Check for updates

J. Appl. Probab. 1–13 (2026)
*doi:*10.1017/jpr.2026.10094

REMARKS ON KOLMOGOROV'S CONSTANT FOR SIMPLE BRANCHING PROCESSES

ANTHONY G. PAKES,* *University of Western Australia*

Abstract

Recent investigations have argued that there is a simple explicit representation for the Kolmogorov constant c associated with the subcritical Galton–Watson branching process. We exhibit examples showing that although this representation can be valid, it more often is not. Our work is presented in terms of the limiting conditional mean population size $\mu = c^{-1}$. The analogous quantity for the Markov branching process is denoted by $\widehat{\mu}$. We show that the simple representation put forward for $\widehat{\mu}$ in fact is an upper bound that is attained only if the offspring-number probability-generating function is quadratic. The conditional mean μ is the limit of a computable increasing sequence (μ_n). Estimates of n are determined ensuring that, for any small positive number ε, $0 < \mu - \mu_n \le \varepsilon$.

Keywords: Galton–Watson process; Markov branching process; Kolmogorov constant; limiting conditional mean size

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60F99; 60J22

1. Introduction

We are concerned with the two Markovian branching process models. The discrete-time model describes the growth of a population of individuals who live for a unit of time and, at the end of their lives, produce $j = 0, 1, \ldots$ offspring with probability p_j and all reproduction events are mutually independent. The number of individuals alive at time $n = 0, 1, \ldots$ is denoted by Z_n. To avoid trivialities we assume that $0 < p_0 + p_1 < 1$. The offspring probability-generating function (PGF) is denoted by $f(s) = \sum_{j=0}^{\infty} p_j s^j$.

The sequence $(Z_n : n \ge 0)$ is a Markov chain whose transition probabilities $p_{ij}^{(n)} = P_i(Z_n = j)$, where $P_i(\cdot) = P(\cdot \mid Z_0 = i)$, are given by the generating function identity

$$\sum_{j=0}^{\infty} p_{ij}^{(n)} s^j = (f_n(s))^i,$$

where $f_0(s) = s$ and $f_n(s) = f(f_{n-1}(s))$ $(n \ge 1)$ is the nth functional iterate of $f(s)$. The power form in this equation reflects the independent development of differing lines of descent.

Clearly the zero state is absorbing – an empty population never recovers. Always assuming that $Z_0 > 0$, denote the time of extinction by $T := \inf\{n \ge 1 : Z_n = 0\}$. We are concerned with

Received 5 August 2025; accepted 9 March 2026.
* Postal address: Department of Mathematics and Statistics, The University ofWestern Australia, 35 Stirling Highway, Crawley, WA, Australia 6009. Email: tony.pakes@uwa.edu.au

© The Author(s), 2026. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

the subcritical case, i.e. the mean number of offspring $m := \sum_{j\geq 1} jp_j < 1$. In this case $p_0 > 0$ and $P_i(T < \infty) = 1$, and the distribution function of T is $P_i(T \leq n) = f_n^i$, where $f_n = f_n(0)$.

An early and fundamental result of the subject, due to Kolmogorov in 1938, asserts the existence of the limit $c = \lim_{n\to\infty} m^{-n}(1 - f_n)$. It is clear that $c \leq 1$ because $\mu_n := E_1(Z_n \mid T > n) = m^n/(1 - f_n) > 1$, where $E_i(\cdot) = E(\cdot \mid Z_0 = i)$. In fact, $c < 1$. Kolmogorov assumed the second moment condition $2b := f''(1) = E_1[Z_1(Z_1 - 1)] < \infty$, and proved that $c > 0$. This yields the exact decay rate of the right-hand tail of the extinction time law: $\lim_{n\to\infty} m^{-n} P_i(T > n) = ic$; see [4].

Independent investigations later rendered Kolmogorov's theorem into an ultimate form. In particular, [10] gives a rather opaque necessary and sufficient condition for $c > 0$ and then some simpler sufficient conditions. It is shown in [6] and in [11] that $c > 0$ if and only if the so-called xlogx condition, $\sum_{j\geq 1} p_j j \log j < \infty$, holds.

There seems to be no known 'explicit' expression for the Kolmogorov constant when it is positive. However, it is claimed in [8, Theorem 1.2] that if $b < \infty$, then

$$c = \left[1 + \frac{b}{m(1-m)}\right]^{-1}. \tag{1.1}$$

This identity is at least implicit in the earlier communication [9] (set $s = 0$ in (7) and (8) of that paper). One of our aims is to show that (1.1) can be valid, but it is not generally true.

We approach this in Section 2 by exhibiting one numerical and several analytical counter-examples. They are presented in terms of the first moment $\mu = c^{-1}$ of the limiting conditional law (LCL), $b_j = \lim_{n\to\infty} P_i(Z_n = j \mid T > n)$ $(j \geq 1)$, where, as the notation suggests, the limit is independent of the initial state i. We also give a very simple proof of an upper estimate of μ derived in [1], obtained there by bounding $f(s)$ with linear-fractional PGFs.

We also consider the Markov branching process which evolves as a Galton–Watson process (GWP) with the additional feature that individual life durations vary according to an exponential law with density function ae^{-at} $(t \geq 0)$, where $a > 0$ is called the split rate, and all life times and reproduction events are independent. If Y_t denotes the population size at time t then $E_i(s^{Y_t}) = (F(s, t))^i$, where the PGF $F(s, t)$ solves the backward Kolmogorov equation $\partial F(s, t)/\partial t = \mathcal{U}(F(s, t))$, $\mathcal{U}(s) = a(f(s) - s)$ and $F(s, 0) = s$.

The time to extinction $\widehat{T} = \inf\{t \geq 0 : Y_t = 0\}$ has the distribution function $F(t) = F(0, t)$ and the analogue of Kolmogorov's theorem is that, with $\nu = a(1 - m)$,

$$\widehat{c} = \lim_{t\to\infty} e^{\nu t}(1 - F(t))$$

exists. This was established by Sevast'yanov in 1951, again assuming $b < \infty$, and in general a little later in [18]. See [4, Theorem 11.1], where an integral expression will yield a representation for $\widehat{c}$. The rendering in [18, Theorem 3] is more explicit where, after a simple change of variable,

$$\widehat{c} = \exp\left[-\int_0^1 \frac{a(1-s) + \mathcal{U}(s)}{(1-s)\mathcal{U}(s)}\, ds\right].$$

The integral is finite if and only if the xlogx condition holds. It is claimed in [7, Theorem 1] that if $b < \infty$, then

$$\widehat{c} = \left[1 + \frac{b}{1-m}\right]^{-1}. \tag{1.2}$$

The counter-examples in Section 2 are also counter-examples to (1.2). The problem of deciding if an arbitrary probability mass function (PMF) $(b_j : j \geq 1)$ on $\mathbb{N}$ satisfying $b_1 > 0$ is

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

the LCL of some subcritical Markov branching process (MBP) is solved in [12, 17]. Examples therein provide further counter-examples to (1.2). All these are cast in terms of $\widehat{\mu} = \widehat{c}^{-1}$, the mean of the LCL. Thus, the assertion (1.2) is equivalent to

$$\widehat{\mu} = \widehat{\mu}_I := 1 + \frac{b}{1-m}. \tag{1.3}$$

We will see for most of our examples that $\widehat{\mu} < \widehat{\mu}_I$. Theorem 3.1 asserts that $\widehat{\mu} \le \widehat{\mu}_I$ with equality only if $f(s)$ is a quadratic function; equivalently, the corresponding MBP is a 'generalised' linear birth and death process. It will follow that strict inequality holds for Examples 2.3–2.5. Our inequality implies a lower-bound estimate of $\widehat{\sigma}^2$, the variance of the LCL.

Referring to the GWP, it is known [13] that $\mu_n = m^n/(1-f_n) \uparrow \mu$ as $n \uparrow \infty$. Bounds for μ have been derived [13, 16] having the form $\mu_n + A_n \le \mu \le \mu_n + B_n$, where $A_n, B_n = O(m^n)$. This raises the question of how large n must be to approximate μ by μ_n to within a prescribed error. We address this question in Section 4. Suffice to say here that large values of n seem to be required only if m is close to unity.

In Section 5 we comment on two issues in [8] that mistakenly lead to (1.1) as an identity valid for all offspring-number laws.

2. The Galton–Watson process

We begin by recalling that if $i, j = 1, 2, \ldots$, then $b_j = \lim_{n\to\infty} P_i(Z_n = j \mid T > n)$ exists independently of the initial state, and $\sum_{j=1}^{\infty} b_j = 1$. This defines the LCL of the process. The PGF $B(s) = \sum_{j=1}^{\infty} b_j s^j$ is the unique PGF solution of the functional equation

$$1 - B(f(s)) = m(1 - B(s)). \tag{2.1}$$

This result was obtained by Yaglom in 1947 assuming that $b < \infty$ and the general assertion was later independently published in [6, 10]. See [2, 3] for monograph accounts, particularly the former.

Iterating (2.1) and setting $s = 0$ yields the identity

$$\frac{1 - B(f_n)}{1 - f_n} = \frac{m^n}{1 - f_n},$$

from which it is obvious [6, 10] that $\mu := B'(1) = c^{-1}$, i.e. the mean of the LCL equals the limit of the conditional expectation $E(Z_n \mid T > n)$. We find it more convenient to work with μ. Thus, (1.1) becomes the assertion that

$$\mu = \mu_I := 1 + \frac{b}{m(1-m)}, \tag{2.2}$$

provided that $b < \infty$.

Example 2.1. We look first at the fractional linear offspring law ([3, pp. 6–7] or [4, p. 7]) where, changing some notation from these references,

$$f(s) = 1 - \frac{\beta}{1-p} + \frac{\beta s}{1 - ps}$$

with $0 < p < 1$ and $0 < \beta < 1 - p$. The offspring mean is $m = \beta/(1-p)^2 < 1$, and $b = \beta p/(1-p)^3 = mp/(1-p)$. It follows that

$$\mu_I = \frac{1 - p - \beta}{(1-p)^2 - \beta} > 1.$$

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

The equation $f(s) = s$ has roots $s = 1$ and [4, p. 9]

$$s = s_0 = \frac{1-p-\beta}{p(1-p)}.$$

In addition, $1 - f_n \sim m^n(1 - s_0^{-1})$. It follows that

$$\mu = (1 - s_0^{-1}) = \frac{1-p-\beta}{(1-p)^2-\beta},$$

i.e. (2.2) is valid in this case. We shall see why in Section 3.

Example 2.2. A numerical counter-example to (2.2) arises from the Poisson offspring law, $f(s) = \mathrm{e}^{-m(1-s)}$. Bounds for μ derived in [15, p. 475] are

$$\frac{2-m}{2(1-m)} \le \mu \le \frac{1+(1-m)^2}{(1-m)(2-m)}. \tag{2.3}$$

Clearly $b = m^2/2$ and hence

$$\mu_I = 1 + \frac{m}{2(1-m)} = \frac{2-m}{2(1-m)},$$

the lower bound (2.3). The proof steps in [15] for (2.3) indicate that the weak inequalities there are in fact strict. This assertion is supported by [15, Table II], where μ is computed to four decimal places for $m = 0.1, 0.2, \ldots, 0.9, 0.95$ and $\mu_I < \mu$ in all cases, although the difference is surprisingly small. We remark that the value given there for μ_I when $m = 0.2$ is incorrect – it should be 1.125.

In both examples we have $\mu_I \le \mu$. Might this be generally true? We shall see below that the answer is no!

A family of offspring laws embedding the fractional linear case was introduced in [14]. This family is indexed by a primary parameter $\theta \in [-1, 1]$, along with secondary parameters. The entire family resolves into nine subfamilies, designated as cases. In all cases, explicit expressions are given for m, $2b$, and the functional iterates $f_n(s)$. We review those cases for which the offspring law is subcritical.

Case 1 has $b < \infty$ in only the boundary case $\theta = 1$ and it is the linear fractional offspring law. The secondary parameters in [14] are $a = (1-p)/\beta$ and $d = (1-p)p/(1-p-\beta)$. We say no more about this case. The other relevant cases specify a parameter $a \in (0, 1)$, which turns out to be the offspring mean, m. Hence we replace a in [14] with m.

Example 2.3. Case 7 has $\theta \in (0, 1)$ and a parameter $A > 1$. There is a second parameter q, which here must have the value $q = 1$ to ensure that $\sum p_j = 1$. Setting $p = A^{-1}$, the functional iterates have the form

$$\begin{aligned} f_n(s) &= A - [m^n(A-s)^{-\theta} + (1-m^n)(A-1)^{-\theta}]^{-1/\theta} \\ &= A[1 - (m^n(1-ps)^{-\theta} + (1-m^n)(1-p)^{-\theta})^{-1/\theta}]. \end{aligned}$$

In addition,

$$\frac{2b}{m(1-m)} = \frac{1+\theta}{A-1} = (1+\theta)\frac{p}{1-p}.$$

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

Hence,

$$\mu_I = 1 + \frac{1+\theta}{2} \cdot \frac{p}{1-p}.$$

Setting $s = 0$, some manipulation with the first form of $f_n(s)$ yields

$$\begin{aligned} f_n &= A - (A-1)[1 - m^n(1-(1-p)^\theta)]^{-1/\theta} \\ &= 1 - \theta^{-1}(A-1)m^n[1-(1-p)^\theta](1+o(1)). \end{aligned}$$

It follows that

$$\mu = \frac{\theta}{A-1} \cdot \frac{1}{1-(1-p)^\theta} = \frac{\theta p}{(1-p)(1-(1-p)^\theta)},$$

manifestly different to μ_I.

Further manipulation reveals that

$$B(s) = \frac{1-(1-ps)^\theta}{1-(1-p)\theta} \cdot \left(\frac{1-p}{1-ps}\right)^\theta,$$

the PGF of the sum of two independent random variables, $S_{p,\theta} + N_{p,\theta}$, where $S_{p,\theta}$ has a power-series law generated by the Sibuya law whose PGF is $1-(1-s)^\theta$, and $N_{p,\theta}$ has a negative binomial law.

Example 2.4. For Case 8 in [14], the relevant parameter choices are $\theta = 0$, $q = 1$, and $A > 1$, and then

$$f_n(s) = A - (A-1)^{1-m^n}(A-s)^{m^n} \tag{2.4}$$

with $2b/m(1-m) = (A-1)^{-1}$. It follows that

$$\mu_I = 1 + \frac{1}{2(A-1)} = \frac{2-p}{2(1-p)},$$

the same algebraic form as for the Poisson case.

Setting $s = 0$ in (2.4) yields

$$\begin{aligned} f_n &= (A-1)\left[\left(\frac{A}{A-1}\right)^{m^n} - 1\right] = (A-1)\left[(1-p)^{-m^n} - 1\right] \\ &= (A-1)m^n(-\log(1-p))(1+o(1)), \end{aligned}$$

whence

$$\mu = [(A-1)(-\log(1-p))]^{-1} = \frac{p}{1-p}(-\log(1-p))^{-1}.$$

Again, $\mu_I \neq \mu$.

We can compare μ_I and μ by ignoring the common factor $1-p$ and observing that

$$\frac{p}{-\log(1-p)} = \frac{1}{1+p/2+p^2/3+\cdots} > \frac{1}{1+(p/2)(1+p+p^2+\cdots)} = \frac{1-p}{1-p/2}.$$

But

$$1 - p/2 - \frac{1-p}{1-p/2} = \frac{p^2/4}{1-p/2} > 0,$$

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

and we conclude that for this case, $\mu_I > \mu$. Hence there is no general order relation between μ and μ_I. Numerical calculation shows for this case that the relative error $\mu_I/\mu - 1$ is within 0.04 if $m \in (0, 0.5]$, and it increases quite quickly if $m \in [0.6, 1)$, e.g. it is 0.113 if $m = 0.6$.

Finally, calculation shows that the LCL has the log-series law,

$$B(s) = \frac{\log(1-ps)}{\log(1-p)}.$$

Example 2.5. Our last example, Case 9 in [14], has $\theta \in (-1, 0)$, $A > 1$, and $q = 1$. Writing $\zeta = -\theta$, the functional iterate is $f_n(s) = A - [m^n(A-s)^\zeta + (1-m^n)(A-1)^\zeta]^{1/\zeta}$ and $b/m(1-m) = (1-\zeta)/(A-1)$. Hence,

$$\mu_I = \frac{2A-1-\zeta}{2(A-1)} = \frac{2-(1+\zeta)p}{2(1-p)}.$$

Setting $s = 0$ we obtain

$$1 - f_n = \zeta^{-1}(A-1)m^n\left[\left(\frac{A}{A-1}\right)^\zeta - 1\right](1+o(1)),$$

and hence

$$\mu = \frac{\zeta p}{1-p}((1-p)^{-\zeta} - 1)^{-1} \neq \mu_I.$$

Numerical calculation suggests that $\mu_I > \mu$. We shall confirm this in Section 3.

The LCL is a power-series law generated from a Sibuya law,

$$B(s) = \frac{1-(1-ps)^\zeta}{1-(1-p)^\zeta}.$$

We end this section by recalling the upper bound for μ derived in [1] using bounding linear-fractional offspring laws.

Theorem 2.1. *If $b < \infty$, then*

$$\mu < 2\mu_I - 1 = 1 + \frac{2b}{m(1-m)}. \tag{2.5}$$

Proof. A double differentiation of (2.1) yields $2b\mu + m^2B''(1) = mB''(1)$. Observing that, in obvious notation, $B''(1) = \sigma_B^2 + \mu^2 - \mu$, division by μ yields

$$2b = m(1-m)\left(\sigma_B^2 + \mu - 1\right) > m(1-m)(\mu - 1),$$

and (2.5) follows. □

This simple proof shows that the weak inequality in [1, p. 331] is in fact strict, and that (2.5) is equivalent to the crudest possible lower bound of σ_B^2. Can (2.5) be refined?

The presentation in [8] is in terms of the non-critical case $m \neq 1$. That the subcritical case is fundamental follows because, as is known, it subsumes the supercritical case in matters of ultimate extinction. Let $\overline{f}(s)$ be an offspring-number PGF satisfying $\overline{f}(0) > 0$ and $\overline{f}'(1) \in (1, \infty]$. Then $\overline{f}(s) = s$ has a unique solution $q \in (0, 1)$, the probability that the associated GWP $(\overline{Z}_n : n \geq 0)$ with $\overline{Z}_1 = 1$ hits the zero state. The Sevast'yanov transformation $f(s) = \overline{f}(qs)/q$

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

determines a subcritical offspring-number law: $m = \overline{f}'(q) < 1$ (denoted by β in [8]). Clearly $f_n(s) = \overline{f}_n(qs)/q$, hence $\overline{f}_n := \overline{f}_n(0) = qf_n \uparrow q$.

A natural definition of a supercritical Kolmogorov constant is $\overline{c} = \lim_{n\to\infty} m^{-n}(q - \overline{f}_n) = qc$, where c is the Kolmogorov constant for the induced subcritical GWP. Note that $\overline{c}$ is denoted by $\mathcal{K}_q$ in [8].

Defining $\overline{b} = \frac{1}{2}\overline{f''}(q) = b/q$ (denoted by b_q in [8]), [8, Theorem 2.1] asserts that

$$\overline{c} = \frac{q}{1 + \overline{b}/m(1-m)} = q\left(1 + \frac{b}{m(1-m)}\right)^{-1}.$$

It follows that examples/counter-examples for the supercritical case correspond to any, such as those above, constructed from a subcritical $f(s)$ whose domain of holomorphy extends sufficiently far beyond the unit disc as to have a fixed point in $(1, \infty)$. For example, the Poisson offspring-number laws.

3. The Markov branching process

We begin by recalling that the subcritical MBP has a stationary measure $(\pi_j : j \geq 1)$ whose generating function is

$$\pi(s) = \int_0^s \frac{1-m}{f(u) - u}\, \mathrm{d}u, \tag{3.1}$$

where $f(s)$ is the offspring PGF of the MBP. Note that this stationary measure is independent of the split rate and it is unique up to scaling. The scaling chosen for (3.1) is the most natural in the sense that the PGF of the LCL is

$$\widehat{B}(s) := \lim_{t\to\infty} E_i\big(s^{Y_t} \mid \widehat{T} > t\big) = 1 - \mathrm{e}^{-\pi(s)}. \tag{3.2}$$

See [4] for the most explicit statements and [3] for a simple proof based on the fact that, for each constant $\delta > 0$, the so-called discrete skeleton $(Y_{\delta n} : n \geq 0)$ is a GWP. It follows from (3.1) and (3.2) that

$$\widehat{\mu} = \lim_{s\uparrow 1} \frac{1 - \widehat{B}(s)}{1-s} = \lim_{s\uparrow 1} \mathrm{e}^{-\pi(s) - \log(1-s)} = \exp\left[-\int_0^1 \left(\frac{1-m}{f(s)-s} - \frac{1}{1-s}\right) \mathrm{d}s\right]. \tag{3.3}$$

On the other hand, it follows from Galton–Watson theory that $\widehat{\mu} = \widehat{c}^{-1}$ and hence (1.2) is equivalent to the assertion that

$$\widehat{\mu} = \widehat{\mu}_I := 1 + \frac{b}{1-m}. \tag{3.4}$$

We now show that (3.4) follows from (2.2). Recall that the mean population size of the MBP is $m_t = E_1(Y_t) = \mathrm{e}^{-\nu t}$, where $\nu = a(1-m)$, the negative of the Malthusian parameter. Also, let $b_t = \frac{1}{2}E_1[Y_t(Y_t - 1)]$. It follows from (2.2) applied to the discrete skeleton that, for any $t > 0$,

$$\widehat{\mu}_I = 1 + \frac{b_t}{m_t(1-m_t)}.$$

A long-known result is that [4, p. 103]

$$\mathrm{Var}(Y_t \mid Y_0 = 1) = \frac{f''(1) - m + 1}{1-m} m_t(1-m_t) = \left(\frac{2b}{1-m} + 1\right) m_t(1-m_t).$$

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

However, the variance equals $2b_t + m_t(1-m_t)$ and hence

$$\frac{b_t}{m_t(1-m_t)} \equiv \frac{b}{1-m},$$

and the expression in (3.4) for $\widehat{\mu}_I$ follows.

It is observed in [14, Section 7] that the θ-offspring laws arise as discrete skeletons of certain MBPs. An alternative approach to constructing examples of LCLs is taken in [12, 17] by finding criteria which ensure that a PMF $(b_j : j \geq 1)$ is the LCL of a subcritical MBP, in fact an uncountable family of MBPs. Examples in these papers provide counter-examples to (3.4).

Example 3.1. Choose a constant $\lambda \in (0, 1)$, set $\theta = \lambda e^{-\lambda}$, and let C(z) be the Cayley tree-counting function, i.e. the solution of the functional equation $C = ze^C$, where $z \geq 0$, which satisfies $0 \leq C(z) \leq e^{-1}$. The Cayley function is related to the better-known principal Lambert function W(z) by $C(z) = -W(-z)$. It is shown in [17, Section 3] that $\widehat{B}(s) = \lambda^{-1}C(\theta s)$ is the PGF of the LCL of MBPs whose offspring PGF is

$$f(s) = [1-(1-m)(1+\lambda)]s + (1-m)s[\lambda\widehat{B}(s) + 1/\widehat{B}(s)],$$

where the offspring mean m is chosen subject to the constraint

$$\frac{2\lambda}{1+\lambda} \leq m < 1.$$

The derivative identity

$$C'(z) = \frac{C(z)}{z(1-C(z))}$$

leads to the evaluations

$$\widehat{\mu} = \frac{1}{1-\lambda} < \widehat{\mu}_I = \frac{1}{2}\left[1 + \frac{1}{(1-\lambda)^2}\right].$$

Example 3.2. Choosing constants $p \in (0, 1)$ and $\alpha \in (0, 1]$, it is shown in [12] that

$$\widehat{B}(s) = 1 - \frac{1-s}{(1-ps)^\alpha}$$

generates an LCL. If $\alpha = 1$, then this law is a shifted geometric distribution. See Section 5 for more on this case. It is obvious that $\widehat{\mu} = (1-p)^{-\alpha}$.

The corresponding offspring probabilities for $j \geq 2$, defining $\rho = (1-\alpha)p/(1-\alpha p)$, are

$$p_j = \frac{1-m}{1-\alpha p}(1-\rho)(p-\rho)\rho^{j-2},$$

where the offspring mean must satisfy

$$1 - \frac{(1-\alpha p)^2}{1-\alpha p^2} \leq m < 1.$$

Observing that $\sum_{j\geq 2} j(j-1)\rho^{j-2} = 2(1-\rho)^{-3}$, some algebra leads to

$$\widehat{\mu}_I = \frac{1-(1-\alpha)p}{1-p}.$$

This equals $\widehat{\mu}$ if $\alpha = 1$, but not if $\alpha < 1$.

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

We end this section by showing that $\widehat{\mu}_I$ always bounds $\widehat{\mu}$ from above, and that this implies a lower bound for the variance $\widehat{\sigma}^2$ of the LCL.

Theorem 3.1.

(i) *If $f''(1) < \infty$, then*

$$\widehat{\mu} \le \widehat{\mu}_I \tag{3.5}$$

and equality holds only if $p_j = 0$ when $j > 2$, i.e. the underlying MBP is a (generalised) linear birth and death process.

(ii) *In addition,*

$$\widehat{\sigma}^2 \ge \widehat{\mu}(\widehat{\mu} - 1), \tag{3.6}$$

and the right-hand side is positive provided $p_0 + p_1 < 1$.

Proof. For (i), referring to (3.3), a Taylor expansion yields

$$f(s) = 1 - m(1-s) + b(s)(1-s)^2,$$

where $\frac{1}{2}f''(s) < b(s) \le b$, and $b(s) \equiv b$ if and only if $f(s)$ is a quadratic function. Hence

$$\frac{1-m}{f(s)-s} = [(1-s)(1+(1-s)b(s)/(1-m))]^{-1},$$

i.e.

$$0 < \frac{1}{1-s} - \frac{1-m}{f(s)-s} = \frac{b(s)/(1-m)}{1+(1-s)b(s)/(1-m)} \le \frac{b/(1-m)}{1+(1-s)b/(1-m)}.$$

Hence, the exponent in (3.3) is bounded above by

$$\int_0^1 \frac{b/(1-m)}{1+ub/(1-m)}\,\mathrm{d}u = \log\left(1 + \frac{b}{1-m}\right),$$

i.e. $\widehat{\mu} \le 1 + b/(1-m) = \widehat{\mu}_I$.

For (ii), let $(b_j : j \ge 1)$ denote an arbitrary PMF satisfying $b_1 > 0$ and having the PGF $\widehat{B}(s)$. Then the function $\beta(s) := (1 - \widehat{B}(s))/\widehat{B}'(s)$ has a power series representation $\beta(s) = \sum_{j=0}^{\infty} \beta_j s^j$ converging in the open unit disc, and $\beta_1 = -1 - 2b_2/b_1 < 0$. Then (b_j) is the PMF of the LCL of a subcritical MBP if and only if $\beta_j \ge 0$ for $j \ge 2$, and the offspring mean satisfies

$$1 + 1/\beta_1 \le m < 1.$$

If these conditions are satisfied, then the offspring-number PGF is $f(s) = s + (1-m)\beta(s)$ and

$$\widehat{\mu}_I = 1 + \frac{1}{2}\beta''(1) = 1 + \frac{\widehat{B}''(1)}{2\widehat{\mu}}.$$

Since $\widehat{B}''(1) = \widehat{\sigma}^2 - \widehat{\mu} + \widehat{\mu}^2$, we see that

$$\frac{\widehat{\sigma}^2 - \widehat{\mu} - \widehat{\mu}^2}{2\widehat{\mu}} = \widehat{\mu}_I - \widehat{\mu},$$

and (3.6) follows from (3.5). □

A consequence of Theorem 3.1(i) is that if (Z_n) is a GWP embeddable in a subcritical MBP, then $\mu \le \mu_I$ with equality if and only if the embedding MBP is a generalised linear birth and death process. This contingency implies that the embedded GWP has a linear fractional offspring-number law. We conclude therefore, that (1.3) is an equality only for Example 2.1.

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

4. Approximating μ

Several papers were published around 1970 concerned with computable bounds for quantities associated with the distribution of the extinction time t of the GWP [1, 5, 13, 16]. As observed in Section 1, an upper bound of the form $\mu \le \mu_n + B_n$ can be distilled from bounds derived in [13, 16]. Assume $b < \infty$ and let $n \to \infty$ in the starred inequality on [16, p. 673]. Then, replacing h there with n, we obtain the evaluation of B_n as

$$B_n(S) := \frac{bm^n}{(1-m)f'(f_n)}. \tag{4.1}$$

Pollak [13] derived bounds for c assuming that the third-order factorial moment $f'''(1) < \infty$. His lower bound depends only on b. Written as an upper bound for μ yields his evaluation of B_n as

$$B_n(P) := \frac{bm^{n-1}}{(1-m)[1-bm^{-1}(1-f_n)]}. \tag{4.2}$$

Choosing a desired error $\varepsilon \in (0, 1)$, we achieve the bounds $0 < \mu - \mu_n \le \varepsilon$ by choosing n such that $B_n \le \varepsilon$. Observe that the denominators in (4.1) and (4.2) increase with n. Recalling that $f_n \uparrow 1$, we see for both cases that $B_n \ge bm^{n-1}/(1-m)$. It follows that an absolute minimum value required for n is

$$n \ge n_\ell = 2 + \left\lfloor \frac{\log(b/\varepsilon) + \log(1-m)^{-1}}{\log m^{-1}} \right\rfloor.$$

Observe next that $B_n(S) \le \varepsilon$ if

$$n \ge \frac{\log[b/(\varepsilon(1-m)f'(f_n))]}{\log m^{-1}}.$$

Values of n satisfying this criterion are obtained by choosing a small natural number k, and then

$$n \ge n_{Sk} := \left(1 + \left\lfloor \frac{\log[b/\varepsilon(1-m)f'(f_k)]}{\log m^{-1}} \right\rfloor\right) \vee k.$$

The simplest choice is $k = 1$, in which case $f'(f_1) = f'(p_0)$.

Recall for the Poisson case that we have $b = m^2/2$. In addition, $p_0 = \mathrm{e}^{-m}$ and $f'(p_0) = m\exp[-m(1-\mathrm{e}^{-m})]$. Hence

$$n_{S1} = \left\lfloor \frac{\log(2/\varepsilon) - \log(1-m) + m(1-\mathrm{e}^{-m})}{\log m^{-1}} \right\rfloor.$$

A similar strategy applied to (4.2) yields the condition

$$n \ge n_{Pk} = \left(1 + \left\lfloor \frac{\log[b/\varepsilon(1-m)(1-bm^{-1}(1-f_k))]}{\log m^{-1}} \right\rfloor\right) \vee k.$$

For the Poisson case with $k = 1$, we have

$$n_{P1} = \left\lfloor \frac{\log(2/\varepsilon) - \log(1-m) - \log(1-(m/2)(1-\mathrm{e}^{-m}))}{\log m^{-1}} \right\rfloor.$$

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

TABLE 1. Minimum values of n required to achieve $\mu - \mu_n \le 10^{-4}$ for Poisson offspring-number laws.

m	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
n_ℓ	3	5	7	9	13	18	27	45	102	224	1305
n_{S1}	3	5	7	9	13	19	28	47	107	235	1367
n_{P1}	3	5	7	9	13	18	27	46	105	231	1342

Defining $\phi(u) = -\log(1 - u/2)$ for $u \in (0, 1)$, it is easily checked that $\phi(u) < u$, implying that $n_{S1} \ge n_{P1}$.

Values of n_ℓ, n_{S1}, and n_{P1} for the Poisson case and $\varepsilon = 10^{-4}$ (following [15]) are displayed in Table 1. We see that there is little relative difference between n_{S1} and n_{P1} if $m \le 0.99$, and that $n_{S1} \ge n_{P1}$, with equality if $m \le 0.5$ and significant difference only if m is very close to unity. In addition, neither n_{S1} nor n_{P1} exceed n_ℓ to any marked degree unless m is close to unity.

Very similar outcomes occur for the geometric offspring-number law $f(s) = (1 + m - ms)^{-1}$. Here, we know that $\mu = (1 - m)^{-1}$. Values of n_{S1} and n_{P1} are identical if $m \le 0.7$, differ by at most 2 if $0.8 \le m \le 0.95$, and differ by 12 if $m = 0.99$. There is little difference, 3 at most, between n_{S1} or n_{P1} and n_ℓ if $m \le 0.8$.

Returning to the Poisson case, we mention that Pollak's upper bound $1.2244 = 1/0.8167$ when $m = 0.3$ [13, Table 1] is *less* than the numerical value $\mu = 1.2327$ in [15, Table II]. The latter value is incorrect. We see from Table 1 that $n_{S1} = n_{P1} = 7$, and computation yields $\mu_6 = 1.222\,257$, $\mu_7 = 1.222\,366$, $\mu_8 = 1.222\,399$ and $\mu_9 = 1.222\,409$. Thus, to within 10^{-4}, $\mu = 1.2224$, consistent with Pollak's bound. In addition, this degree of accuracy is achieved by μ_7, as designed, but not by μ_6.

5. Final remarks

We conclude by discussing two matters from [8]. The identity (2.2) is essentially the subcritical version of [8, Theorem 1.2]. This follows from [8, Lemma 2.2], which asserts that a function $\Delta(s)$ defined in the first display on p. 933 is constant-valued. In our notation (see [8, (2.8)]),

$$\Delta(s) = \frac{\mu}{1 - B(s)} - \frac{1}{1 - s}.$$

[An aside: It follows that $\Delta(s) = \sum_{j \ge 1} \Delta_j s^j$, where $\Delta_j = \mu u_j - 1$ and (u_j) is a renewal sequence such that $u_j \to \mu^{-1}$.] Specifically, [8, Lemma 2.2] is asserting that $\Delta(s) \equiv \gamma$. This has consequences for the allowable form of $f(s)$, as we now explain.

Defining $p = \gamma/(1 + \gamma) \in (0, 1)$, [8, Lemma 2.2] asserts that

$$B(s) = 1 - \mu(1 - p)\frac{1 - s}{1 - ps}.$$

But $B(0) = 0$, whence $\mu(1 - p) = 1$, i.e.

$$B(s) = s\frac{1 - p}{1 - ps},$$

the PGF of a shifted geometric law, mentioned in Example 3.2.

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

Conversely, if $p \in (0, 1)$ and $B(s)$ has this form, then it follows from the functional equation that it corresponds to offspring PGFs having the form

$$f(s\,;m) = \frac{1 - m + (m - p)s}{1 - mp - p(1 - m)s}.$$

This specifies the linear-fractional offspring laws, and [8, Theorem 2.1 and Lemma 2.2] are valid for these laws.

It follows from our counter-examples that $\Delta(s)$ is not, in general, constant-valued. The algebraic detail in [8] is correct up to the display following [8, (2.35)]. The next display giving the desired contradiction is not correct. It results from incorrect use of the algebra of asymptotic equivalence. To illustrate the error, let $u(s) \equiv 1$ and $v(s) = s = 1 - (1 - s)$. If ν is any real number then, as $s \to 1$, $u(s) \sim (v(s))^{\nu} = (1 - (1 - s))^{\nu} = 1 - \nu(1 - s)(1 + o(1))$. We emphasise that this is valid for all real ν, but it is invalid to equate the coefficients of $1 - s$ on each side of the asymptotic equivalence, i.e. to adduce the conclusion $\nu = 0$. This is exactly the error committed in going to the second unlabelled display in [8].

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AGRESTI, A. (1974). Bounds on the extinction time distribution of a branching process. *Adv. Appl. Prob.* **6**, 322–335.
[2] ASMUSSEN, S. AND HERING, H. (1983). *Branching Processes*. Birkhäuser, Boston.
[3] ATHREYA, K. B. AND NEY, P. E. (1973). *Branching Processes*. Springer, Berlin.
[4] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.
[5] HEATHCOTE, C. R. AND SENETA, E. (1966). Inequalities for branching processes. *J. Appl. Prob.* **3**, 261–267. (Correction: *J. Appl. Prob.* **4**, 215.)
[6] HEATHCOTE, C. R., SENETA, E. AND VERE-JONES, D. (1967). A refinement of two theorems in the theory of branching processes. *Theory Prob. Appl.* **12**, 297–301.
[7] IMOMOV, A. A. (2025). On the Kolmogorov constant explicit form in the theory of stochastic continuous-time Markov branching systems. *J. Theoret. Prob.* **38**, 32.
[8] IMOMOV, A. A. AND MURTAZAEV, M. (2024). On the Kolmogorov constant explicit form in the theory of discrete-time stochastic branching systems. *J. Appl. Prob.* **61**, 927–941.
[9] IMOMOV, A. A. AND NAZAROV, Z. (2023). Limit theorems for the positive recurrent Q-process. In *Information Technologies and Mathematical Modelling: Queueing Theory and Applications*, eds A. Dudin, A. Nazarov and A. Moiseev. Springer, Cham, pp. 1–15.
[10] JOFFE, A. (1967). On the Galton–Watson process with mean less than one. *Ann. Math. Statist.* **38**, 264–266.
[11] NAGAEV, A. V. AND BADALBAEV, I. S. (1967). A refinement of certain theorems on branching random processes. *Litovsk. Mat. Sb.* **7**, 129–136 [in Russian].
[12] PAKES, A. G. (2025). On the stationary measure for Markov branching processes. *Mathematics* **13**, 1082.
[13] POLLAK, E. (1969). Bounds for certain branching processes. *J. Appl. Prob.* **6**, 201–204.
[14] SAGITOV, S. AND LINDO, A. (2016). A special family of Galton–Watson processes with explosions. In *Branching Processes and Their Applications* (Lect. Notes Statist. **219**), eds I. M. del Puerto et al. Springer, Cham, pp. 237–254.
[15] SENETA, E. (1967). On the transient behaviour of a Poisson branching process. *J. Austral. Math. Soc.* **7**, 465–480.

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press

[16] SENETA, E. (1968). On asymptotic properties of sub-critical branching processes. *J. Austral. Math. Soc.* **8**, 671–682.
[17] TCHORBADJIEFF, A., MAYSTER, P. AND PAKES, A. G. (2024). On subcritical Markov branching processes with a specified limiting conditional law. *Stoch. Quality Control* **39**, 9–23.
[18] ZOLOTAREV, V. M. (1957). More exact statements of several theorems in the theory of branching processes. *Theory Prob. Appl.* **11**, 245–253.

https://doi.org/10.1017/jpr.2026.10094 Published online by Cambridge University Press