a certain method for cataloguing a large network of computers (see Mallows and Meloche (2005)), it is possible to search for active computers by sending packets to them in sequence, using their IP addresses. Whenever an active computer is found, it can be co-opted to help in the remaining search. As a simple model for this situation, consider a Bernoulli process Z on the positive integers, with $\mathrm{P}(Z_n = \{0, 1\}) = \{1 - p, p\}$ (in an obvious notation). The positions of the 1s represent the (unknown) positions of active computers. We denote the number of computers active at time t by K_t. We start at time $t = 0$ at position $X_0 = 0$ with one active computer ($K_0 = 1$), and move to the right with speed 1. At time $t = 1$ we are at position $X_1 = 1$ and if $Z_1 = 1$ we have found a second computer, so that $K_1 = 2$. If $Z_1 = 0$ then $K_1 = 1$. In general, if at time t we are at position X_t and have K_t active computers, then at time $t + 1$ we will be in position $X_{t+1} = X_t + K_t$ and will have found $K_{t+1} = K_t + \sum_{i=X_t+1}^{X_t+K_t} Z_i$ active computers.

It is easy to establish that $\mathrm{E}(K_t) = s^t$ and $p\,\mathrm{E}(X_t) = s^t - 1$, where $s = 1 + p$. We find that, as $t \to \infty$, both K_t/s^t and pX_t/s^t have asymptotically the same distribution, whose CF $\phi(u)$ satisfies

$$\phi(su) = (1 - p)\phi(u) + p\phi(u)^2.$$

This limiting distribution is thus B-stable with $f(u) = (1 - p)u + pu^2$. A similar analysis can be performed with computers appearing on the real line according to a Poisson process. This general problem seemed interesting to the authors.

4. Strictly stable random variables

First we describe the family of strictly stable RVs. These depend on three parameters as follows. Here $k(\alpha) = 1 - |1 - \alpha|$ and c is positive. Then $\psi(t) = \psi(t, \alpha, \beta, c)$ is the CF of a strictly stable RV if and only if $\psi(t) = \mathrm{e}^{w(t)}$, where

- $\alpha = 2$, $\beta = 0$, and $w(t) = -ct^2$ (i.e. Gaussian with mean 0); or
- $0 < \alpha < 2$ ($\alpha \neq 1$), $-1 \leq \beta \leq 1$, and $w(t) = -c|t|^\alpha \exp(-\frac{1}{2}\pi \mathrm{i}\beta k(\alpha)\mathrm{sgn}(t))$; or
- $\alpha = 1$, $-\infty < \beta < \infty$, and $w(t) = -c|t| + \mathrm{i}\beta t$ (i.e. translated Cauchy).

Remark 1. The asymmetric stable distributions with $\alpha = 1$ are not strictly stable; they cannot be located to satisfy the property of strict stability. Neither Feller (1971) nor Loève (1960) pointed out that translated Cauchy laws are strictly stable.

5. B-stability when E(Y_0) is finite and nonzero

Theorem 1. *If* $\mathrm{E}(Y_0)$ *is finite and nonzero then* $Y_n/\mathrm{E}(Y_0)\mu^n$ *converges, in distribution, to* Z.

Proof. As above, define $X_0 = 1$ and $X_n = J^{*X_{n-1}}$, so that $\{X_n\}$ is a branching process. Then X_n/μ^n is a martingale and, so, converges almost everywhere to an RV Z, say. Let N_n be a sequence of independent RVs with $N_n \sim X_n$. Then $N_n/\mu^n \to Z$ in distribution. By Lemma 2, we can assume (for each n separately) that Y_n is the sum of N_n copies of Y_0. Then

$$\frac{Y_n}{\mu^n} = \frac{Y_n}{N_n}\frac{N_n}{\mu^n} \tag{2}$$

https://doi.org/10.1239/jap/1118777189 Published online by Cambridge University Press

and, as $n \to \infty$, we have $N_n \to \infty$, in distribution, so the first fraction on the right-hand side of (2) tends to the constant $\mathrm{E}(Y)$ by the strong law of large numbers, while the second fraction on the right-hand side of (2) tends to Z, as shown above.

6. B-stability when $\mathrm{E}(Y)$ is not finite

There are positive B-stable RVs that do not have finite means. Suppose that $0 < \alpha < 1$ and S_α is a positive, strictly stable RV with index α. The Laplace transform of the distribution of S_α is $\mathrm{E}(\mathrm{e}^{-sS_\alpha}) = \mathrm{e}^{-s^\alpha}$, for $s > 0$. Define

$$Y = S_\alpha Z^{1/\alpha},$$

where Z is (as above) the limit law of X_n/μ^n, and S_α and Z are independent. Suppose that the Laplace transform of the distribution of Z is $\zeta(s)$. Then ζ satisfies

$$f(\zeta(s)) = \zeta(\mu s).$$

The Laplace transform $g(s)$ of the distribution of Y is

$$g(s) = \mathrm{E}(\mathrm{e}^{-sY}) = \mathrm{E}(\mathrm{E}(\mathrm{e}^{-sY} \mid Z)) = \mathrm{E}(\mathrm{e}^{-s^\alpha Z}) = \zeta(s^\alpha),$$

and we have

$$f(g(s)) = g(cs),$$

with $c = \mu^{1/\alpha}$.

The next theorem shows that this construction also works when the stable RV S_α is not necessarily positive, for $\alpha < 1$, and it also works when $\alpha = 1$, in the strictly stable case.

Theorem 2. *If Y_0 is in the domain of attraction of a strictly stable RV S_α, with $0 < \alpha \leq 1$, then $Y_n/\mu^{n/\alpha}$ converges, in distribution, to $S_\alpha Z^{1/\alpha}$, where S_α and Z are independent.*

Proof. We use the same device as in Theorem 1, introducing a branching process $\{X_n\}$, defining independent $N_n \sim X_n$, and defining Y_n to be the sum of N_n independent copies of Y_0. Then we obtain

$$\frac{Y_n}{\mu^n} = A_n B_n^{1/\alpha},$$

where

$$A_n = \frac{Y_n}{N_n^{1/\alpha}}, \qquad B_n = \frac{N_n}{\mu^n}.$$

We need to be careful because now both factors tend to nontrivial limits, and we need to establish that they are independent in the limit. We are assuming that, as $n \to \infty$, $Y_0^{*n}/n^{1/\alpha} \to S_\alpha$, so, given any $\varepsilon > 0$, we can find n_0 sufficiently large that, for $m > n_0$, we have

$$\left| \mathrm{P}\left(\frac{Y_0^{*m}}{m^{1/\alpha}} < y\right) - \mathrm{P}(S_\alpha < y) \right| < \varepsilon, \tag{3}$$

for all y. Since $X_n \to \infty$ with probability 1, there exists an n_1 sufficiently large that, for $n > n_1$, we have $\mathrm{P}(N_n \leq n_0) < \varepsilon$. Thus, for $n > n_1$ with probability at least $1 - \varepsilon$ for each value m of X_n, we have (3). It follows that, asymptotically, A_n and B_n are independent, with $A_n \sim S_\alpha$ and $B_n \sim Z$.

7. B-stability with zero means

When $\mathrm{E}(Y_0) = 0$ we get a similar class of limits; the following theorem can be proved in the same way as Theorem 2.

Theorem 3. *If Y_0 is in the domain of attraction of a strictly stable RV S_α, with $1 < \alpha \le 2$ and $\mathrm{E}(Y_0) = 0$, then $Y_n/\mu^{n/\alpha}$ converges, in distribution, to $S_\alpha Z^{1/\alpha}$, where S_α and Z are independent.*

Of course, if $\mathrm{E}(Y_0)$ is nonzero then Theorem 1 applies.

8. Examples

Two tractable examples of branching processes were given by Harris (1963).

Example 1. Take $\mu > 1$ and

$$f(u) = \frac{u}{\mu - (\mu - 1)u},$$

so that J is a geometric RV (shifted by 1), with mean $\mathrm{E}(J) = \mu$. It is easily verified that $f_n(u)$ is of the same form, with μ replaced by μ^n. Taking the limit as $n \to \infty$, we find that Z is exponential, i.e. $\mathrm{E}(\mathrm{e}^{-sZ}) = 1/(1+s)$. From Section 4, we know that the CF of S_α is of the form $\mathrm{e}^{w(t)}$, so that

$$\mathrm{E}(\mathrm{e}^{\mathrm{i}tY}) = \mathrm{E}(\mathrm{e}^{\mathrm{i}tS_\alpha Z^{1/\alpha}}) = \mathrm{E}(\mathrm{e}^{w(tZ^{1/\alpha})}) = \mathrm{E}(\mathrm{e}^{w(t)Z}) = \frac{1}{1 - w(t)}.$$

Example 2. Take

$$f(u) = \frac{u}{(\mu - (\mu - 1)u^k)^{1/k}},$$

for some integer $k \ge 2$. Again, $\mathrm{E}(J) = \mu$, and f_n is obtained by replacing μ by μ^n. Here $\mathrm{e}^{\mathrm{i}tZ} = 1/(1 - \mathrm{i}t)^{1/k}$.

Two new examples are as follows. Firstly, take $f(u) = u^2/(2 - u^2)$. Then $\mathrm{E}(\mathrm{e}^{-sZ}) = 1/\cosh(\sqrt{s})$. Secondly, take $f(u) = u^2/(2-u)^2$. Then $\mathrm{E}(\mathrm{e}^{-sZ}) = 1/(\cosh(\sqrt{s}))^2$.

9. The case $\mathrm{P}(J = 0) > 0$

Suppose that $p_0 = \mathrm{P}(J = 0) > 0$, while $\mathrm{E}(J) > 1$. A classical argument shows that there is a unique ξ in $(0, 1)$ such that $f(\xi) = \xi$ and $\mathrm{P}(X_n = 0) \to \xi$. Thus, $\mathrm{P}(Z = 0) = \xi$. We define an *associated RV* J' with PGF $f'(u)$, where

$$f(\xi + (1 - \xi)u) = \xi + (1 - \xi) f'(u), \qquad 0 \le u \le 1.$$

Then $\mathrm{E}(J') = \mathrm{E}(J) = \mu$ and $\mathrm{P}(J' = 0) = 0$. Corresponding to J and J' we have branching processes $\{X_n\}$ and $\{X'_n\}$, respectively; the limits of X_n/μ^n and X'_n/μ^n are Z and Z' with Laplace transforms $\zeta(s)$ and $\zeta'(s)$, respectively.

Lemma 3. *We have $\zeta(s) = \xi + (1 - \xi)\zeta'(s)$.*

Proof. This follows immediately from the definitions.

Thus, the conditional distribution of Z given $\{Z > 0\}$ is simply Z'. We immediately obtain the following theorem.

Theorem 4. *Theorems 1–3 apply when* $\mathrm{P}(J = 0) > 0$ *(provided that* $\mathrm{E}(J) > 1$*).*

10. Miscellaneous remarks

Remark 2. (*A case with* $\mathrm{E}(J)$ *infinite.*) Suppose that $\mathrm{E}(u^J) = 1 - (1-u)^\gamma$, for some γ in (0, 1). We find that $\mathrm{E}(u^{X_k}) = 1 - (1-u)^{\gamma^k}$, and some messy computations show that the median m_k of X_k is approximately e^{1/γ^k}, and that the distribution of X_k/m_k approaches an improper distribution, i.e. $\mathrm{P}(X_k/m_k < x) \to \frac{1}{2}$ for all $x > 0$. The assumption that $\mathrm{E}(J)$ is finite seems to be essential to obtain proper limits.

Remark 3. The distribution of Z is determined by the distribution of J, but not every (positive) RV Z can be obtained. Given a Laplace transform $\phi(s) = \mathrm{E}(\mathrm{e}^{-sZ})$ and a constant $\mu > 1$, we can always define the function $f(u) = \phi(\mu\phi^{-1}(u))$, but this need not be a proper PGF. For example, if Z is exponential and $\phi(s) = 1/(1+s)$ then we find that $f(u) = u/(\mu - (\mu-1)u)$ for any $\mu > 1$. However, if $\phi(s) = 1/\cosh(\sqrt{s})$ then $f(u)$ is a proper PGF only when μ is the square of an integer.

Remark 4. It is known that stable distributions are unimodal, but B-stable distributions need not be. Suppose that $f(u) = (u + u^k)/2$ with k very large, so $\mu = (k+1)/2$. The density of Z is close to $(a(z) + b(z))/2$, where $a(z)$ is approximately Gaussian, with mean 2 and variance $1/k$, and $b(z)$ is concentrated near $z = 0$.

Remark 5. It is known that stable distributions are infinitely divisible, but B-stable distributions need not be. For example, suppose that

$$f(u) = \frac{u^n + u^{n+1} + \cdots + u^{2n}}{n+1},$$

with n large. Then Z is approximately uniform on $(\frac{1}{2}, \frac{3}{2})$ and cannot be written as $U + U'$ with $U \sim U'$.

Remark 6. If J has moments up to order k, say, then the defining relation (1) allows the moments of Z to be computed up to the same order. The computations are simplified if we use cumulants instead of moments; it follows from (1) that the cumulant-generating functions of J and Z, respectively $g(t)$ and $h(t)$, are related by

$$g(h(t)) = h(\mu t).$$

Remark 7. There are several open questions. We have shown that for each distribution (of J) on the nonnegative integers, with $\mathrm{E}(J) > 1$, there is a three-parameter family of B-stable distributions. We cannot prove that these are the only possibilities. We do not know of a convenient characterization of the possible limit distributions. We would like a fuller understanding of the construction of $\phi(\mu\phi^{-1}(u))$ described above. There are several ways in which the classical branching process can be generalized; can our constructions be generalized in similar ways?

References

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications,* Vol. 2, 2nd edn. John Wiley, New York.
Harris, T. (1963). *Branching Processes*. Prentice-Hall, Englewood Cliffs, NJ.
Loève, M. (1960). *Probability Theory*. Van Nostrand, Princeton, NJ.
Mallows, C. L. and Meloche, J. (2005). Searching for searchers. Submitted.