The Annals of Applied Probability
1994, Vol. 4, No. 3, 779–790

LARGE DEVIATION RATES FOR BRANCHING PROCESSES—I. SINGLE TYPE CASE[1]

BY K. B. ATHREYA

Iowa State University

Let $\{Z_n\}_0^\infty$ be a Galton–Watson branching process with offspring distribution $\{p_j\}_0^\infty$. We assume throughout that $p_0 = 0$, $p_j \neq 1$ for any $j \geq 1$ and $1 < m = \sum j p_j < \infty$. Let $W_n = Z_n m^{-n}$ and $W = \lim_n W_n$. In this paper we study the rates of convergence to zero as $n \to \infty$ of

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon\right), \qquad P(|W_n - W| > \varepsilon),$$
$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \middle| W \geq a\right)$$

for $\varepsilon > 0$ and $a > 0$ under various moment conditions on $\{p_j\}$. It is shown that the rate for the first one is geometric if $p_1 > 0$ and supergeometric if $p_1 = 0$, while the rates for the other two are always supergeometric under a finite moment generating function hypothesis.

1. Introduction and summary of results. Let $\{Z_n\}_0^\infty$ be a Galton–Watson brancing process with offspring distribution $\{p_j\}_0^\infty$. We assume throughout that $p_0 = 0$, $p_j \neq 1$ for any $j \geq 1$ and $1 < m = \sum j p_j < \infty$.

It is known [1] that $Z_{n+1} Z_n^{-1} \to m$ w.p.1. and that $W_n \equiv Z_n m^{-n}$ converges to a r.v. W w.p.1. The goal of this paper is to study the rates of decay as $n \to \infty$ of

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \epsilon\right) \quad \text{and} \quad P(|W_n - W| > \varepsilon).$$

Besides being of some interest in its own right, this seems to be of some importance in algorithmic tree structures in computer science (see [6] and [8]). In [6] it is mentioned that the execution of a canonical algorithm for evaluating uniform AND/OR trees in some probabilistic models can be viewed as a branching process. The probability that the running time of this algorithm deviates from its expected value corresponds to the tail probabilities and large deviations associated with branching processes. Also, from a statistical inference point of view the first quantity above is of some interest

Received February 1993; revised October 1993.
[1] Supported in part by NSF Grant DMS-92-04938.
AMS 1991 *subject classifications*. 60J80, 60F10.
Key words and phrases. Large deviation, branching processes.

since Z_{n+1}/Z_n is a reasonable estimate of m. We establish the following results:

THEOREM 1. *Assume* $p_1 > 0$ *and* $E(exp(\theta_0 Z_1) \mid Z_0 = 1) < \infty$ *for some* $\theta_0 > 0$. *Then for all* $\varepsilon > 0$,

$$\lim_{n \to \infty} \frac{1}{p_1^n} P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \middle| Z_0 = 1\right) = \sum_k \phi(k, \varepsilon) q_k < \infty, \tag{1.1}$$

where

$$\phi(k, \varepsilon) = P\big(|\overline{X}_k - m| > \varepsilon\big), \tag{1.2}$$

$\overline{X}_k$ *being the mean* $(1/k)\sum_1^k X_i$ *of* k *i.i.d. r.v.* $\{X_i\}$ *with distribution* $\{p_j\}$, *and* $\{q_k\}$ *is defined via the generating function* $Q(s) = \sum_1^\infty q_k s^k$, $0 \le s < 1$, *being the unique solution of the functional equation*

$$Q(f(s)) = p_1 Q(s), \qquad f(s) = \sum_1^\infty p_j s^j, \qquad 0 \le s < 1, \tag{1.3}$$

subject to

$$Q(0) = 0, \qquad Q(s) < \infty \quad \text{for } 0 \le s < 1, \qquad Q(1) = \infty. \tag{1.4}$$

The next theorem and corollary establish (1.1) under conditions weaker than $E(e^{\theta_0 Z_1} \mid Z_0 = 1)$ being finite for some $\theta_0 > 0$. This is the main result of this paper.

THEOREM 2. *Assume that* $p_1 > 0$ *and that there exist constants* C_ε *and* $r > 0$ *such that* $p_1 m^r > 1$ *and* $\phi(k, \varepsilon) \le C_\varepsilon / k^r$ *for all* k, *where* $\phi(k, \varepsilon)$ *is as in* (2). *Then* (1) *holds.*

COROLLARY 1. *Assume* $p_1 > 0$ *and* $E(Z_1^{2r+\delta} \mid Z_0 = 1) < \infty$ *for some* $r \ge 1$ *and* $\delta > 0$ *such that* $p_1 m^r > 1$. *Then* (1) *holds.*

The next result shows that the rate of decay of $P(|Z_{n+1}/Z_n - m| > \varepsilon)$ when $p_1 = 0$ is *supergeometric.*

THEOREM 3. *Assume* $p_1 = 0$ *and let* $k = \inf\{j\colon j \ge 2,\ p_j \ne 0\}$. *Assume* $E(exp(\theta_0 Z_1) \mid Z_0 = 1) < \infty$ *for some* $\theta_0 > 0$. *Then, for all* $\varepsilon > 0$, *there exist constants* $0 < C(\varepsilon) < \infty$ *and* $0 < \lambda(\varepsilon) < 1$ *such that*

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon\right) \le C(\varepsilon)(\lambda(\varepsilon))^{k^n}.$$

The next result is needed in Theorem 6, which gives the rate of decay of $P(|W_n - W| > \varepsilon)$, and is also of interest in computer science (see [6] and [8]).

THEOREM 4. *Let $E(\exp(\theta_0 Z_1) \mid Z_0 = 1) < \infty$ for some $\theta_0 > 0$. Then $\exists\, \theta_1 > 0$ such that*

$$C_1 = \sup_n E(\exp(\theta_1 W_n)) < \infty.$$

The next result shows that the decay rate of $P(|W_n - W| > \varepsilon)$ is supergeometric.

THEOREM 5. *Let $E(\exp(\theta_0 Z_1) \mid Z_0 = 1) < \infty$ for some $\theta_0 > 0$. Then there exist constants C_3 and $\lambda > 0$ such that*

$$P(|W - W_n| \ge \varepsilon) \le C_3 \exp\left(-\lambda \varepsilon^{2/3} (m^{1/3})^n\right).$$

The next result shows that, conditioned on W being positive, the rate of decay of $P(|Z_{n+1}/Z_n - m| > \varepsilon)$ is supergeometric. A heuristic argument for this is that $W \ge a$ means that $Z_n \ge a' m^n$ for large n where $0 < a' < a$ and hence if $\phi(k, \varepsilon)$ decays exponentially with k, $P(|Z_{n+1}/Z_n - m| > \varepsilon | Z_n \ge am^n)$ should be of the order $\exp(-ca'm^n)$, where c is a constant.

THEOREM 6. *Let $E(\exp(\theta_0 Z_1) \mid Z_0 = 1) < \infty$ for some $\theta_0 > 0$. Then there exist constants C_4 and $\lambda > 0$ such that for all $\varepsilon > 0$, $a > 0$, we can find $0 < I(\varepsilon) < \infty$ such that*

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \middle| W \ge a\right) \le C_4 \exp(-a\gamma I(\varepsilon) m^n) + C_3 \exp\left(-\lambda (a(1-\gamma))^{2/3} (m^{1/3})^n\right)$$

for every $0 < \gamma < 1$ and hence (for $\gamma = \frac{1}{2}$) $\le$ (const.)$\exp(-\lambda(a/2)^{2/3}(m^{1/3})^n)$.

The generalization of the above results to the multitype case is contained in Vidyashankar's thesis [10].

The next section is devoted to some preliminary results. The proofs of the above theorems are in the last section.

2. Some preliminary results. Let $f(s) = \sum_0^\infty p_j s^j$, where $\{p_j\}$ is a probability distribution and $0 \le s \le 1$. Let $f_2(s) = f(f(s))$ and $f_n(s) = f((\cdots(s))\cdots)$ be the nth iterate of f for $n \ge 1$. It is well known [1] that if $f(s) = E(s^{Z_1} \mid Z_0 = 1)$, where $\{Z_n\}_0^\infty$ is a Galton–Watson branching process with offspring distribution $\{p_j\}$, then $f_n(s) = E(s^{Z_n} \mid Z_0 = 1)$. In this section we study the rate of convergence of $f_n(s)$ and its inverse $g_n(s)$ as $n \to \infty$.

PROPOSITION 1. *For $0 \le s < 1$, $f_n(s) \to q$, where q is the smallest root in $[0, 1]$ of $f(s) = s$. This $q < 1$ iff $m = f'(1-) = \sum j p_j > 1$.*

PROPOSITION 2. *Let $p_0 = 0$ and $p_1 > 0$. Then $q = 0$ and there exists $0 \le q_k < \infty$ such that*

$$\lim_{n \to \infty} \frac{f_n(s)}{p_1^n} = \sum_1^\infty q_k s^k \equiv Q(s) < \infty \quad \text{for } 0 \le s < 1 \tag{2.1}$$

with $q_1 = 1$. *Further*, $Q(s)$ *is the unique solution of the functional equation*

$$(2.2)\qquad \begin{aligned} Q(f(s)) &= p_1 Q(s), \qquad 0 \le s < 1, \\ Q(0) &= 0, \qquad Q(s) \neq 0 \quad \text{for } x \neq q. \end{aligned}$$

Consequently, for all $1 \le r, j < \infty$,

$$(2.3)\qquad \lim_n \frac{P(Z_n = j \mid Z_0 = r)}{p_1^{nr}} = q_{rj} \quad \text{exists}$$

where q_{rj} *satisfies* $\Sigma_1^\infty q_{rj} s^j = (\Sigma_1^\infty q_k s^k)^r$ *for* $0 \le s < 1$.

For the proofs of Propositions 1 and 2, see [1]. Assertion (2.3) follows from (2.1) since $E(s^{Z_n} \mid Z_1 = r) = (E(s^{Z_n} \mid Z_0 = 1))^r = (f_n(s))^r$. The next proposition shows that if p_0 and p_1 are both zero, then $f_n(s)$ decays at a supergeometric rate. This is referred to as the Böttcher case in the literature (see [4]). In this case we have the following proposition. Although most of this proposition is known (see [5] and [7]), we supply the proof below for completeness.

PROPOSITION 3. *Let* $p_0 = 0 = p_1$. *Let* $k = \inf\{j\colon j \ge 1,\ p_j \neq 0\}$. *Then*

$$(2.4)\qquad f_n(s) = s^{k^n} p_k^{\Sigma_0^{n-1} k^j} (R_n(s))^{k^n},$$

where $\lim_n R_n(s) = R(s)$ *exists uniformly in* $[0,1]$, *with* $R(0) = 1$, $R(1) < \infty$. *Further*,

$$(2.5)\qquad \left(\frac{R_n(s)}{R(s)}\right)^{k^n} \to 1 \quad \text{for } 0 \le s \le 1,$$

and hence

$$(2.6)\qquad f_n(s) \sim \left(p_k^{-1/(k-1)}\right)\left(p_k^{1/(k-1)} s R(s)\right)^{k^n}.$$

Also $R(\cdot)$ *satisfies the functional equation*

$$(2.7)\qquad f(s) R(f(s)) = p_k (s R(s))^k$$

and is the unique solution of (2.7), *subject to the condition* $R(0) > 0$ *and* $R(\cdot)$ *is continuous in* $[0,1]$.

PROOF. Since $f(s) = \Sigma_k^\infty p_j s^j$ we may write $f(s) = p_k s^k (1 + \gamma g(s))$, where $\gamma = (1 - p_k)/p_k$, $g(s) = \Sigma_{j=k+1}^\infty p_j s^{j-k}/(1 - p_k)$. Thus $0 < \gamma < \infty$ and $g(\cdot)$ is a probability generating function.

By definition,

$$f_{n+1}(s) = f(f_n(s)) = p_k (f_n(s))^k (1 + \gamma g(f_n(s))).$$

Let $h_n(s) \equiv (f_n(s))^{1/k^n}$. Then $h_n(\cdot)$ satisfies

$$h_{n+1}(s) = \left(p_k(1 + \gamma g(f_n(s)))\right)^{1/k^{(n+1)}} h_n(s),$$

iterating which yields

$$h_n(s) = s\prod_{j=0}^{n-1}\Big(p_k\big(1+\gamma g(f_j(s))\big)\Big)^{1/k^{(j+1)}}$$

and hence

$$f_n(s) = s^{k^n} p_k^{1+k+\cdots+k^{n-1}}(R_n(s))^{k^n}, \tag{2.8}$$

where $R_n(s) = \prod_{j=0}^{n-1}(1+\gamma g(f_j(s)))^{1/k^{j+1}}$. Since $k>1$ and $g(1)=1$ we have, for $0 \le s \le 1$,

$$0 \le \log R_n(s) \le \left(\sum_0^\infty \frac{1}{k^{j+1}}\right)(\log(1+\gamma)) < \infty,$$

and hence $R_n(s)$ converges uniformly on $[0,1]$. This proves (2.4).

If $k=1$, (2.8) yields $f_n(s) = sp_1^n R_n(s)$ and the convergence of $R_n(s)$ is established by showing that for $0 \le s < 1$, $\gamma g(s) \le \gamma s$ and $\Sigma_j(f_j(s)) < \infty$, thus yielding a proof of (2.2).

Next, letting

$$R(s) = \prod_{j=0}^{\infty}\big(1+\gamma(f_j(s))\big)^{1/k^{j+1}}, \tag{2.9}$$

we note that

$$\begin{aligned} 0 \le k^n \log\left(\frac{R(s)}{R_n(s)}\right) &= k^n \sum_{j=n}^{\infty} \frac{1}{k^{j+1}}\log\big(1+\gamma g(f_j(s))\big)\\ &\le \sum_{r=0}^{\infty}\frac{1}{k^{r+1}}\log(1+\gamma g(f_{n+r}(s))). \end{aligned}$$

For $0 \le s < 1$, $f_j(s) \to 0$. Also, $g(0)=0$ and $0 \le 1+\gamma g(f_{n+r}(s)) \le (1+\gamma)$ and $\Sigma_0^\infty 1/k^{r+1} < \infty$ for $k>1$. Thus by the dominated convergence theorem,

$$\lim_n k^n \log\frac{R(s)}{R_n(s)} = 0.$$

For $s=1$, $k^n\log(R(s)/R_n(s)) = \Sigma_0^\infty k^{-(r+1)}(\log(1+\gamma))$ and thus is independent of n. Next it is easy to see that $R(\cdot)$ defined in (2.9) satisfies (2.7). To prove uniqueness, if $\tilde{R}(\cdot)$ is another solution of (2.7) such that $\tilde{R}(\cdot)$ is continuous at 0 and $\tilde{R}(0) \ne 0$, then $r(s) = (\tilde{R}(s)/R(s))$ for $0 \le s \le 1$ satisfies

$$r(f(s)) = (r(s))^k, \tag{2.10}$$

so that for $0 \le s < 1$, $r(s) = (r(f(s)))^{1/k} = (r(f_n(s)))^{1/k^n}$. By continuity at 0, $r(f_n(s)) \to r(0) > 0$ [since $0 \le s > 1 \Rightarrow f_n(s) \to 0$] and hence for $k>1$, $(r(f_n(s)))^{1/k^n} \to 1 \Rightarrow r(s) = 1$ for $0 \le s < 1$.

Since $r(s)$ is continuous at 1, $r(1)=1$ and so $R(s) = \tilde{R}(s)$ for $0 \le s \le 1$. □

We shall have occasion to use the inverse function $g(s)$ of $f(s)$ defined by

$$f(g(s)) = s \quad \text{for } 0 \le s < \infty.$$

For $0 \le s \le 1$, $g(s)$ is well defined and $g(s) \ge s$. Also since $f(s) \ge s$ for $s \ge 1$, $g(s)$ is well defined for $1 \le s \le f(s_0)$ and $g(s) \le s$. Thus, the iterates g_n of g are such that they are nondecreasing in $[0, 1]$ and nonincreasing in $[1, f(s_0)]$. The next proposition shows that the rate of convergence of $g_n(\cdot)$ is geometric.

PROPOSITION 4. *Let $f(s_0) < \infty$ for some $s_0 > 1$. Then for $1 \le s \le f(s_0)$, $g_n(s) \downarrow 1$ and*

$$\tilde{Q}_n(s) \equiv m^n(g_n(s) - 1) \downarrow \tilde{Q}(s), \tag{2.11}$$

where $\tilde{Q}(\cdot)$ is the unique solution of the functional equation

$$\tilde{Q}(f(s)) = m\tilde{Q}(s) \quad \text{for } 1 \le s \le f(s_0) \tag{2.12}$$

subject to

$$\begin{gathered} 0 < \tilde{Q}(s) < \infty \quad \text{for } 1 < s \le f(s_0), \\ \tilde{Q}(1) = 0, \qquad \tilde{Q}'(1) = 1. \end{gathered} \tag{2.13}$$

The proof is similar to that of Theorems 1 and 2 of [1, page 40] and is omitted.

It is also known (see [1], page 42) that for $0 \le s \le 1$, $g_n(s) \uparrow 1$ and $m^n(1 - g_n(s))$ increases to a finite positive limit iff $E(Z_1 \log Z_1 \mid Z_0 = 1) < \infty$.

3. Proofs of Theorems 1–4.

PROOF OF THEOREM 1. This theorem is in [2].

PROOF OF THEOREM 2. By conditioning on $\mathscr{F}_n = \sigma(Z_0, Z_1, \dots, Z_n)$,

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon\right) = \sum_k \phi(k, \varepsilon) P(Z_n = k).$$

By assumption,

$$h_n(k) \equiv \frac{\phi(k, \varepsilon) P(Z_n = k)}{p_1^n} \le \frac{C_\varepsilon}{k^r} \frac{P(Z_n = k)}{p_1^n} \equiv h'_n(k), \quad \text{say.}$$

By (2.3),

$$\begin{aligned} h_n(k) &\to q_k \phi(k, \varepsilon) \equiv h(k), \quad \text{say}, \\ h'_n(k) &\to C_\varepsilon \frac{q_k}{k^r}. \end{aligned}$$

If we show that

$$\sum_k h'_n(k) \to \sum_k C_\varepsilon \frac{q_k}{k^r} < \infty,$$

then by a slight modification of the Lebesgue dominated convergence theorem (see [9], page 270), we get that

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \middle| Z_0 = 1\right) = \sum_k h_n(k) \to \sum_k h(k) < \infty.$$

However,

$$\sum \frac{1}{k^r} \frac{P(Z_n = k)}{p_1^n} = \frac{E(Z_n^{-r})}{p_1^n}.$$

For any nonnegative r.v. X and $0 < p < \infty$,

$$\begin{aligned} EX^{-p} &= E\left(\frac{1}{\Gamma(p)} \int_0^\infty e^{-tX} t^{p-1}\, dt\right) \\ &= \frac{1}{\Gamma(p)} \int_0^\infty E(e^{-tX}) t^{p-1}\, dt. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{E(Z_n^{-r})}{p_1^n} &= \frac{1}{\Gamma(p)} \int_0^\infty \frac{f_n(e^{-t})}{p_1^n} t^{r-1}\, dt \\ &= \frac{1}{\Gamma(p)} \int_0^1 \frac{f_n(s)}{p_1^n} k(s)\, ds, \end{aligned}$$

where

$$k(s) = \frac{|\log s|^{r-1}}{s}.$$

Since $(f_n(s))/p_1^n \uparrow Q(s)$, by the monotone convergence theorem

$$\Gamma(p) \frac{E(Z_n^{-r})}{p_1^n} \uparrow \int_0^1 Q(s) k(s)\, ds.$$

So the proof will be complete if we show $\int_0^1 Q(s)k(s)\, ds < \infty$. Let $g(s) = f^{-1}(s)$ be the inverse of f defined by $f(g(s)) = s$ for $0 \le s \le 1$. It follows that $g_n(s)$, the nth iterate of g, satisfies $f_n(g_n(s)) = s$, $g_{n+1}(s) \ge g_n(s)$ and for $0 < s < 1$, $g_n(s) \uparrow 1$ and $f_n(s) \downarrow 0$. Fix $0 < t_0 < 1$. Then $t_n = g_n(t_0) \uparrow 1$. Also since Q satisfies (4),

$$\begin{aligned} I_n &= \int_{t_n}^{t_{n+1}} Q(s)k(s)\, ds = \int_{t_n}^{t_{n+1}} \frac{Q(f(s))}{p_1} k(s)\, ds \\ &= \int_{t_{n-1}}^{t_n} Q(u) \frac{k(g(u))g'(u)\, du}{p_1} = \int_{t_{n-1}}^{t_n} Q(u)k(u) \left(\frac{k(g(u))g'(u)}{p_1 k(u)}\right) du. \end{aligned}$$

Since $g'(u) = 1/(f'(g(u)))$ and $|\log s|/(1-s) \to 1$ as $s \uparrow 1$, $(k(g(u))/(p_1 k(u)))g'(u) \to 1/(p_1 m^r)$, where $m = f'(1)$. Thus, if $p_1 m^r > 1$, then for any $0<\lambda<(p_1 m^r)^{-1}<1$ there exists an n_0 such that for $u \geq g_{n_0}(t_0)$, $k(g(u))g'(u)/(p_1 k(u)) < \lambda$. Thus, for $n \geq n_0 + 2$,

$$I_n \leq \lambda I_{n-1} \Rightarrow \sum_{n=n_0+2}^{\infty} I_n \leq I_{n_0+1} \sum_1^{\infty} \lambda^j < \infty$$

$$\Rightarrow \int_{t_{n_0}}^{1} Q(s)k(s)\,ds < \infty. \qquad \square$$

Proof of Corollary 1. Since $E(Z_1^{2r+\delta} \mid Z_0 = 1) < \infty$ for $r \geq 1$, $\delta > 0$, we have that

$$C_r = \sup_k E\left|\sqrt{k}\,\frac{(\bar{X}_k - m)}{\sigma}\right|^{2r}$$

is finite and so by Markov's inequality, $\phi(k,\varepsilon) \leq (1/\varepsilon^{2r})C_r/k^r$. $\square$

Proof of Theorem 3. Since $E(\exp(\theta_0 Z_1) \mid Z_0 = 1) < \infty$ for some $\theta_0 > 0$, for $\varepsilon > 0$, there exist C_ε and $0 < \rho_\varepsilon < 1$ such that $\phi(k,\varepsilon) \leq C_\varepsilon \rho_\varepsilon^k$ for all k. Thus,

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \middle| Z_0 = 1\right) \leq C_\varepsilon \Sigma \rho_\varepsilon^k P(Z_n = k \mid Z_0 = 1)$$

$$= C_\varepsilon f_n(\rho_\varepsilon).$$

Now use (2.6). $\square$

Proof of Theorem 4. Although it is possible to deduce this theorem from the work of [3] we give the following proof due to its brevity and for the sake of completeness. By hypothesis $K \equiv f(s_0) < \infty$ for $s_0 = e^{\theta_0}$. So $f_2(s) \leq K$ if $0 \leq f(s) \leq s_0$, that is, if $0 \leq s \leq g(s_0)$. Similarly, $f_3(s) \leq K$ if $0 \leq f(s) \leq g(s_0)$, that is, if $0 \leq s \leq g_2(s_0)$. More generally,

$$f_n(s) \leq K \quad \text{if } 0 \leq s \leq g_{n-1}(s_0).$$

Now, since $W_n = Z_n m^{-n}$, $E(\exp(\theta W_n) \mid Z_0 = 1) = f_n(\exp(\theta/m^n))$. Thus $E(\exp(\theta W_n) \mid Z_0 = 1) \leq K$ if $\theta \leq m^n \log g_{n-1}(s_0)$. Since $g_n(s_0) \downarrow 1$, $\log g_n(s_0) \sim (g_n(s_0) - 1)$. By Proposition 4, $f(s_0) < \infty$ for $1 < s_0$ implies $m^n \log g_{n-1}(s_0) \to m\tilde{Q}(s_0)$, which is positive and finite. Since $g_n(s_0) > 1$ for all $n \geq 1$, we can choose

$$\theta_1 = \inf_n m^n \log g_{n-1}(s_0) \quad \text{and} \quad C_1 = K. \qquad \square$$

Proof of Theorem 5. First we need an estimate. Let $\phi(\theta_1) = E(\exp(\theta_1 W)) < \infty$ for all $\theta \leq \theta_1$. So, if $\{W^{(i)}\}_1^\infty$ are i.i.d. copies of W and $S_k = \Sigma_1^k(W^{(i)} - 1)$, then for $\theta \leq \theta_1$,

$$E\big(\exp\big(\theta\big(S_k/\sqrt{k}\big)\big)\big) = \left(\phi\left(\frac{\theta}{\sqrt{k}}\right)e^{-\theta/\sqrt{k}}\right)^k = \left(1 + \frac{1}{k}\,\frac{\big(\phi(\theta/\sqrt{k})e^{-\theta/\sqrt{k}} - 1\big)}{(\theta^2/k)}\,\theta^2\right)^k.$$

However, $\sup_{|u| \le 1} |(\phi(u)e^{-u} - 1)/u^2| = c < \infty$ since $\lim_{u \to 0}(\phi(u)e^{-u} - 1)/u^2 = \frac{1}{2}(\mathrm{Var}(W)) < \infty$. If $\theta_2 = \min(\theta_1, 1)$, then $\sup_{|\theta| \le \theta_2}(\phi(\theta/\sqrt{k})e^{-\theta/\sqrt{k}})^k \le e^c = C_2$, say. We have used the fact that for $x > 0$, $(1 + x/k)^k \le e^x$.

Now we proceed with the proof the Theorem 5. We begin by noting that (see Theorem 2 on page 55 of [1])

$$W - W_n = \lim_{k \to \infty} (W_{n+k} - W_n) = \frac{1}{m^n} \sum_{j=1}^{Z_n} (W^{(j)} - 1),$$

where $W^{(j)}$ is the limit r.v. in the line of descent initiated by the jth parent of the nth generation. By conditional independence,

$$P((W - W_n) > \varepsilon | Z_0, Z_1, \dots, Z_n) = \psi(Z_n, m^n\varepsilon),$$

where $\psi(k, \eta) = P(S_k \ge \eta)$. However,

$$P(S_k \ge \eta) = P\left(\frac{S_k}{\sqrt{k}} \ge \frac{\eta}{\sqrt{k}}\right) \le \exp\left(-\frac{\theta_2 \eta}{\sqrt{k}}\right) C_2 \quad \text{(by our estimate)}.$$

Thus,

$$P(W - W_n > \varepsilon) = E\psi(Z_n, m^n\varepsilon) \le C_2 E\left(\exp\left(-\frac{\theta_2 m^n \varepsilon}{\sqrt{Z_n}}\right)\right) = C_2 E\left(\exp\left(-\theta_2 \varepsilon m^{n/2} \frac{1}{\sqrt{W_n}}\right)\right).$$

For $\lambda > 0$,

$$E\big(\exp(-\lambda(1/\sqrt{W_n}))\big) = \lambda \int_0^\infty e^{-\lambda u} P\left(\frac{1}{\sqrt{W_n}} \le u\right) du = \lambda \int_0^\infty e^{-\lambda u} P\left(W_n \ge \frac{1}{u^2}\right) du \le \lambda C_1 \int_0^\infty e^{-\lambda u} \exp\left(-\frac{\theta_1}{u^2}\right) du \quad \text{(by Theorem 4)} = C_1 \int_0^\infty e^{-t} \exp\left(-\frac{\theta_1 \lambda^2}{t^2}\right) dt.$$

Thus,

$$P(W - W_n \ge \varepsilon) \le C_2 C_1 \int_0^\infty e^{-t} \exp\left(-\frac{\theta_1 \lambda_n^2}{t^2}\right) dt,$$

where $\lambda_n = \theta_2 \varepsilon m^{n/2}$. However, for $\lambda > 0$,

$$I(\lambda) \equiv \int_0^\infty e^{-t} e^{-\lambda^2/t^2}\,dt = \int_0^{k(\lambda)} + \int_{k(\lambda)}^\infty \le \exp\left(-\frac{\lambda^2}{k^2(\lambda)}\right) + e^{-k(\lambda)}.$$

Choose $k(\lambda) = \lambda^{2/3}$. Then $I(\lambda) \le 2\exp(-\lambda^{2/3})$. Thus

$$P(W - W_n \ge \varepsilon) \le 2C_2C_1 \exp\left(\sqrt{\theta_1}\,\theta_2 \varepsilon m^{n/2}\right)^{2/3} = C_3 \exp\left(-\lambda\left(m^{1/3}\right)^n \varepsilon^{2/3}\right),$$

where $C_3 = 2C_1C_2$, $\lambda = (\sqrt{\theta_1}\ \theta_2)^{2/3}$. Similar arguments hold for $P(W_n - W \ge \varepsilon)$. □

PROOF OF THEOREM 6.

$$\begin{aligned}
P&\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \,\middle|\, W \ge a\right)\\
&= P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon, W \ge a\right)\frac{1}{P(W \ge a)}\\
&= p_a\left(P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon, W_n \le a\gamma, W \ge a\right)\right.\\
&\qquad \left.+P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon, W_n \ge a\gamma, W \ge a\right)\right)\\
&= p_a(a_{n1} + a_{n2}) \quad \text{say,}
\end{aligned}$$

where $0 < \gamma < 1$ and $p_a = 1/(P(W \ge a))$. Clearly,

$$\begin{aligned}
a_{n2} &\le P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon, W_n \ge a\gamma\right)\\
&= C_4 \exp(-I(\varepsilon) a\gamma m^n),
\end{aligned}$$

where C_4 and $I(\varepsilon)$ are such that $P(|\overline{X}_k| \ge \varepsilon) \le C_4 e^{-kI(\varepsilon)}$ and $\overline{X}_k = (X_1 + \cdots + X_k)/k$, $\{X_i\}$ being i.i.d. as $Z_1 - m$ with $Z_0 = 1$. [Such a C_4 and $I(\varepsilon)$ exist by Chernoff type bounds since $E(\exp(\theta_1 Z_1)) < \infty$ for $\theta_1 > 0$.] Now

$$\begin{aligned}
a_{n1} &\le P(W - W_n \ge a(1-\gamma))\\
&\le C_4 \exp\left(-\lambda(a(1-\gamma))^{2/3}\left(m^{1/3}\right)^n\right) \quad \text{(by Theorem 5)},
\end{aligned}$$

$$\begin{aligned}
P&\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \,\middle|\, W \ge a\right)\\
&\le p_a\left(C_4 \exp(-I(\varepsilon) a\gamma m^n) + C_3 \exp\left(-\lambda(a(1-\gamma))^{2/3}\left(m^{1/3}\right)^n\right)\right).
\end{aligned}$$

Since the only condition on γ is that $0 < \gamma < 1$ and the second term goes to zero slower than the first term, we can say that there exist C_5 and λ (C_5 may depend on γ) such that

$$P\left(\left|\frac{Z_{n+1}}{Z_n} - m\right| > \varepsilon \middle| W \geq a\right) \leq C_5 \exp\left(-\lambda(a(1-\gamma))^{2/3}(m^{1/3})^n\right). \qquad \square$$

REMARK. Let $I_i(\cdot)$, $i = 1, 2$, be two functions from $(0, \infty)$ to $(0, \infty)$ such that, for each $x > 0$, if

$$\phi_{ni}(x) = P\left(\frac{1}{n}\sum_{j=1}^{n} X_{ij} - m > x\right), \qquad i = 1, 2,$$

where $\{X_{1j}\}_{j=1}^{\infty}$ are i.i.d. with distribution $\{p_j\}^{\infty}$, and $\{X_{2j}\}_{j=1}^{\infty}$ are i.i.d. with distribution the same as W, then $n^{-1}\log \phi_{ni}(x) \to -I_i(x)$, $i = 1, 2$. A sufficient condition for this is that $\sum_j p_j s^j < \infty$ for all $s > 0$.

By conditioning on $\mathscr{F}_n \equiv \sigma(Z_0, Z_1, \dots, Z_n)$, we see that for any $\varepsilon > 0$,

$$P\left(\frac{Z_{n+1}}{Z_n} - m > \varepsilon \middle| \mathscr{F}_n\right) = \phi_{Z_n 1}(\varepsilon)$$

and

$$P(W - W_n > \varepsilon \mid \mathscr{F}_n) = \phi_{Z_n 2}(\varepsilon).$$

Since $Z_n/m^n \to W$ w.p.1, we see that an almost sure large deviation result holds: On the set $\{W > 0\}$,

$$\frac{1}{m^n}\log P\left(\left(\frac{Z_{n+1}}{Z_n} - m\right) > \varepsilon \middle| \mathscr{F}_n\right) \to -WI_1(\varepsilon)$$

and

$$\frac{1}{m^n}\log P((W - W_n) > \varepsilon \mid \mathscr{F}_n) \to -WI_2(\varepsilon).$$

Thus on almost all sample paths in $\{W > 0\}$, the large deviation probabilities decay at the same supergeometric rate and yet overall unconditionally the rate is only geometric as asserted by Theorem 2. This interesting sharp contrast between conditional and unconditional large deviation rates was pointed out to us by the referee.

Acknowledgments. The author would like to thank the referee for a thorough reading of the paper and for making many important suggestions to improve the paper as well as for pointing out some very relevant references. The author would also like to thank Professor Yali Amit for raising the questions that led to Theorem 6 and also for referring us to the work of Miller and O'Sullivan [8]. The author also extends his thanks to his student A. Vidyashankar for several useful discussions.

REFERENCES

[1] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.

[2] ATHREYA, K. B. and VIDYASHANKAR, A. (1993). *Large Deviation Results for Branching Processes. A Festschrift in Honour of Gopinath Kallianpur*. Springer, Berlin.

[3] BIGGINS, J. D. and SHANBHAG, D. N. (1981). Some divisibility problems in branching processes. *Math. Proc. Cambridge Philos. Soc.* **90** 321–330.

[4] BINGHAM, N. H. (1988). On the limit of a supercritical branching process. *J. Appl. Probab.* **25A** 215–228.

[5] DUBUC, S. (1971). Problémes relatifs á l'iteration des fonctions suggérées par les processus en cascade. *Ann. Inst. Fourier* **21** 151–171.

[6] KARP, R. and ZHANG, Y. (1983). Tail probabilities for finite supercritical branching processes, Technical Report, Dept. Computer Science and Engineering, Southern Methodist University, Dallas, TX.

[7] KUCZUMA, M. (1968). *Functional Equations in a Single Variable*. PWN, Warsaw.

[8] MILLER, H. I. and O'SULLIVAN, J. A. (1992). Entropies and combinatorics of random branching processes and context free languages. *IEEE Trans. Inform. Theory* **38** 1292–1311.

[9] ROYDEN, H. L. (1987). *Real Analyis*, 3rd ed. Macmillan, New York.

[10] VIDYASSHANKAR, A. (1993). Large deviation results for branching processes in fixed and random environments. Thesis, Dept. Mathematics, Iowa State Univ.

DEPARTMENT OF MATHEMATICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011