J. Appl. Probab. **60,** 1281–1292 (2023)
*doi:*10.1017/jpr.2022.134

SELF-NORMALIZED CRAMÉR MODERATE DEVIATIONS FOR A SUPERCRITICAL GALTON–WATSON PROCESS

XIEQUAN FAN,* *Northeastern University at Qinhuangdao*
QI-MAN SHAO,** *Southern University of Science and Technology*

Abstract

Let $(Z_n)_{n\geq 0}$ be a supercritical Galton–Watson process. Consider the Lotka–Nagaev estimator for the offspring mean. In this paper we establish self-normalized Cramér-type moderate deviations and Berry–Esseen bounds for the Lotka–Nagaev estimator. The results are believed to be optimal or near-optimal.

Keywords: Self-normalized processes; Lotka–Nagaev estimator; Berry–Esseen bounds; offspring mean

2020 Mathematics Subject Classification: Primary 60F10; 60J80
Secondary 62F12; 62F03

1. Introduction

A Galton–Watson process can be described as follows:

$$Z_0 = 1, \quad Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i} \quad \text{for } n \geq 0,$$

where $X_{n,i}$ is the offspring number of the ith individual of the generation n. Moreover, the random variables $(X_{n,i})_{i\geq 1}$ are independent of each other with common distribution law

$$\mathbb{P}(X_{n,i} = k) = p_k, \quad k \in \mathbb{N},$$

and are also independent of Z_n.

An important task in statistical inference for Galton–Watson processes is to estimate the average offspring number of an individual m, usually termed the offspring mean. Clearly we have

$$m = \mathbb{E}Z_1 = \mathbb{E}X_{n,i} = \sum_{k=0}^{\infty} kp_k.$$

Let v denote the standard variance of Z_1, that is,

$$v^2 = \mathbb{E}(Z_1 - m)^2.$$

Received 25 August 2021; revision received 9 December 2022.
* Postal address: School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao, 066004, Hebei, China. Email address: fanxiequan@hotmail.com
** Postal address: Department of Statistics and Data Science, SICM, National Center for Applied Mathematics Shenzhen, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China. Email address: shaoqm@sustech.edu.cn

To avoid triviality, assume that $v > 0$. For estimation of the offspring mean m, the Lotka–Nagaev [12, 14] estimator Z_{n+1}/Z_n plays an important role. Throughout the paper we assume that

$$p_0 = 0.$$

Then the Lotka–Nagaev estimator is well-defined $\mathbb{P}$-a.s. For the Galton–Watson processes, Athreya [1] has established large deviations for the normalized Lotka–Nagaev estimator (see also Chu [3] for self-normalized large deviations); Ney and Vidyashankar [15, 16] and He [9] obtained sharp rate estimates for the large deviation behavior of the Lotka–Nagaev estimator; Maaouia and Touati [13] established a self-normalized central limit theorem (CLT) for the maximum likelihood estimator of m; Bercu and Touati [2] proved an exponential inequality for the Lotka–Nagaev estimator via self-normalized martingale methods. Alternative approaches for obtaining self-normalized exponential inequalities can be found in de la Peña, Lai, and Shao [4]. Despite the fact that the Lotka–Nagaev estimator is well studied, there is no result for self-normalized Cramér moderate deviations for the Lotka–Nagaev estimator. The main purpose of this paper is to fill this gap.

Let us briefly introduce our main result. Assume that $n_0, n \in \mathbb{N}$. Notice that, by the classical CLT for independent and identically distributed (i.i.d.) random variables,

$$\mathbf{X}_{n_n,n} = \left(\sqrt{Z_{n_0}}\left(\frac{Z_{n_0+1}}{Z_{n_0}} - m\right), \ldots, \sqrt{Z_{n_0+n-1}}\left(\frac{Z_{n_0+n}}{Z_{n_0+n-1}} - m\right)\right)$$

asymptotically behaves like a vector of i.i.d. Gaussian random variables with mean 0 and variance v^2 (even if n_0 depends on n) and the convergence rate to Gaussian distribution is exponential; see Kuelbs and Vidyashankar [11]. Because

$$\frac{1}{n}\sum_{k=n_0}^{n_0+n-1} Z_k\left(\frac{Z_{k+1}}{Z_k} - m\right)^2$$

is an estimator of the offspring variance v^2, it is natural to compare the self-normalized sum

$$M_{n_0,n} := \frac{(nv^2)^{-1/2}\sum_{k=n_0}^{n_0+n-1}\sqrt{Z_k}(Z_{k+1}/Z_k - m)}{\sqrt{(nv^2)^{-1}\sum_{k=n_0}^{n_0+n-1} Z_k(Z_{k+1}/Z_k - m)^2}} = \frac{\sum_{k=n_0}^{n_0+n-1}\sqrt{Z_k}(Z_{k+1}/Z_k - m)}{\sqrt{\sum_{k=n_0}^{n_0+n-1} Z_k(Z_{k+1}/Z_k - m)^2}}$$

to the tail of the Gaussian distribution. This is the main purpose of the paper. Assume that $\mathbb{E}Z_1^{2+\rho} < \infty$ for some $\rho \in (0, 1]$. We prove the following self-normalized Cramér moderate deviations for the Lotka–Nagaev estimator. It holds that

$$\mathbb{P}(\pm\, M_{n_0,n} \geq x) = (1 - \Phi(x))(1 + \mathrm{o}(1)) \tag{1}$$

uniformly for $x \in [0,\ \mathrm{o}(n^{\rho/(4+2\rho)}))$ as $n \to \infty$; see Theorem 2.1. This type of result is user-friendly in the statistical inference of m, since in practice we usually do not know the variance v^2 or the distribution of Z_1. Let $\kappa_n \in (0, 1)$. Assume that

$$|\ln \kappa_n| = \mathrm{o}\big(n^{\rho/(2+\rho)}\big), \qquad n \to \infty.$$

From (1) we can easily obtain a $1 - \kappa_n$ confidence interval for m, for n large enough. Clearly, the right-hand side of (1) and $M_{n_0,n}$ do not depend on v^2, so the confidence interval of m does

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

not depend on v^2; see Proposition 3.1. Due to these significant advantages, the limit theory for self-normalized processes is attracting more and more attention. We refer to Jing, Shao, and Wang [10] and Fan *et al.* [8] for closely related results.

The paper is organized as follows. In Section 2 we present Cramér moderate deviations for the self-normalized Lotka–Nagaev estimator, provided that $(Z_n)_{n\geq 0}$ can be observed. In Section 3 we present some applications of our results in statistics. The remaining sections are devoted to the proofs of theorems.

2. Main results

Assume that the total populations $(Z_k)_{k\geq 0}$ of all generations can be observed. For $n_0, n \in \mathbb{N}$, recall the definition of $M_{n_0,n}$:

$$M_{n_0,n} = \frac{\sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}(Z_{k+1}/Z_k - m)}{\sqrt{\sum_{k=n_0}^{n_0+n-1} Z_k(Z_{k+1}/Z_k - m)^2}}.$$

Here n_0 may depend on n. For instance, we can take n_0 as a function of n. We may take $n_0 = 0$. However, in real-world applications it may happen that we know historical data $(Z_k)_{n_0\leq k\leq n_0+n}$ for some $n_0 \geq 2$, as well as the increment n of generation numbers, but do not know the data $(Z_k)_{0\leq k\leq n_0-1}$. In such a case $M_{0,n}$ is no longer applicable to estimating m, whereas $M_{n_0,n}$ is suitable. Motivated by this problem, it would be better to consider the more general case $n_0 \geq 0$ instead of taking $n_0 = 0$. As $(Z_k)_{k=n_0,\ldots,n_0+n}$ can be observed, $M_{n_0,n}$ can be regarded as a time-type self-normalized process for the Lotka–Nagaev estimator Z_{k+1}/Z_k. The following theorem gives a self-normalized Cramér moderate deviation result for the Galton–Watson processes.

Theorem 2.1. *Assume that* $\mathbb{E}Z_1^{2+\rho} < \infty$ *for some* $\rho \in (0, 1]$.

(i) *If* $\rho \in (0, 1)$, *then for all* $x \in [0,\ o(\sqrt{n}))$,

$$\left| \ln \frac{\mathbb{P}(M_{n_0,n} \geq x)}{1 - \Phi(x)} \right| \leq C_\rho \left(\frac{x^{2+\rho}}{n^{\rho/2}} + \frac{(1+x)^{1-\rho(2+\rho)/4}}{n^{\rho(2-\rho)/8}} \right), \tag{2}$$

where C_ρ *depends only on the constants* ρ, v *and* $\mathbb{E}Z_1^{2+\rho}$.

(ii) *If* $\rho = 1$, *then for all* $x \in [0,\ o(\sqrt{n}))$,

$$\left| \ln \frac{\mathbb{P}(M_{n_0,n} \geq x)}{1 - \Phi(x)} \right| \leq C \left(\frac{x^3}{\sqrt{n}} + \frac{\ln n}{\sqrt{n}} + \frac{(1+x)^{1/4}}{n^{1/8}} \right), \tag{3}$$

where C *depends only on the constants* v *and* $\mathbb{E}Z_1^3$.

In particular, inequalities (2) *and* (3) *together imply that*

$$\frac{\mathbb{P}(M_{n_0,n} \geq x)}{1 - \Phi(x)} = 1 + o(1) \tag{4}$$

uniformly for $x \in [0,\ o(n^{\rho/(4+2\rho)}))$ *as* $n \to \infty$. *Moreover, the same inequalities remain valid when*

$$\frac{\mathbb{P}(M_{n_0,n} \geq x)}{1 - \Phi(x)}$$

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

is replaced by

$$\frac{\mathbb{P}(M_{n_0,n} \leq -x)}{\Phi(-x)}.$$

Notice that the mean of a standard normal random variable is 0. By the maximum likelihood method, it is natural to let $M_{n_0,n} = 0$; then we have

$$\sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}\left(\frac{Z_{k+1}}{Z_k} - m\right) = 0,$$

which implies that

$$\overline{m}_n := \frac{1}{\sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}} \sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}\left(\frac{Z_{k+1}}{Z_k}\right)$$

can be regarded as a random weighted Lotka–Nagaev estimator for m.

Equality (4) implies that $\mathbb{P}(M_{n_0,n} \leq x) \to \Phi(x)$ as n tends to ∞. Thus Theorem 2.1 implies the central limit theory for $M_{n_0,n}$. Moreover, equality (4) states that the relative error of normal approximation for $M_{n_0,n}$ tends to zero uniformly for $x \in [0,\ \mathrm{o}(n^{\rho/(4+2\rho)}))$ as $n \to \infty$.

Theorem 2.1 implies the following moderate deviation principle (MDP) result for the time-type self-normalized Lotka–Nagaev estimator.

Corollary 2.1. *Assume the conditions of Theorem 2.1. Let $(a_n)_{n\geq 1}$ be any sequence of real numbers satisfying $a_n \to \infty$ and $a_n/\sqrt{n} \to 0$ as $n \to \infty$. Then, for each Borel set B,*

$$-\inf_{x\in B^o} \frac{x^2}{2} \leq \liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{M_{n_0,n}}{a_n} \in B\right) \leq \limsup_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{M_{n_0,n}}{a_n} \in B\right) \leq -\inf_{x\in \overline{B}} \frac{x^2}{2},$$

where B^o and $\overline{B}$ denote the interior and the closure of B, respectively.

Remark 2.1. From (2) and (3), it is easy to derive the following Berry–Esseen bound for the self-normalized Lotka–Nagaev estimator:

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(M_{n_0,n} \leq x) - \Phi(x)| \leq \frac{C_\rho}{n^{\rho(2-\rho)/8}},$$

where C_ρ depends only on the constants ρ, v and $\mathbb{E}Z_1^{2+\rho}$. When $\rho > 1$, by the self-normalized Berry–Esseen bound for martingales in Fan and Shao [6], we can get a Berry–Esseen bound of order $n^{-\rho/(6+2\rho)}$.

The last remark gives a self-normalized Berry–Esseen bound for the Lotka–Nagaev estimator, while the next theorem presents a normalized Berry–Esseen bound for the Lotka–Nagaev estimator. Denote

$$H_{n_0,n} = \frac{1}{\sqrt{nv}} \sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}\left(\frac{Z_{k+1}}{Z_k} - m\right).$$

Notice that the random variables $(X_{k,i})_{1\leq i\leq Z_k}$ have the same distribution as Z_1, and that $(X_{k,i})_{1\leq i\leq Z_k}$ are independent of Z_k. Then, for the Galton–Watson processes, it holds that

$$\mathbb{E}[(Z_{k+1} - mZ_k)^2 \mid Z_k] = \mathbb{E}\left[\left(\sum_{i=1}^{Z_k} (X_{k,i} - m)\right)^2 \,\middle|\, Z_k\right] = Z_k v^2.$$

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

It is easy to see that

$$H_{n_0,n} = \sum_{k=n_0}^{n_0+n-1} \frac{1}{\sqrt{nv^2/Z_k}} \left(\frac{Z_{k+1}}{Z_k} - m \right).$$

Thus $H_{n_0,n}$ can be regarded as a normalized process for the Lotka–Nagaev estimator Z_{k+1}/Z_k. We have the following normalized Berry–Esseen bounds for the Galton–Watson processes.

Theorem 2.2. *Assume the conditions of Theorem 2.1.*

(i) *If* $\rho \in (0, 1)$, *then*

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(H_{n_0,n} \le x) - \Phi(x)| \le \frac{C_\rho}{n^{\rho/2}}, \tag{5}$$

where C_ρ *depends only on* ρ, v *and* $\mathbb{E}Z_1^{2+\rho}$.

(ii) *If* $\rho = 1$, *then*

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(H_{n_0,n} \le x) - \Phi(x)| \le C\frac{\ln n}{\sqrt{n}}, \tag{6}$$

where C *depends only on* v *and* $\mathbb{E}Z_1^3$.

Moreover, the same inequalities remain valid when $H_{n_0,n}$ *is replaced by* $-H_{n_0,n}$.

The convergence rates of (5) and (6) are identical to the best possible convergence rates of the Berry–Esseen bounds for martingales; see Theorem 2.1 of Fan [5] and the associated comment. Notice that $H_{n_0,n}$ is a martingale with respect to the natural filtration.

3. Applications

Cramér moderate deviations certainly have many applications in statistics.

3.1. *p*-value for hypothesis testing

Self-normalized Cramér moderate deviations can be applied to hypothesis testing of m for the Galton–Watson processes. When $(Z_k)_{k=n_0,\dots,n_0+n}$ can be observed, we can use Theorem 2.1 to estimate the p-value. Assume that $\mathbb{E}Z_1^{2+\rho} < \infty$ for some $0 < \rho \le 1$, and that $m > 1$. Let $(z_k)_{k=n_0,\dots,n_0+n}$ be the observed value of the $(Z_k)_{k=n_0,\dots,n_0+n}$. In order to estimate the offspring mean m, we can make use of the Harris estimator [2] given by

$$\widehat{m}_n = \frac{\sum_{k=n_0}^{n_0+n-1} Z_{k+1}}{\sum_{k=n_0}^{n_0+n-1} Z_k}.$$

Then the observation for the Harris estimator is

$$\widehat{m}_n = \frac{\sum_{k=n_0}^{n_0+n-1} z_{k+1}}{\sum_{k=n_0}^{n_0+n-1} z_k}.$$

By Theorem 2.1, it is easy to see that

$$\frac{\mathbb{P}(M_{n_0,n} \ge x)}{1-\Phi(x)} = 1 + \mathrm{o}(1) \quad \text{and} \quad \frac{\mathbb{P}(M_{n_0,n} \le -x)}{1-\Phi(x)} = 1 + \mathrm{o}(1) \tag{7}$$

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

uniformly for $x \in [0, \mathrm{o}(n^{\rho/(4+2\rho)}))$. Notice that $1 - \Phi(x) = \Phi(-x)$. Thus, when $|\widetilde{m}_n| = \mathrm{o}(n^{\rho/(4+2\rho)})$, by (7), the probability $\mathbb{P}(M_{n_0,n} > |\widetilde{m}_n|)$ is almost equal to $2\Phi(-|\widetilde{m}_n|)$, where

$$\widetilde{m}_n = \frac{\sum_{k=n_0}^{n_0+n-1} \sqrt{z_k}(z_{k+1}/z_k - \widehat{m}_n)}{\sqrt{\sum_{k=n_0}^{n_0+n-1} z_k(z_{k+1}/z_k - \widehat{m}_n)^2}}.$$

3.2. Construction of confidence intervals

Assume the data $(Z_k)_{k\geq 0}$ can be observed. Cramér moderate deviations can also be applied to the construction of confidence intervals of m. We use Theorem 2.1 to construct confidence intervals.

Proposition 3.1. *Assume that* $\mathbb{E}Z_1^{2+\rho} < \infty$ *for some* $\rho \in (0, 1]$. *Let* $\kappa_n \in (0, 1)$. *Assume that*

$$|\ln \kappa_n| = \mathrm{o}\big(n^{\rho/(2+\rho)}\big). \tag{8}$$

Let

$$a_{n_0,n} = \left(\sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}\right)^2 - (\Phi^{-1}(1-\kappa_n/2))^2 \sum_{k=n_0}^{n_0+n-1} Z_k,$$

$$b_{n_0,n} = 2(\Phi^{-1}(1-\kappa_n/2))^2 \sum_{k=n_0}^{n_0+n-1} Z_{k+1} - 2\left(\sum_{k=n_0}^{n_0+n-1} \frac{Z_{k+1}}{\sqrt{Z_k}}\right)\left(\sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}\right),$$

$$c_{n_0,n} = \left(\sum_{k=n_0}^{n_0+n-1} \frac{Z_{k+1}}{\sqrt{Z_k}}\right)^2 - (\Phi^{-1}(1-\kappa_n/2))^2 \sum_{k=n_0}^{n_0+n-1} \frac{Z_{k+1}^2}{Z_k}.$$

Then $[A_{n_0,n}, B_{n_0,n}]$, *with*

$$A_{n_0,n} = \frac{-b_{n_0,n} - \sqrt{b_{n_0,n}^2 - 4a_{n_0,n}c_{n_0,n}}}{2a_{n_0,n}}$$

and

$$B_{n_0,n} = \frac{-b_{n_0,n} + \sqrt{b_{n_0,n}^2 - 4a_{n_0,n}c_{n_0,n}}}{2a_{n_0,n}},$$

is a $1-\kappa_n$ *confidence interval for m, for n large enough.*

Proof. Notice that $1 - \Phi(x) = \Phi(-x)$. Theorem 2.1 implies that

$$\frac{\mathbb{P}(M_{n_0,n} \geq x)}{1-\Phi(x)} = 1 + \mathrm{o}(1) \quad \text{and} \quad \frac{\mathbb{P}(M_{n_0,n} \leq -x)}{1-\Phi(x)} = 1 + \mathrm{o}(1) \tag{9}$$

uniformly for $0 \leq x = \mathrm{o}(n^{\rho/(4+2\rho)})$; see (4). Notice that the inverse function Φ^{-1} of a standard normal distribution function Φ has the following asymptotic expansion:

$$\Phi^{-1}(1-p) = \sqrt{\ln(1/p^2) - \ln\ln(1/p^2) - \ln(2\pi)} + \mathrm{o}(p), \quad p \searrow 0.$$

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

In particular, this says that for any positive sequence $(\kappa_n)_{n\geq 1}$ that converges to zero, as $n\to\infty$, we have

$$\Phi^{-1}(1-\kappa_n/2)=\sqrt{2|\ln\kappa_n|}+\mathrm{o}(\sqrt{|\ln\kappa_n|}).$$

Thus, when κ_n satisfies the condition (8), the upper $(\kappa_n/2)$th quantile of a standard normal distribution is of order $\mathrm{o}(n^{\rho/(4+2\rho)})$. Then, applying (9) to the last equality, we complete the proof of Proposition 3.1. Notice that $A_{n_0,n}$ and $B_{n_0,n}$ are solutions of the following equation:

$$\frac{\sum_{k=n_0}^{n_0+n-1}\sqrt{Z_k}(Z_{k+1}/Z_k-x)}{\sqrt{\sum_{k=n_0}^{n_0+n-1}Z_k(Z_{k+1}/Z_k-x)^2}}=\Phi^{-1}(1-\kappa_n/2).$$

This completes the proof of Proposition 3.1. □

Assume the parameter v^2 is known. When v^2 is known, we can apply normalized Berry–Esseen bounds (see Theorem 2.2) to construct confidence intervals.

Proposition 3.2. *Assume that* $\mathbb{E}Z_1^{2+\rho}<\infty$ *for some* $\rho\in(0,1]$. *Let* $\kappa_n\in(0,1)$. *Assume that*

$$|\ln\kappa_n|=\mathrm{o}(\ln n). \tag{10}$$

Then $[A_n, B_n]$, *with*

$$A_n=\frac{\sum_{k=n_0}^{n_0+n}Z_{k+1}/\sqrt{Z_k}-\sqrt{n}v\Phi^{-1}(1-\kappa_n/2)}{\sum_{k=n_0}^{n_0+n}\sqrt{Z_k}}$$

and

$$B_n=\frac{\sum_{k=n_0}^{n_0+n}Z_{k+1}/\sqrt{Z_k}+\sqrt{n}v\Phi^{-1}(1-\kappa_n/2)}{\sum_{k=n_0}^{n_0+n}\sqrt{Z_k}},$$

is a $1-\kappa_n$ *confidence interval for m, for n large enough.*

Proof. Theorem 2.2 implies that

$$\frac{\mathbb{P}(H_{n_0,n}\geq x)}{1-\Phi(x)}=1+\mathrm{o}(1)\quad\text{and}\quad\frac{\mathbb{P}(H_{n_0,n}\leq -x)}{1-\Phi(x)}=1+\mathrm{o}(1) \tag{11}$$

uniformly for $0\leq x=\mathrm{o}(\sqrt{\ln n})$. The upper $(\kappa_n/2)$th quantile of a standard normal distribution satisfies

$$\Phi^{-1}(1-\kappa_n/2)=\mathrm{O}(\sqrt{|\ln\kappa_n|}),$$

which, by (10), is of order $\mathrm{o}(\sqrt{\ln n})$. Proposition 3.2 follows from applying (11) to $H_{n_0,n}$. □

4. Proof of Theorem 2.1

In the proof of Theorem 2.1 we will use the following lemma (see Corollary 2.3 of Fan *et al.* [7]), which gives self-normalized Cramér moderate deviations for martingales.

Lemma 4.1. *Let* $(\eta_k,\mathcal{F}_k)_{k=1,\ldots,n}$ *be a finite sequence of martingale differences. Assume that there exist a constant* $\rho\in(0,1]$ *and numbers* $\gamma_n>0$ *and* $\delta_n\geq 0$ *satisfying* $\gamma_n,\delta_n\to 0$ *such that for all* $1\leq i\leq n$,

$$\mathbb{E}[|\eta_k|^{2+\rho}\mid\mathcal{F}_{k-1}]\leq\gamma_n^{\rho}\mathbb{E}[\eta_k^2\mid\mathcal{F}_{k-1}] \tag{12}$$

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

and

$$\left|\sum_{k=1}^{n}\mathbb{E}[\eta_k^2 \mid \mathcal{F}_{k-1}]-1\right| \leq \delta_n^2 \quad a.s. \tag{13}$$

Denote

$$V_n=\frac{\sum_{k=1}^{n}\eta_k}{\sqrt{\sum_{k=1}^{n}\eta_k^2}}$$

and

$$\widehat{\gamma}_n(x,\rho)=\frac{\gamma_n^{\rho(2-\rho)/4}}{1+x^{\rho(2+\rho)/4}}.$$

(i) *If* $\rho\in(0,1)$*, then for all* $0\leq x=\mathrm{o}(\gamma_n^{-1})$*,*

$$\left|\ln\frac{\mathbb{P}(V_n\geq x)}{1-\Phi(x)}\right|\leq C_\rho\big(x^{2+\rho}\gamma_n^\rho+x^2\delta_n^2+(1+x)(\delta_n+\widehat{\gamma}_n(x,\rho))\big).$$

(ii) *If* $\rho=1$*, then for all* $0\leq x=\mathrm{o}(\gamma_n^{-1})$*,*

$$\left|\ln\frac{\mathbb{P}(V_n\geq x)}{1-\Phi(x)}\right|\leq C\big(x^3\gamma_n+x^2\delta_n^2+(1+x)(\delta_n+\gamma_n|\ln\gamma_n|+\widehat{\gamma}_n(x,1))\big).$$

Now we are in a position to prove Theorem 2.1. Denote

$$\hat{\xi}_{k+1}=\sqrt{Z_k}(Z_{k+1}/Z_k-m),$$

$\mathcal{F}_{n_0}=\{\emptyset,\Omega\}$ and $\mathcal{F}_{k+1}=\sigma\{Z_i\colon n_0\leq i\leq k+1\}$ for all $k\geq n_0$. Notice that $X_{k,i}$ is independent of Z_k. Then it is easy to verify that

$$\begin{aligned}\mathbb{E}[\hat{\xi}_{k+1}\mid\mathcal{F}_k]&=Z_k^{-1/2}\mathbb{E}[Z_{k+1}-mZ_k\mid\mathcal{F}_k]\\&=Z_k^{-1/2}\sum_{i=1}^{Z_k}\mathbb{E}[X_{k,i}-m\mid\mathcal{F}_k]\\&=Z_k^{-1/2}\sum_{i=1}^{Z_k}\mathbb{E}[X_{k,i}-m]\\&=0.\end{aligned}$$

Thus $(\hat{\xi}_k,\mathcal{F}_k)_{k=n_0+1,\dots,n_0+n}$ is a finite sequence of martingale differences. Notice that $X_{k,i}-m$, $i\geq1$, are centered and independent random variables. Thus the following equalities hold:

$$\begin{aligned}\sum_{k=n_0}^{n_0+n-1}\mathbb{E}[\hat{\xi}_{k+1}^2\mid\mathcal{F}_k]&=\sum_{k=n_0}^{n_0+n-1}Z_k^{-1}\mathbb{E}[(Z_{k+1}-mZ_k)^2\mid\mathcal{F}_k]\\&=\sum_{k=n_0}^{n_0+n-1}Z_k^{-1}\mathbb{E}\left[\left(\sum_{i=1}^{Z_k}(X_{k,i}-m)\right)^2\,\middle|\,\mathcal{F}_k\right]\\&=\sum_{k=n_0}^{n_0+n-1}Z_k^{-1}Z_k\mathbb{E}[(X_{k,i}-m)^2]\\&=nv^2.\end{aligned}$$

https://doi.org/10.1017/jpr.2022.134 Published online by Cambridge University Press

Moreover, it is easy to see that

$$\mathbb{E}[|\hat{\xi}_{k+1}|^{2+\rho} \mid \mathcal{F}_k] = Z_k^{-1-\rho/2}\mathbb{E}[|Z_{k+1} - mZ_k|^{2+\rho} \mid \mathcal{F}_k]$$
$$= Z_k^{-1-\rho/2}\mathbb{E}\left[\left|\sum_{i=1}^{Z_k}(X_{k,i} - m)\right|^{2+\rho} \,\middle|\, \mathcal{F}_k\right]. \tag{14}$$

By Rosenthal's inequality, we have

$$\mathbb{E}\left[\left|\sum_{i=1}^{Z_k}(X_{k,i} - m)\right|^{2+\rho} \,\middle|\, \mathcal{F}_k\right] \le C'_\rho\left(\left(\sum_{i=1}^{Z_k}\mathbb{E}(X_{k,i} - m)^2\right)^{1+\rho/2} + \sum_{i=1}^{Z_k}\mathbb{E}|X_{k,i} - m|^{2+\rho}\right)$$
$$\le C'_\rho\big(Z_k^{1+\rho/2}v^{2+\rho} + Z_k\mathbb{E}|Z_1 - m|^{2+\rho}\big).$$

Since the set of extinction of the process $(Z_k)_{k\ge 0}$ is negligible with respect to the annealed law $\mathbb{P}$, we have $Z_k \ge 1$ for any k. From (14), by the last inequality and the fact $Z_k \ge 1$, we deduce that

$$\mathbb{E}[|\hat{\xi}_{k+1}|^{2+\rho} \mid \mathcal{F}_k] \le C'_\rho\big(v^\rho + \mathbb{E}|Z_1 - m|^{2+\rho}/v^2\big)v^2$$
$$= C'_\rho\big(v^\rho + \mathbb{E}|Z_1 - m|^{2+\rho}/v^2\big)\mathbb{E}[\hat{\xi}^2_{k+1} \mid \mathcal{F}_k]$$
$$\le C'_\rho\big(v^\rho + 2^{1+\rho}\big(\mathbb{E}Z_1^{2+\rho} + m^{2+\rho}\big)/v^2\big)\mathbb{E}[\hat{\xi}^2_{k+1} \mid \mathcal{F}_k].$$

By Jensen's inequality, we have $m^{2+\rho} = (\mathbb{E}Z_1)^{2+\rho} \le \mathbb{E}Z_1^{2+\rho}$. Thus we have

$$\mathbb{E}[|\hat{\xi}_{k+1}|^{2+\rho} \mid \mathcal{F}_k] \le C_\rho\big(v^\rho + \mathbb{E}Z_1^{2+\rho}/v^2\big)\mathbb{E}[\hat{\xi}^2_{k+1} \mid \mathcal{F}_k].$$

Let $\eta_k = \hat{\xi}_{n_0+k}/(\sqrt{n}v)$ and $\mathcal{F}_k = \mathcal{F}_{n_0+k}$. Then $(\eta_k, \mathcal{F}_k)_{k=1,\ldots,n}$ is a martingale difference sequence and satisfies conditions (12) and (13) with $\delta_n = 0$ and

$$\gamma_n = \big(C_\rho\big(v^\rho + \mathbb{E}Z_1^{2+\rho}/v^2\big)\big)^{1/\rho}/(\sqrt{n}v).$$

Clearly,

$$M_{n_0,n} = \frac{\sum_{k=1}^n \eta_k}{\sqrt{\sum_{k=1}^n \eta_k^2}}.$$

Applying Lemma 4.1 to $(\eta_k, \mathcal{F}_k)_{k=1,\ldots,n}$, we obtain the desired inequalities. Notice that for any $\rho \in (0, 1]$ and all $x \ge 0$, the following inequality holds:

$$\frac{1+x}{1+x^{\rho(2+\rho)/4}} \le C_\rho(1+x)^{1-\rho(2+\rho)/4}.$$

5. Proof of Corollary 2.1

We only give a proof of Corollary 2.1 for $\rho \in (0, 1)$. The proof for $\rho = 1$ is similar. We first show that for any Borel set $B \subset \mathbb{R}$,

$$\limsup_{n\to\infty}\frac{1}{a_n^2}\ln\mathbb{P}\left(\frac{M_{n_0,n}}{a_n} \in B\right) \le -\inf_{x\in\overline{B}}\frac{x^2}{2}. \tag{15}$$

When $B=\emptyset$, the last inequality is obvious, with $-\inf_{x\in\emptyset} x^2/2=-\infty$. Thus we may assume that $B\neq\emptyset$. Let $x_0=\inf_{x\in B}|x|$. Clearly, we have $x_0\geq\inf_{x\in\overline{B}}|x|$. Then, by Theorem 2.1, it follows that for $\rho\in(0,\,1)$ and $a_n=\mathrm{o}(\sqrt{n})$,

$$\mathbb{P}\bigg(\frac{M_{n_0,n}}{a_n}\in B\bigg)\leq\mathbb{P}(|M_{n_0,n}|\geq a_n x_0)$$
$$\leq 2(1-\Phi(a_n x_0))\exp\bigg\{C_\rho\bigg(\frac{(a_n x_0)^{2+\rho}}{n^{\rho/2}}+\frac{(1+a_n x_0)^{1-\rho(2+\rho)/4}}{n^{\rho(2-\rho)/8}}\bigg)\bigg\}.$$

Using the inequalities

$$\frac{1}{\sqrt{2\pi}(1+x)}\mathrm{e}^{-x^2/2}\leq 1-\Phi(x)\leq\frac{1}{\sqrt{\pi}(1+x)}\mathrm{e}^{-x^2/2},\qquad x\geq 0, \tag{16}$$

and the fact that $a_n\to\infty$ and $a_n/\sqrt{n}\to 0$, we obtain

$$\limsup_{n\to\infty}\frac{1}{a_n^2}\ln\mathbb{P}\bigg(\frac{M_{n_0,n}}{a_n}\in B\bigg)\leq-\frac{x_0^2}{2}\leq-\inf_{x\in\overline{B}}\frac{x^2}{2},$$

which gives (15).

Next we prove that

$$\liminf_{n\to\infty}\frac{1}{a_n^2}\ln\mathbb{P}\bigg(\frac{M_{n_0,n}}{a_n}\in B\bigg)\geq-\inf_{x\in B^o}\frac{x^2}{2}. \tag{17}$$

When $B^o=\emptyset$, the last inequality is obvious, with $-\inf_{x\in\emptyset} x^2/2=-\infty$. Thus we may assume that $B^o\neq\emptyset$. Since B^o is an open set, for any given small $\varepsilon_1>0$ there exists an $x_0\in B^o$ such that

$$0<\frac{x_0^2}{2}\leq\inf_{x\in B^o}\frac{x^2}{2}+\varepsilon_1.$$

Again by the fact that B^o is an open set, for $x_0\in B^o$ and all sufficiently small $\varepsilon_2\in(0,\,|x_0|]$, we have $(x_0-\varepsilon_2,\,x_0+\varepsilon_2]\subset B^o$. Without loss of generality, we may assume that $x_0>0$. Clearly, we have

$$\mathbb{P}\bigg(\frac{M_{n_0,n}}{a_n}\in B\bigg)\geq\mathbb{P}(M_{n_0,n}\in(a_n(x_0-\varepsilon_2),\,a_n(x_0+\varepsilon_2)])$$
$$=\mathbb{P}(M_{n_0,n}\geq a_n(x_0-\varepsilon_2))-\mathbb{P}(M_{n_0,n}\geq a_n(x_0+\varepsilon_2)x). \tag{18}$$

Again by Theorem 2.1, it is easy to see that for $a_n\to\infty$ and $a_n=\mathrm{o}(\sqrt{n})$,

$$\lim_{n\to\infty}\frac{\mathbb{P}(M_{n_0,n}\geq a_n(x_0+\varepsilon_2))}{\mathbb{P}(M_{n_0,n}\geq a_n(x_0-\varepsilon_2))}=0.$$

From (18), by the last line and Theorem 2.1, for all n large enough and $a_n=\mathrm{o}(\sqrt{n})$ it holds that

$$\mathbb{P}\bigg(\frac{M_{n_0,n}}{a_n}\in B\bigg)\geq\frac{1}{2}\mathbb{P}(M_{n_0,n}\geq a_n(x_0-\varepsilon_2))$$
$$\geq\frac{1}{2}(1-\Phi(a_n(x_0-\varepsilon_2)))\exp\bigg\{-C_\rho\bigg(\frac{(a_n x_0)^{2+\rho}}{n^{\rho/2}}+\frac{(1+a_n x_0)^{1-\rho(2+\rho)/4}}{n^{\rho(2-\rho)/8}}\bigg)\bigg\}.$$

Using (16) and the fact that $a_n \to \infty$ and $a_n/\sqrt{n} \to 0$, after some calculations we get

$$\liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{M_{n_0,n}}{a_n} \in B\right) \geq -\frac{1}{2}(x_0 - \varepsilon_2)^2.$$

Letting $\varepsilon_2 \to 0$, we deduce that

$$\liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{M_{n_0,n}}{a_n} \in B\right) \geq -\frac{x_0^2}{2} \geq -\inf_{x\in B^o} \frac{x^2}{2} - \varepsilon_1.$$

Since ε_1 can be arbitrarily small, we get (17). Combining (15) and (17), we complete the proof of Corollary 2.1.

6. Proof of Theorem 2.2

Recall the martingale differences $(\eta_k, \mathcal{F}_k)_{k=1,\ldots,n}$ defined in the proof of Theorem 2.1. Then η_k satisfies conditions (12) and (13) with $\delta_n = 0$ and

$$\gamma_n = \left(C_\rho\left(v^\rho + \mathbb{E}Z_1^{2+\rho}/v^2\right)\right)^{1/\rho}/\sqrt{n}v.$$

Clearly, we have $H_{n_0,n} = \sum_{k=1}^{n} \eta_k$. Applying Theorem 2.1 of Fan [5] to $(\eta_k, \mathcal{F}_k)_{k=1,\ldots,n}$, we obtain the desired inequalities.

Acknowledgements

The authors would like to thank the two referees for their helpful remarks and suggestions.

Funding information

The research is partially supported by the National Nature Science Foundation of China NSFC (grants 12031005 and 11971063) and Shenzhen Outstanding Talents Training Fund, China.

Competing interests

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

References

[1] Athreya, K. B. (1994). Large deviation rates for branching processes I: Single type case. *Ann. Appl. Prob.* **4**, 779–790.

[2] Bercu, B. and Touati, A. (2008). Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Prob.* **18**, 1848–1869.

[3] Chu W. (2018). Self-normalized large deviation for supercritical branching processes. *J. Appl. Prob.* **55**, 450–458.

[4] De la Peña, V. H., Lai, T. L. and Shao, Q. M. (2009). *Self-Normalized Processes, Limit Theorems and Statistical Applications*. Springer.

[5] Fan, X. (2019). Exact rates of convergence in some martingale central limit theorems. *J. Math. Anal. Appl.* **469**, 1028–1044.

[6] Fan, X. and Shao, Q. M. (2018). Berry–Esseen bounds for self-normalized martingales. *Commun. Math. Statist.* **6**, 13–27.

[7] Fan, X., Grama, I., Liu, Q. and Shao, Q. M. (2019). Self-normalized Cramér type moderate deviations for martingales. *Bernoulli* **25**, 2793–2823.

[8] FAN, X., GRAMA, I., LIU, Q. AND SHAO, Q. M. (2020). Self-normalized Cramér type moderate deviations for stationary sequences and applications. *Stoch. Process. Appl.* **130**, 5124–5148.
[9] HE, H. (2016). On large deviation rates for sums associated with Galton–Watson processes. *Adv. Appl. Prob.* **48**, 672–690.
[10] JING, B. Y., SHAO, Q. M. AND WANG, Q. (2003). Self-normalized Cramér-type large deviations for independent random variables. *Ann. Prob.* **31**, 2167–2215.
[11] KUELBS, J. AND VIDYASHANKAR, A. N. (2011). Weak convergence results for multiple generations of a branching process. *J. Theor. Prob.* **24**, 376–396.
[12] LOTKA, A. (1939). Théorie analytique des associations biologiques. *Actualités Sci. Ind.* **780**, 123–136.
[13] MAAOUIA, F. AND TOUATI, A. (2005). Identification of multitype branching processes. *Ann. Statist.* **33**, 2655–2694.
[14] NAGAEV, S. V. (1967). On estimating the expected number of direct descendants of a particle in a branching process. *Theory Prob. Appl.* **12**, 314–320.
[15] NEY, P. E. AND VIDYASHANKAR A. N. (2003). Harmonic moments and large deviation rates for supercritical branching processes. *Ann. Appl. Prob.* **13**, 475–489.
[16] NEY, P. E. AND VIDYASHANKAR A. N. (2004). Local limit theory and large deviations for supercritical branching processes. *Ann. Appl. Prob.* **14**, 1135–1166.