J. Appl. Prob. **55,** 450–458 (2018)
*doi:*10.1017/jpr.2018.29

SELF-NORMALIZED LARGE DEVIATION FOR SUPERCRITICAL BRANCHING PROCESSES

WEIJUAN CHU,* *Hohai University*

Abstract

We consider a supercritical branching process $(Z_n,\ n \geq 0)$ with offspring distribution $(p_k,\ k \geq 0)$ satisfying $p_0 = 0$ and $p_1 > 0$. By applying the self-normalized large deviation of Shao (1997) for independent and identically distributed random variables, we obtain the self-normalized large deviation for supercritical branching processes, which is the self-normalized version of the result obtained by Athreya (1994). The self-normalized large deviation can also be generalized to supercritical multitype branching processes.

Keywords: Self-normalized large deviation; supercritical branching process; multitype branching process

2010 Mathematics Subject Classification: Primary 60J80; 60F10
Secondary 60J27

1. Introduction

Let $(Z_n,\ n \geq 0)$ be a supercritical branching process with $Z_0 = 1$, offspring distribution $p_k = \mathbb{P}(X = k)$, $k = 0, 1, 2, \ldots$, and mean $\mu = \mathbb{E}X \in (1, \infty)$. That is,

$$Z_0 = 1, \qquad Z_{n+1} = X_{n,1} + X_{n,2} + \cdots + X_{n,Z_n}, \quad n \geq 0, \tag{1}$$

where $X_{n,k}$ is the number of offspring of the kth individual in generation n, and $(X_{n,k},\ n \geq 0,\ k \geq 1)$ are independent and identically distributed (i.i.d.) with the same distribution as X. To avoid the deterministic case, we suppose that $p_k < 1$ for any nonnegative integer k. Without loss of generality, we also assume that $p_0 = 0$ throughout this paper. Then we have $Z_n \to \infty$ almost surely (a.s.) as $n \to \infty$. Thus, according to the strong law of large numbers, we have, as $n \to \infty$,

$$\frac{Z_{n+1}}{Z_n} \to \mu \quad \text{a.s.}$$

One of the interesting topics is to consider the convergence rate of

$$\mathbb{P}\bigg(\bigg|\frac{Z_{n+1}}{Z_n} - \mu\bigg| > \varepsilon\bigg) \quad \text{for } \varepsilon > 0 \text{ as } n \to \infty.$$

In [1], the author obtained the following theorem.

Theorem 1. ([1, Theorem 1].) *Assume that $p_1 > 0$ and $\mathbb{E}[\exp(\theta Z_1) \mid Z_0 = 1] < \infty$ for some $\theta > 0$. Then, for all $\varepsilon > 0$,*

$$\lim_{n\to\infty} \frac{1}{p_1^n}\mathbb{P}\bigg(\bigg|\frac{Z_{n+1}}{Z_n} - \mu\bigg| > \varepsilon \;\bigg|\; Z_0 = 1\bigg) = \sum_k \phi(k, \varepsilon) q_k < \infty, \tag{2}$$

Received 12 May 2016; revision received 12 January 2018.
* Postal address: College of Science, Hohai University, No. 8 Focheng Road West, Nanjing, Jiangsu Province, 210098, P. R. China. Email address: chuwj@hhu.edu.cn

where $\phi(k,\varepsilon)=\mathbb{P}(|\bar{X}_k-\mu|>\varepsilon)$, $\bar{X}_k$ *is the mean* $(1/k)\sum_{i=1}^k X_i$ *of* k *i.i.d. random variables* $\{X_i\}$ *with distribution* $\{p_j,\ j\geq 1\}$, *and* $\{q_k\}$ *is defined via the generating function* $Q(s)=\sum_{k=1}^{\infty}q_k s^k$, $0\leq s<1$, *the unique solution of the functional equation*

$$Q(f(s))=p_1Q(s),\quad \textit{where } f(s)=\sum_{j=1}^{\infty}p_js^j,\ 0\leq s<1,$$

subject to

$$Q(0)=0,\qquad Q(1)=\infty,\qquad Q(s)<\infty\quad \textit{for } 0\leq s<1.$$

Weaker conditions for the large deviation in (2) to hold do exist; see, for example, [1, Theorem 2 and Corollary 1], where the exponential moment condition is replaced by $p_1\mu^r>1$ and $\mathbb{P}(|\bar{X}_k-\mu|>\varepsilon)\leq C_\varepsilon k^{-r}$ for some constants C_ε and $r>0$, or by $\mathbb{E}[Z_1^{2\alpha+\delta}\mid Z_0=1]<\infty$ for some $\alpha\geq 1$ and $\delta>0$ such that $p_1\mu^\alpha>1$. Other related results with weaker moment conditions can be found in [4], [6], and [7].

All the abovementioned results require the use of moment conditions for the offspring X. In this paper we use the self-normalized large deviation obtained by Shao [9], recalled in Theorem 2 below, to establish the self-normalized large deviation for supercritical branching processes in the $p_1>0$ case. First we introduce the self-normalized large deviation results for the i.i.d. random variables.

Let $Y,Y_1,\dots,Y_n$ be i.i.d. random variables with $\mathbb{P}(Y\neq 0)>0$, and let $S_n=\sum_{i=1}^n Y_i$. Shao [9] obtained the large deviation for S_n in the self-normalized version without any moment conditions.

Theorem 2. ([9, Theorem 1.1].) *Let* $Y,Y_1,Y_2,\dots$ *be i.i.d. random variables. Assume that either* $\mathbb{E}Y\geq 0$ *or* $\mathbb{E}Y^2=\infty$. *Let* $V_n^2=\sum_{i=1}^n Y_i^2$. *Then*

$$\lim_{n\to\infty}\mathbb{P}(S_n\geq x\sqrt{n}V_n)^{1/n}=\sup_{c\geq 0}\inf_{t\geq 0}\mathbb{E}[\mathrm{e}^{t(cY-x(Y^2+c^2)/2)}]$$

for $x>\mathbb{E}Y/(\mathbb{E}Y^2)^{1/2}$, *where* $\mathbb{E}Y/(\mathbb{E}Y^2)^{1/2}=0$ *if* $\mathbb{E}Y^2=\infty$.

Now we consider the self-normalized large deviation for supercritical branching processes. That is, we find a nondeterministic normalization $(R_n,\ n\geq 1)$ for $|Z_{n+1}/Z_n-\mu|$, such that

$$R_n\bigg|\frac{Z_{n+1}}{Z_n}-\mu\bigg|\to 0\quad\text{in probability as } n\to\infty$$

with convergence rate p_1^n. The following theorem is the main result of this paper.

Theorem 3. *If* $p_0=0$, $p_1>0$, *and the offspring mean* $\mu\in(1,\infty)$, *then, for all* $x>0$,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{p_1^n}\mathbb{P}\bigg(\frac{\sqrt{Z_n}}{\sqrt{\sum_{i=1}^{Z_n}(X_{n,i}-Z_{n+1}/Z_n)^2}}\bigg|\frac{Z_{n+1}}{Z_n}-\mu\bigg|\geq x\;\bigg|\;Z_0=1\bigg)\\ =\sum_{k=1}^{\infty}q_k\psi(k,x)\\ <\infty,\end{aligned}\tag{3}$$

https://doi.org/10.1017/jpr.2018.29 Published online by Cambridge University Press

where the $\{X_{n,i}\}$ *are the same as in (1),*

$$\psi(k,x)=\mathbb{P}\biggl(\biggl|\sum_{i=1}^{k}(X_i-\mu)\biggr|\geq x\sqrt{k}\sqrt{\sum_{i=1}^{k}(X_i-\bar{X}_k)^2}\biggr)$$

$\bar{X}_k$ *is the mean* $(1/k)\sum_{i=1}^{k}X_i$ *of* k *i.i.d. random variables* $\{X_i\}$ *with distribution* $\{p_j,\ j\geq 1\}$, *and the* $\{q_k\}$ *are as defined in Theorem 1.*

The self-normalized large deviations usually have some certain statistics applications. This result can be used for statistical inference and to construct a confidence interval for the mean μ for instance.

2. Proof of Theorem 3

Recalling that Z_n is the number of individuals of the branching process in generation n, and $X_{n,i},\ 1\leq i\leq Z_n$, is the number of offspring of the ith individual in generation n, we define

$$V(n)^2=\sum_{i=1}^{Z_n}(X_{n,i}-\mu)^2,\qquad \bar{X}(n)=\frac{1}{Z_n}\sum_{i=1}^{Z_n}X_{n,i}, \tag{4}$$

$$V_k^2=\sum_{i=1}^{k}(X_i-\mu)^2,\qquad \bar{X}_k=\frac{1}{k}\sum_{i=1}^{k}X_i, \tag{5}$$

where $(X_i,\ i\geq 1)$ is an independent random variable series and distributed as the offspring number X. Then we have $\bar{X}(n)=Z_{n+1}/Z_n$ and

$$\sum_{i=1}^{Z_n}(X_{n,i}-\bar{X}(n))^2=V(n)^2-Z_n(\mu-\bar{X}(n))^2. \tag{6}$$

First we state a convergence property about the generating function of Z_n that will be useful later in this section. Define the generating function of Z_n as

$$f_n(s)=\mathbb{E}[s^{Z_n}\mid Z_0=1],\qquad |s|\leq 1.$$

Lemma 1. ([1, Propositions 2 and 3].) *If* $p_1>0$ *then*

$$\lim_{n\to\infty}\frac{f_n(s)}{p_1^n}=\sum_{k=1}^{\infty}q_k s^k, \tag{7}$$

$$\lim_{n\to\infty}\frac{\mathbb{P}(Z_n=k\mid Z_0=1)}{p_1^n}=q_k\quad \textit{for all } k\geq 1, \tag{8}$$

where $(q_k,\ k\geq 1)$ *is the same as in Theorem 1.*

For any $x>0$, by (6), the probability in (3) can be written as

$$\mathbb{P}\biggl(\biggl|\frac{\sum_{i=1}^{Z_n}(X_{n,i}-\mu)}{\sqrt{V(n)^2-Z_n(\mu-\bar{X}(n))^2}}\frac{1}{\sqrt{Z_n}}\biggr|\geq x\biggr).$$

Now we are ready to prove Theorem 3.

https://doi.org/10.1017/jpr.2018.29 Published online by Cambridge University Press

Proof of Theorem 3. For any integer $n \geq 1$, by the total probability formula, the independence of Z_n and $(X_{n,i},\ i \geq 1)$, and (4) and (5), we obtain

$$\begin{aligned}
&\mathbb{P}\biggl(\biggl|\sum_{i=1}^{Z_n}(X_{n,i}-\mu)\biggr| \geq x\sqrt{Z_n}\sqrt{V(n)^2 - Z_n(\mu-\bar{X}(n))^2}\biggr) \\
&\quad = \sum_{k=1}^{\infty}\mathbb{P}(Z_n = k)\mathbb{P}\biggl(\biggl|\sum_{i=1}^{k}(X_i-\mu)\biggr| \geq x\sqrt{k}\sqrt{V_k^2 - k(\mu-\bar{X}_k)^2}\biggr) \\
&\quad := \sum_{k=1}^{\infty}\mathbb{P}(Z_n = k)\psi(k,x),
\end{aligned}$$

where $\psi(k,x)$ is defined in Theorem 3 and $\{X_i,\ i \geq 1\}$ is an independent random series, distributed as X. According to the value of $\sum_{i=1}^{k}(X_i-\mu)$, $\psi(k,x)$ can be divided into two parts:

$$\begin{aligned}
\psi(k,x) &= \mathbb{P}\biggl(\sum_{i=1}^{k}(X_i-\mu) \geq x\sqrt{k}\sqrt{V_k^2 - k(\mu-\bar{X}_k)^2}\biggr) \\
&\qquad + \mathbb{P}\biggl(\sum_{i=1}^{k}(\mu-X_i) \geq x\sqrt{k}\sqrt{V_k^2 - k(\mu-\bar{X}_k)^2}\biggr) \\
&:= I(k) + J(k).
\end{aligned}$$

Now if both $I(k)$ and $J(k)$ converge to 0 quickly enough as $k \to \infty$ then, by a slight modification of the Lebesgue dominated convergence theorem (see [8, Proposition 18, p. 270]) together with (7) and (8), we can obtain Theorem 3.

Now we prove that both $I(k)$ and $J(k)$ converge to 0 exponentially as $k \to \infty$. For any $0 < \varepsilon < 1$, $I(k)$ becomes

$$\begin{aligned}
I(k) &= \mathbb{P}\biggl(\sum_{i=1}^{k}(X_i-\mu) \geq x\sqrt{k}\sqrt{V_k^2 - k(\mu-\bar{X}_k)^2},\ k(\mu-\bar{X}_k)^2 < \varepsilon V_k^2\biggr) \\
&\qquad + \mathbb{P}\biggl(\sum_{i=1}^{k}(X_i-\mu) \geq x\sqrt{k}\sqrt{V_k^2 - k(\mu-\bar{X}_k)^2},\ k(\mu-\bar{X}_k)^2 \geq \varepsilon V_k^2\biggr) \\
&\leq \mathbb{P}\biggl(\sum_{i=1}^{k}(X_i-\mu) \geq x\sqrt{(1-\varepsilon)k}V_k\biggr) + \mathbb{P}(k(\mu-\bar{X}_k)^2 \geq \varepsilon V_k^2). \qquad (9)
\end{aligned}$$

Note that

$$k(\mu-\bar{X}_k)^2 = \frac{1}{k}\biggl(\sum_{i=1}^{k}(X_i-\mu)\biggr)^2.$$

Thus, for the second part of (9), we have

$$\begin{aligned}
\mathbb{P}(k(\mu-\bar{X}_k)^2 \geq \varepsilon V_k^2) &= \mathbb{P}\biggl(\biggl(\sum_{i=1}^{k}(X_i-\mu)\biggr)^2 \geq \varepsilon k V_k^2\biggr) \\
&= \mathbb{P}\biggl(\sum_{i=1}^{k}(X_i-\mu) \geq \sqrt{\varepsilon k}V_k\biggr) + \mathbb{P}\biggl(\sum_{i=1}^{k}(\mu-X_i) \geq \sqrt{\varepsilon k}V_k\biggr). \qquad (10)
\end{aligned}$$

https://doi.org/10.1017/jpr.2018.29 Published online by Cambridge University Press

By applying Theorem 2 with variables $Y_i = X_i - \mu$ for constant $\delta > 0$ to be determined later, we obtain, for the first part of (10),

$$\mathbb{P}\bigg(\sum_{i=1}^{k}(X_i - \mu) \geq \sqrt{\varepsilon k} V_k\bigg) \leq (1+\delta)^k \rho_1^k \quad \text{for large enough integer } k, \tag{11}$$

where

$$\rho_1 := \sup_{c\geq 0}\inf_{t\geq 0} \mathbb{E}\big[\exp\big(t\big(c(X-\mu) - \sqrt{\varepsilon}\tfrac{1}{2}((X-\mu)^2 + c^2)\big)\big] \in (0,1).$$

For the second part of (10), we can apply Theorem 2 with variables $Y_i = \mu - X_i$ to obtain

$$\mathbb{P}\bigg(\sum_{i=1}^{k}(\mu - X_i) \geq \sqrt{\varepsilon k} V_k\bigg) \leq (1+\delta)^k \rho_2^k \quad \text{for large enough } k, \tag{12}$$

where

$$\rho_2 := \sup_{c\geq 0}\inf_{t\geq 0} \mathbb{E}\big[\exp\big(t\big(c(\mu - X) - \sqrt{\varepsilon}\tfrac{1}{2}((\mu - X)^2 + c^2)\big)\big)\big] \in (0,1).$$

Similarly, the first probability in (9) can be estimated by

$$\mathbb{P}\bigg(\sum_{i=1}^{k}(X_i - \mu) \geq x\sqrt{(1-\varepsilon)k} V_k\bigg) \leq (1+\delta)^k \rho_3^k \quad \text{for large enough } k, \tag{13}$$

where

$$\rho_3 := \sup_{c\geq 0}\inf_{t\geq 0} \mathbb{E}\big[\exp\big(t\big(c(X-\mu) - x\sqrt{1-\varepsilon}\tfrac{1}{2}((X-\mu)^2 + c^2)\big)\big)\big] \in (0,1).$$

The same argument can be applied to estimate $J(k)$ by dealing with $-X_i$ instead. Indeed,

$$J(k) \leq C_2(1+\delta)^k \rho_4^k \quad \text{for all } k \geq 1 \tag{14}$$

for some constants $\rho_4 \in (0,1)$ and $C_2 > 0$. Now we define $\rho = \max\{\rho_1, \rho_2, \rho_3, \rho_4\}$ and choose a constant $\delta \in (0,1)$ such that $(1+\delta)\rho < 1$. Therefore, by (9)–(14), we obtain

$$\psi(k,x) = I(k) + J(k) \leq C(1+\delta)^k \rho^k \quad \text{for all } k \geq 1$$

with C being a positive constant. Therefore,

$$0 \leq h_n(k) := \frac{\mathbb{P}(Z_n = k)}{p_1^n}\psi(k,x) \leq C\frac{\mathbb{P}(Z_n = k)}{p_1^n}(1+\delta)^k\rho^k =: Cg_n(k).$$

By (8), for any $k \geq 1$,

$$\frac{\mathbb{P}(Z_n = k)}{p_1^n} \to q_k \quad \text{as } n \to \infty.$$

Thus, for any $k \geq 1$, $g_n(k) \to q_k(1+\delta)^k\rho^k$ and $h_n(k) \to q_k\psi(k,x)$ as $n \to \infty$. Moreover, by (7), we have

$$\lim_{n\to\infty}\sum_{k=1}^{\infty} g_n(k) = \lim_{n\to\infty}\frac{f_n((1+\delta)\rho)}{p_1^n} = \sum_{k=1}^{\infty} q_k(1+\delta)^k\rho^k < \infty.$$

https://doi.org/10.1017/jpr.2018.29 Published online by Cambridge University Press

Therefore, by the dominated convergence theorem (see, for example, [8, Proposition 18]), we have

$$\lim_{n\to\infty}\sum_{k=1}^{\infty}h_n(k)=\sum_{k=1}^{\infty}q_k\psi(k,x),$$

which is our result in Theorem 3. □

3. Self-normalized large deviation for supercritical multitype branching processes

In this section we consider the self-normalized large deviation for supercritical multitype branching processes. Assume that $\boldsymbol{Z}_n(i)=(Z_n(i,1),\dots,Z_n(i,d))$ is a d-type branching process initiated from a single particle of type i. For ease of exposition, we consider only the $d=2$ case, that is,

$$\boldsymbol{Z}_0(i)=\boldsymbol{I}_i,\qquad \boldsymbol{Z}_n(i)=\sum_{j=1}^{2}\sum_{k=1}^{Z_{n-1}(i,j)}\boldsymbol{X}_{n-1}^{k}(j),$$

where $\boldsymbol{I}_1=(1,0)$, $\boldsymbol{I}_2=(0,1)$, and $\boldsymbol{X}_{n-1}^k(j)=(X_{n-1}^k(j,1),X_{n-1}^k(j,2))$ are the offspring of the kth individual of type j in generation $n-1$. Here $(\boldsymbol{X}_{n-1}^k(j),\ n=1,2,\dots,\ k=1,2,\dots)$ are independent and have the same distribution as $\boldsymbol{X}(j)=(X(j,1),X(j,2))$, $j=1,2$. These random vectors take values in $\mathbb{Z}_+^2$, the set of two-dimensional vectors with elements being nonnegative integers.

Throughout this section, vectors are set as bold characters. In particular, we use $\boldsymbol{0}$ and $\boldsymbol{1}$ to respectively denote the two-dimensional zero vector and the vector with each element being 1. The partial ordering of two-dimensional vectors is defined as $\boldsymbol{x}=(x_1,x_2)\preceq(\prec)\boldsymbol{y}=(y_1,y_2)$ if and only if $x_i\le(<)y_i$ for each i.

Next we state the notation, definitions, and assumptions about the multitype branching processes that will be used in this section.

Let $M=(\mu_{ij})_{i,j=1}^2$ with $\mu_{ij}=\mathbb{E}X(i,j)$ the mean matrix, which is positively regular ($M^n>0$ for some integer $n>0$). We also assume that its maximal eigenvalue is $\rho>1$, which corresponds to the supercritical case, and $\boldsymbol{t}=(t_1,t_2)$ and $\boldsymbol{u}=(u_1,u_2)$ are strictly positive left and right eigenvectors corresponding to ρ, normalized so that $\boldsymbol{1}^\top\boldsymbol{u}=1$ and $\boldsymbol{t}^\top\boldsymbol{u}=1$.

For $\boldsymbol{s}=(s_1,s_2)$ with $0\le s_i\le 1$, $i=1,2$, define

$$f^{(i)}(\boldsymbol{s})=\mathbb{E}\boldsymbol{s}^{\boldsymbol{Z}_1(i)}=\sum_{\boldsymbol{j}}p^{(i)}(\boldsymbol{j})s_1^{j_1}s_2^{j_2},$$

where $\boldsymbol{j}=(j_1,j_2)\in\mathbb{Z}_+^2$ and

$$p^{(i)}(\boldsymbol{j})=\mathbb{P}(\boldsymbol{Z}_1=\boldsymbol{j}\mid \boldsymbol{Z}_0=\boldsymbol{I}_i).$$

If we define $\boldsymbol{f}(\boldsymbol{s})=(f^{(1)}(\boldsymbol{s}),f^{(2)}(\boldsymbol{s}))$ then $\boldsymbol{f}(\boldsymbol{s})=\boldsymbol{s}$ has a unique solution $\boldsymbol{e}=(e_1,e_2)$ with each $e_i\in[0,1)$, $i=1,2$. In addition, e_i is the extinction probability of $(\boldsymbol{Z}_n(i),\ n\ge 0)$, that is,

$$e_i=\mathbb{P}\Big(\lim_{n\to\infty}\boldsymbol{Z}_n=\boldsymbol{0}\ \Big|\ \boldsymbol{Z}_0=\boldsymbol{I}_i\Big).$$

After the (multitype) Sevastyanov transformation, $\boldsymbol{f}(\boldsymbol{0})$ can be zero. Then we have $\boldsymbol{e}=\boldsymbol{0}$. We make this assumption throughout this section. In this case, we define the matrix

$$A=\bigg(\frac{\partial f^{(i)}(\boldsymbol{s})}{\partial s_j}\bigg)_{i,j=1}^{2}\bigg|_{\boldsymbol{s}=\boldsymbol{0}},$$

https://doi.org/10.1017/jpr.2018.29 Published online by Cambridge University Press

and assume that there exists a constant $\gamma \in (0, 1)$ such that $(\gamma^{-n} A^n)_{n\geq 1}$ converges to a matrix which is nonzero and finite. In fact, this is the so-called Schröder case (see [5]). For more details about the multitype branching processes, we refer the reader to [2, Chapter V].

The following result is the large deviation of $(\boldsymbol{Z}_n(i),\ n \geq 0)$ obtained in [3].

Theorem 4. ([3, Theorem 2].) *Assume that the abovementioned assumptions hold and that*

$$\max_i \mathbb{E}[(\boldsymbol{1}\cdot \boldsymbol{Z}_1^\top)^{2r} \mid \boldsymbol{Z}_0 = \boldsymbol{I}_i] < \infty,$$

where r is such that $\rho^r\gamma > 1$. Let $\boldsymbol{l} = (l_1, l_2)$ be a nonzero vector with $l_1 \neq l_2$. Then, for every $\varepsilon > 0$ and $i = 1, 2$, the limit

$$\lim_{n\to\infty} \gamma^{-n}\mathbb{P}\bigg(\bigg|\frac{\boldsymbol{l}\cdot \boldsymbol{Z}_{n+1}^\top}{\boldsymbol{1}\cdot \boldsymbol{Z}_n^\top} - \frac{\boldsymbol{l}\cdot(\boldsymbol{Z}_n M)^\top}{\boldsymbol{1}\cdot \boldsymbol{Z}_n^\top}\bigg| > \varepsilon \ \bigg|\ \boldsymbol{Z}_0 = \boldsymbol{I}_i\bigg) \tag{15}$$

exists, and is positive and finite.

For simplicity, in the remainder of this section we omit the initial particle of type i in $\boldsymbol{Z}_n(i) = (Z_n(i, 1), Z_n(i, 2))$, and denote the process as $(\boldsymbol{Z}_n,\ n \geq 0)$ and $\boldsymbol{Z}_n = (Z_n(1), Z_n(2))$. Thus, the vectors in (15) become

$$\begin{aligned}\boldsymbol{1}\cdot \boldsymbol{Z}_n^\top &= Z_n(1) + Z_n(2),\\ \boldsymbol{l}\cdot \boldsymbol{Z}_{n+1}^\top - \boldsymbol{l}\cdot(\boldsymbol{Z}_n M)^\top &= \sum_{k=1}^{Z_n(1)} (l_1(X_n^k(1, 1) - \mu_{11}) + l_2(X_n^k(1, 2) - \mu_{12}))\\ &\quad + \sum_{k=1}^{Z_n(2)} (l_1(X_n^k(2, 1) - \mu_{21}) + l_2(X_n^k(2, 2) - \mu_{22})).\end{aligned}$$

Then the self-normalized version is

$$\frac{S(n)}{\sqrt{(Z_n(1) + Z_n(2))V(n)}},$$

where

$$\begin{aligned}S(n) &= \sum_{k=1}^{Z_n(1)} (l_1(X_n^k(1, 1) - \mu_{11}) + l_2(X_n^k(1, 2) - \mu_{12}))\\ &\quad + \sum_{k=1}^{Z_n(2)} (l_1(X_n^k(2, 1) - \mu_{21}) + l_2(X_n^k(2, 2) - \mu_{22}))\end{aligned} \tag{16}$$

and

$$\begin{aligned}V(n)^2 &= \sum_{k=1}^{Z_n(1)} (l_1^2(X_n^k(1, 1) - \bar{X}_n(1, 1))^2 + l_2^2(X_n^k(1, 2) - \bar{X}_n(1, 2))^2)\\ &\quad + \sum_{k=1}^{Z_n(2)} (l_1^2(X_n^k(2, 1) - \bar{X}_n(2, 1))^2 + l_2^2(X_n^k(2, 2) - \bar{X}_n(2, 2))^2)\end{aligned} \tag{17}$$

with

$$\bar{X}_n(i, j) = \frac{1}{Z_n(i)} \sum_{k=1}^{Z_n(i)} X_n^k(i, j), \qquad i, j = 1, 2.$$

In a similar fashion as in the single-type case, using the total probability formula and then the

https://doi.org/10.1017/jpr.2018.29 Published online by Cambridge University Press

dominated convergence theorem, we obtain the self-normalized large deviation for the multitype case.

Theorem 5. *With* $S(n)$ *and* $V(n)$ *defined in (16) and (17), the limit*

$$\lim_{n\to\infty} \gamma^{-n}\mathbb{P}\bigg(\frac{S(n)}{\sqrt{(Z_n(1)+Z_n(2))}V(n)} > x \,\bigg|\, \boldsymbol{Z}_0 = \boldsymbol{I}_i\bigg)$$

exists, and is positive and finite.

For $i = 1, 2$, let $(\boldsymbol{X}^k(i) = (X^k(i,1), X^k(i,2)),\ k \geq 1)$ be a sequence of independent random vectors that have the same distribution as $\boldsymbol{X}(i)$. For $i, j = 1, 2$, and integers m, n, define

$$\begin{aligned}
Y_{i,j}(k) &= l_j(X^k(i,j) - \mu_{ij}),\\
S_{ij}(n) &= \sum_{k=1}^{n} Y_{i,j}(k),\\
S(n,m) &= S_{11}(n) + S_{12}(n) + S_{21}(m) + S_{22}(m),\\
\bar{X}_n(i,j) &= \frac{1}{n}\sum_{k=1}^{n} X^k(i,j),\\
V_{ij}(n)^2 &= \sum_{k=1}^{n} Y_{i,j}(k)^2,\\
V(n,m)^2 &= V_{11}(n)^2 + V_{12}(n)^2 + V_{21}(m)^2 + V_{22}(m)^2,\\
V_{n,m}^2 &= \sum_{j=1}^{2}\sum_{k=1}^{n} l_j^2(X^k(1,j) - \bar{X}_n(1,j))^2 + \sum_{j=1}^{2}\sum_{k=1}^{m} l_j^2(X^k(2,j) - \bar{X}_n(2,j))^2,\\
\varepsilon(n,m) &= V(n,m)^2 - V_{n,m}^2 = n\sum_{j=1}^{2}(\bar{X}_n(1,j) - \mu_{1j})^2 + m\sum_{j=1}^{2}(\bar{X}_n(2,j) - \mu_{2j})^2.
\end{aligned}$$

We are now ready to prove Theorem 5.

Proof of Theorem 5. Similarly as in the proof of Theorem 3, after using the total probability formula in

$$\mathbb{P}\bigg(\frac{S(n)}{\sqrt{(Z_n(1)+Z_n(2))}V(n)} > x \,\bigg|\, \boldsymbol{Z}_0 = \boldsymbol{I}_i\bigg),$$

it remains to verify the dominated convergence, which is ensured by the exponential convergence to 0 as $n, m \to \infty$ of

$$\mathbb{P}(S(n,m) \geq x\sqrt{n+m}V_{n,m}).$$

For any constant $\delta \in (0, 1)$, we have

$$\begin{aligned}
&\mathbb{P}(S(n,m) \geq x\sqrt{n+m}V_{n,m})\\
&\quad\leq \mathbb{P}(S(n,m) \geq x\sqrt{n+m}V_{n,m},\ \varepsilon(n,m) \leq \delta V(n,m)^2) + \mathbb{P}(\varepsilon(n,m) > \delta V(n,m)^2)\\
&\quad\leq \mathbb{P}(S(n,m) \geq x\sqrt{1-\delta}\sqrt{n+m}V(n,m)) + \mathbb{P}(\varepsilon(n,m) > \delta V(n,m)^2)\\
&\quad:= \mathrm{I} + \mathrm{II}, \qquad (18)
\end{aligned}$$

where

$$\begin{aligned}\mathrm{I} &\le \mathbb{P}(S(n,m)^2 \ge x^2(1-\delta)(n+m)V(n,m)^2)\\ &\le \mathbb{P}(4(S_{11}(n)^2+\cdots+S_{22}(m)^2) \ge x^2(1-\delta)(n+m)V(n,m)^2)\\ &\le \mathbb{P}(4(S_{11}(n)^2+\cdots+S_{22}(m)^2) \ge x^2(1-\delta)(nV_{11}(n)^2+\cdots+mV_{22}(m)^2))\\ &\le \mathbb{P}(4S_{11}(n)^2 \ge x^2(1-\delta)nV_{11}(n)^2)+\cdots+\mathbb{P}(4S_{22}(m)^2 \ge x^2(1-\delta)mV_{22}(m)^2)\end{aligned}$$

and

$$\begin{aligned}\mathrm{II} &\le \sum_{j=1}^{2}\mathbb{P}(n(\bar{X}_n(1,j)-\mu_{1j})^2 > \delta V_{1j}^2(n)) + \sum_{j=1}^{2}\mathbb{P}(m(\bar{X}_m(2,j)-\mu_{2j})^2 > \delta V_{2j}^2(m))\\ &= \mathbb{P}\bigg(\bigg|\sum_{k=1}^{n}(X^k(1,1)-\mu_{11})\bigg| \ge \sqrt{\delta n}V_{1,1}(n)\bigg)\\ &\quad + \mathbb{P}\bigg(\bigg|\sum_{k=1}^{n}(X^k(1,2)-\mu_{12})\bigg| \ge \sqrt{\delta n}V_{1,2}(n)\bigg)\\ &\quad + \mathbb{P}\bigg(\bigg|\sum_{k=1}^{m}(X^k(2,1)-\mu_{21})\bigg| \ge \sqrt{\delta m}V_{2,1}(m)\bigg)\\ &\quad + \mathbb{P}\bigg(\bigg|\sum_{k=1}^{m}(X^k(2,2)-\mu_{22})\bigg| \ge \sqrt{\delta m}V_{2,2}(m)\bigg).\end{aligned}$$

From Theorem 2, we see that the probability in (18) converges to 0 exponentially as $n, m \to \infty$. We can then obtain Theorem 5 by the same arguments used in the proof of Theorem 3 and the results of [3, Theorem 1]. □

Acknowledgements

The research of Weijuan Chu was supported by the National Natural Science Foundation of China (under grant numbers 11201224, 2017B05314, and 11771209). The anonymous referee is also gratefully acknowledged for a careful reading and insightful comments, from which the author has benefited greatly.

References

[1] Athreya, K. B. (1994). Large deviation rates for branching processes. I. Single type case. *Ann. Appl. Prob.* **4,** 779–790.
[2] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes*. Springer, New York.
[3] Athreya, K. B. and Vidyashankar, A. N. (1995). Large deviation rates for branching processes. II. The multitype case. *Ann. Appl. Prob.* **5,** 566–576.
[4] He, H. (2016). On large deviation rates for sums associated with Galton–Watson processes. *Adv. Appl. Prob.* **48,** 672–690.
[5] Jones, O. D. (2004). Large deviations for supercritical multitype branching processes. *J. Appl. Prob.* **41,** 703–720.
[6] Ney, P. E. and Vidyashankar, A. N. (2003). Harmonic moments and large deviation rates for supercritical branching processes. *Ann. Appl. Prob.* **13,** 475–489.
[7] Ney, P. E. and Vidyashankar, A. N. (2004). Local limit theory and large deviations for supercritical branching processes. *Ann. Appl. Prob.* **14,** 1135–1166. (Correction: **16** (2006), 2272.)
[8] Royden, H. L. (1988). Real Analysis, 3rd edn. Macmillan, New York.
[9] Shao, Q.-M. (1997). Self-normalized large deviations. *Ann. Prob.* **25,** 285–328.