# Cramér moderate deviations for a supercritical Galton-Watson process

Paul Doukhan[a], Xiequan Fan[b], Zhi-Qiang Gao[c]

[a] *CY University, AGM UMR 8088, site Saint-Martin, 95000 Cergy-Pontoise, France*
[b] *Center for Applied Mathematics, Tianjin University, Tianjin 300072, China*
[c] *Laboratory of Mathematics and Complex Systems (Ministry of Education), School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China*

**Abstract**

Let $(Z_n)_{n\geq 0}$ be a supercritical Galton-Watson process. The Lotka-Nagaev estimator $Z_{n+1}/Z_n$ is a common estimator for the offspring mean. In this paper, we establish some Cramér moderate deviation results for the Lotka-Nagaev estimator via a martingale method. Applications to construction of confidence intervals are also given.

*Keywords:* Cramér moderate deviations; Lotka-Nagaev estimator; offspring mean

*2000 MSC:* primary 60G42; 60F10; 60E15; secondary 62F03

## 1. Introduction

Let $(X_i)_{i\geq 1}$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and positive variance $\sigma^2$. Denote by $S_n = \sum_{i=1}^n X_i$ the partial sums of $(X_i)_{i\geq 1}$. Assume $\mathbb{E}\exp\{c_0|X_1|\} < \infty$ for some constant $c_0 > 0$. Cramér [3] has established the following asymptotic moderate deviation expansion: for all $0 \leq x = o(n^{1/2})$,

$$\left| \ln \frac{\mathbb{P}(S_n > x\sigma\sqrt{n})}{1-\Phi(x)} \right| = O\left(\frac{1+x^3}{\sqrt{n}}\right) \quad \text{as} \quad n \to \infty, \tag{1.1}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x \exp\{-t^2/2\}dt$ is the standard normal distribution. The results of type (1.1) are usually called as Cramér moderate deviations. After the seminal work of Cramér, a number of Cramér moderate deviations have been established for various settings. See, for instance, Linnik [11] and [6] for independent random variables, Fan, Grama and Liu [4] for martingales (see also Puhalskii [18] for large deviation principles), Grama, Liu and Miqueu [8] and Fan, Hu and Liu [5] for a supercritical branching process in a random environment, and Beknazaryan, Sang and Xiao [2] for random fields. In this paper, we are going to establish Cramér moderate deviations for a supercritical Galton-Watson process.

A Galton-Watson process is defined as follows

$$Z_0 = 1, \qquad Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}, \qquad \text{for } n \geq 0, \tag{1.2}$$

where $X_{n,i}$ is the offspring number of the $i$-th individual of the generation $n$ and $Z_n$ stands for the total population of the generation $n$. Moreover, $(X_{n,i})_{i\geq 1}$ are independent of each other with a common distribution law $\mathbb{P}(X_{n,i} = k) = p_k,\ k \geq 0$, and are also independent to $Z_n$. Denote by $m$ the offspring mean of an individual, then it holds

$$m = \mathbb{E}Z_1 = \mathbb{E}X_{n,i} = \sum_{k=0}^{\infty} kp_k, \quad n, i \geq 1.$$

*February 15, 2022*

Denote by $v$ the standard variance of $Z_1$, then we have

$$v^2 = \mathbb{E}(Z_1 - m)^2 = \mathrm{Var}(X_{n,i}) = \mathrm{Var}(Z_1). \tag{1.3}$$

To avoid triviality, we assume that $v$ is positive. The Lotka-Nagaev estimator $Z_{n+1}/Z_n$ is a common estimator for the offspring mean $m$. Throughout the paper, we assume that $p_0 = 0$, then the Lotka-Nagaev estimator $Z_{n+1}/Z_n$ is well defined $\mathbb{P}$-a.s. Athreya [1] has established large deviation rates for the Lotka-Nagaev estimator with $Z_1$ satisfying Cramér's condition. See also Ney and Vidyashankar [16, 17] with a much weaker assumption that $\mathbb{P}(Z_1 \geq x) \sim ax^{1-\alpha}, x \to \infty$, for two constants $\alpha > 2$ and $a > 0$. Fleischmann and Wachtel [7] considered a generalization of the Lotka-Nagaev estimator $S_{Z_n}/Z_n$, where $S_n = \sum_{i=1}^n X_i$ is independent of $Z_n$ and $\mathbb{P}(Z_1 \geq x) = \mathbb{P}(X_1 \geq x) \sim ax^{-\beta}, x \to \infty$, for a constant $\beta > 2$. See also He [9] when $X_1$ is in the domain of attraction of a stable law. For the Galton-Watson processes with immigration, we refer to Liu and Zhang [13] and Li and Li [12] for the rates of convergence of the Lotka-Nagaev estimator. In this paper, we establish some Cramér moderate deviation results for the Lotka-Nagaev estimator via a martingale method. Notice that the Cramér moderate deviation results for a supercritical branching process in a random environment (BPRE) stated in [8] do not implies our results, because the random environment for BPRE cannot be degenerate and they considered the estimator $\frac{1}{n} \ln Z_n$ instead of the Lotka-Nagaev estimator.

The paper is organized as follows. In Section 2, we present our main results, including Cramér moderate deviations and moderate deviation principles for the Lotka-Nagaev estimator. 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

### *2.1. The data $(Z_k)_{n_0 \leq k \leq n_0+n}$ can be observed*

Let $n_0, n \in \mathbb{N}$. Denote

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

In usual, one takes $n_0 = 0$. Here we consider the more general cases that $n_0$ may depend on $n$. It is easy to check that

$$\mathbb{E}\Big[\sqrt{Z_k}(\frac{Z_{k+1}}{Z_k} - m)\Big| Z_0, ..., Z_k\Big] = 0 \quad \text{and} \quad \mathrm{Var}\Big(\sqrt{Z_k}(\frac{Z_{k+1}}{Z_k} - m)\Big) = v^2.$$

Thus $H_{n_0,n}$ is a standardized martingale. Denote

$$\widehat{m}_n = \frac{1}{\sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}} \sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}(\frac{Z_{k+1}}{Z_k}) \tag{2.1}$$

the random weighted Lotka-Nagaev estimator. Then $H_{n_0,n}$ can be rewritten in the following form

$$H_{n_0,n} = \frac{\widehat{m}_n - m}{v\sqrt{n}} \sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}.$$

Thus $(H_{n_0,n})_{n\geq 1}$ is the standardized process for the estimator $\widehat{m}_n$. We have the following Cramér moderate deviation result with respect to $H_{n_0,n}$.

**Theorem 2.1.** *Assume that there exists a positive constant $c$ such that*

$$\mathbb{E}|Z_1 - m|^l \leq \frac{1}{2}\, l!\, (l-1)^{-l/2}\, c^{l-2}\, \mathbb{E}(Z_1 - m)^2, \qquad l \geq 2. \tag{2.2}$$

*Then the following equalities hold for all* $0 \le x = o(\sqrt{n})$,

$$\left| \ln \frac{\mathbb{P}(H_{n_0,n} \ge x)}{1-\Phi(x)} \right| = O\bigg( \frac{x^3}{\sqrt{n}} + (1+x)\frac{\ln n}{\sqrt{n}} \bigg) \tag{2.3}$$

*and*

$$\left| \ln \frac{\mathbb{P}(H_{n_0,n} \le -x)}{\Phi(-x)} \right| = O\bigg( \frac{x^3}{\sqrt{n}} + (1+x)\frac{\ln n}{\sqrt{n}} \bigg) \tag{2.4}$$

*as* $n \to \infty$.

**Remark 2.1.** *Let us make some comments on Theorem 2.1.*

1. *It is worth noting that if* $Z_1 \le m + c_2$*, then condition (2.2) is satisfied with* $c = \frac{1}{3} 2^{3/2} \max\{m,\ c_2\}$.
2. *Sub-Gaussian random variable also satisfies condition (2.2), that is, if there exists a positive constant* $c_1 > 0$ *such that*

$$\mathbb{P}(Z_1 - m \ge x) \le c_1 \exp\{-x^2/c_1\}, \quad x \ge 0,$$

*then condition (2.2) is satisfied. Indeed, it is easy to see that for all* $l \ge 2$,

$$\begin{aligned} \mathbb{E}|Z_1 - m|^l &\le m^l \mathbb{P}(Z_1 - m < 0) + \int_0^\infty l\, x^{l-1} \mathbb{P}(Z_1 - m \ge x) dx \\ &\le m^l + \int_0^\infty l\, x^{l-1} c_1 \exp\{-x^2/c_1\} dx = m^l + c_1 (\sqrt{\frac{c_1}{2}})^{l-1} \int_0^\infty l\, y^{l-1} \exp\{-y^2/2\} dy \\ &\le m^l + c_1 (\sqrt{\frac{c_1}{2}})^{l-1} (l-1)!!. \end{aligned}$$

*By Stirling's formula* $n! = \sqrt{2\pi n}\, n^n e^{-n} e^{\frac{1}{12 n \theta_n}}$ *for some* $0 \le \theta_n \le 1$*, we deduce that for all* $l \ge 2$,

$$(l-1)!! \le \sqrt{l!} \le l!\,(l-1)^{-l/2} 2^{l-2}.$$

*Thus, we have* $\mathbb{E}|Z_1 - m|^l \le m^l + l!(l-1)^{-l/2} c_1 (\sqrt{\frac{c_1}{2}})^{l-1} 2^{l-2}$*, which implies (2.2) with* $c$ *large enough.*

It is easy to see that for all $n \ge 3$,

$$\frac{x^3}{\sqrt{n}} + (1+x)\frac{\ln n}{\sqrt{n}} \le 2(1+x^3)\frac{\ln n}{\sqrt{n}}, \quad x \ge 0.$$

Using the inequality $|e^x - 1| \le e^\alpha |x|$ valid for $|x| \le \alpha$, from Theorem 2.1, we obtain the following result about the equivalence to the normal tail.

**Corollary 2.1.** *Assume the condition of Theorem 2.1. Then for all* $n \ge 3$ *and all* $0 \le x \le n^{1/6}$,

$$\frac{\mathbb{P}(H_{n_0,n} \ge x)}{1-\Phi(x)} = 1 + O\bigg( \frac{x^3}{\sqrt{n}} + (1+x)\frac{\ln n}{\sqrt{n}} \bigg) \quad \text{and} \quad \frac{\mathbb{P}(H_{n_0,n} \le -x)}{\Phi(-x)} = 1 + O\bigg( \frac{x^3}{\sqrt{n}} + (1+x)\frac{\ln n}{\sqrt{n}} \bigg). \tag{2.5}$$

*In particular, it implies that*

$$\frac{\mathbb{P}(H_{n_0,n} \ge x)}{1-\Phi(x)} = 1 + o(1) \quad \text{and} \quad \frac{\mathbb{P}(H_{n_0,n} \le -x)}{\Phi(-x)} = 1 + o(1) \tag{2.6}$$

*holds uniformly for* $0 \le x = o(n^{1/6})$ *as* $n \to \infty$.

Theorem 2.1 also implies the following moderate deviation principle (MDP) result. An analogy for a BPRE, but with respect to $\ln Z_n$, we refer to Huang and Liu [10].

**Corollary 2.2.** *Assume the condition of Theorem 2.1. Let $(a_n)_{n\geq 1}$ be any sequence of real positive numbers satisfying $a_n \to \infty$ and $a_n/\sqrt{n} \to 0$ as $n \to \infty$. Then for each Borel set $B$,*

$$-\inf_{x\in B^o} \frac{x^2}{2} \leq \liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\Big(\frac{H_{n_0,n}}{a_n} \in B\Big) \leq \limsup_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\Big(\frac{H_{n_0,n}}{a_n} \in B\Big) \leq -\inf_{x\in \overline{B}} \frac{x^2}{2}, \tag{2.7}$$

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

The proof of Corollary 2.2 is given in Section 5.

*2.2. The data $Z_{n+1}$ and $Z_n$ can be observed*

For the Galton-Watson process, it holds

$$\mathbb{E}[(Z_{n+1} - mZ_n)^2|Z_n] = \mathbb{E}[(\sum_{i=1}^{Z_n}(X_{n,i} - m))^2|Z_n] = Z_n\upsilon^2.$$

Thus the following one

$$R_n = \frac{\sqrt{Z_n}}{\upsilon}\Big(\frac{Z_{n+1}}{Z_n} - m\Big)$$

is a normalized process for the Lotka-Nagaev estimator. When $Z_1$ satisfies the Cramér condition (cf. (2.8)), we have the following Cramér moderate deviation result for the normalized Lotka-Nagaev estimator $R_n$.

**Theorem 2.2.** *Assume there exists a constant $\kappa_0 > 0$ such that*

$$\mathbb{E}\exp\{\kappa_0 Z_1\} < \infty. \tag{2.8}$$

*Then*

$$\Big|\ln \frac{\mathbb{P}(R_n \geq x)}{1 - \Phi(x)}\Big| = O\Big(\frac{1 + x^3}{\sqrt{n}}\Big) \tag{2.9}$$

*holds uniformly for $0 \leq x = o(\sqrt{n})$ as $n \to \infty$. In particular, it implies that*

$$\frac{\mathbb{P}\big(R_n \geq x\big)}{1 - \Phi(x)} = 1 + o(1) \tag{2.10}$$

*holds uniformly for $0 \leq x = o(n^{1/6})$ as $n \to \infty$.*

Clearly, the ranges of validity for (2.9) and (2.10) coincide with the case of classical Cramér moderate deviation result [3].

As $Z_1 \geq 0$, we still have the following Cramér moderate deviation result for the normalized Lotka-Nagaev estimator $R_n$ under a weaker moment condition.

**Theorem 2.3.** *Assume that $\mathbb{E}Z_1^{2+\rho} < \infty$ for some $\rho \in (0, 1]$. Then*

$$\Big|\ln \frac{\mathbb{P}(R_n \leq -x)}{\Phi(-x)}\Big| = O\Big(\frac{1 + x^{2+\rho}}{n^{\rho/2}}\Big) \tag{2.11}$$

*holds uniformly for $0 \leq x = o(\sqrt{n})$ as $n \to \infty$. In particular, it implies that*

$$\frac{\mathbb{P}\big(R_n \leq -x\big)}{\Phi(-x)} = 1 + o(1) \tag{2.12}$$

*holds uniformly for $0 \leq x = o(n^{\rho/(4+2\rho)})$ as $n \to \infty$.*

Clearly, condition (2.8) implies that $\mathbb{E}Z_1^3 < \infty$. Thus, with condition (2.8), Theorem 2.3 implies that (2.11) holds with $\rho = 1$. By an argument similar to the proof of Corollary 2.2, we have following MDP result for $R_n$.

**Corollary 2.3.** *Assume the condition of Theorem 2.2. Let $(a_n)_{n\geq 1}$ be a 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 $B$,*

$$-\inf_{x\in B^o} \frac{x^2}{2} \leq \liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\Big(\frac{R_n}{a_n} \in B\Big) \leq \limsup_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\Big(\frac{R_n}{a_n} \in B\Big) \leq -\inf_{x\in \overline{B}} \frac{x^2}{2}, \tag{2.13}$$

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

Under the Linnik condition [11] (instead of Cramér's condition (2.8)), we have the following Cramér type moderate deviation result for the normalized Lotka-Nagaev estimator $R_n$.

**Theorem 2.4.** *Assume that there exist two constants $\iota_0 > 0$ and $\tau \in (0, \frac{1}{6}]$ such that*

$$\mathbb{E}\exp\{\iota_0 Z_1^{\frac{4\tau}{2\tau+1}}\} < \infty. \tag{2.14}$$

*Then*

$$\frac{\mathbb{P}(R_n \geq x)}{1-\Phi(x)} = 1 + o(1) \tag{2.15}$$

*holds uniformly for $x \in [0, o(n^\tau))$ as $n \to \infty$.*

Inequality (2.15) states that the relative error of normal approximation for $R_n$ tends to zero uniformly for $x \in [0, o(n^\tau))$. The range of validity for (2.15) coincides with the Cramér moderate deviation of Linnik [11] for i.i.d. random variables.

## 3. Applications to construction of confidence intervals

*3.1. The data $(Z_k)_{n_0\leq k\leq n_0+n}$ can be observed*

Cramér moderate deviation results can be applied to construction of confidence intervals for $m$. Recall $\widehat{m}_n$ defined by (2.1). By Theorem 2.1, we have the following result for the confidence interval for $m$.

**Proposition 3.1.** *Assume the condition of Theorem 2.1. Let $\kappa_n \in (0,1)$. Assume*

$$\big|\ln \kappa_n\big| = o\big(n^{1/3}\big). \tag{3.1}$$

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

$$A_{n_0,n} = \widehat{m}_n - \frac{v\sqrt{n}\,\Phi^{-1}(1-\kappa_n/2)}{\sum_{k=n_0}^{n_0+n-1}\sqrt{Z_k}} \quad \text{and} \quad B_{n_0,n} = \widehat{m}_n + \frac{v\sqrt{n}\,\Phi^{-1}(1-\kappa_n/2)}{\sum_{k=n_0}^{n_0+n-1}\sqrt{Z_k}},$$

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

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

$$\frac{\mathbb{P}(H_{n_0,n} \geq x)}{1-\Phi(x)} = 1 + o(1) \quad \text{and} \quad \frac{\mathbb{P}(H_{n_0,n} \leq -x)}{\Phi(-x)} = 1 + o(1) \tag{3.2}$$

uniformly for $0 \leq x = o(n^{1/6})$. Notice that the inverse function $\Phi^{-1}$ of the the standard normal distribution function $\Phi$ satisfies the following asymptotic expansion

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

By (5.2) and (3.1), it is easy to see that the upper $(\kappa_n/2)$th quantile of a standard normal distribution $\Phi^{-1}(1-\kappa_n/2) = -\Phi^{-1}(\kappa_n/2) = O(\sqrt{|\ln \kappa_n|})$ is of order $o\big(n^{1/6}\big)$. Then applying the last equality to (3.2), we have

$$\mathbb{P}\big(H_{n_0,n} \geq \Phi^{-1}(1-\kappa_n/2)\big) \sim \kappa_n/2 \quad \text{and} \quad \mathbb{P}\big(H_{n_0,n} \leq -\Phi^{-1}(1-\kappa_n/2)\big) \sim \kappa_n/2$$

as $n \to \infty$. Clearly, $H_{n_0,n} \leq \Phi^{-1}(1-\kappa_n/2)$ means that $m \geq A_{n_0,n}$, while $H_{n_0,n} \geq -\Phi^{-1}(1-\kappa_n/2)$ means $m \leq B_{n_0,n}$. This completes the proof of Proposition 3.1. □

*3.2. The data $Z_{n+1}$ and $Z_n$ can be observed*

When $Z_{n+1}$ and $Z_n$ can be observed, we can make use of Theorem 2.4 to construct confidence intervals.

**Proposition 3.2.** *Assume the condition of Theorem 2.4. Let $\kappa_n \in (0,1)$. Assume*

$$\big|\ln \kappa_n\big| = o\big(n^{2\tau}\big). \tag{3.3}$$

*Let*

$$\Delta_n = \frac{\Phi^{-1}(1-\kappa_n/2)}{Z_n} \upsilon.$$

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

$$A_n = \frac{Z_{n+1}}{Z_n} - \Delta_n \quad \textit{and} \quad B_n = \frac{Z_{n+1}}{Z_n} + \Delta_n,$$

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

*Proof.* Theorem 2.4 implies that

$$\frac{\mathbb{P}(R_n \geq x)}{1-\Phi\left(x\right)} = 1 + o(1) \quad \text{and} \quad \frac{\mathbb{P}(R_n \leq -x)}{1-\Phi\left(x\right)} = 1 + o(1) \tag{3.4}$$

uniformly for $0 \leq x = o(n^{\tau})$. When $\kappa_n$ satisfies the condition (3.3), the upper $(\kappa_n/2)$th quantile of a standard normal distribution satisfies $\Phi^{-1}(1-\kappa_n/2) = O(\sqrt{|\ln \kappa_n|})$, which is of order $o\big(n^{\tau}\big)$. Using (3.4), by an argument similar to the proof of Proposition 3.1, we obtain the $1-\kappa_n$ confidence interval for $m$. □

## 4. Proof of Theorem 2.1

Let $(\xi_i, \mathcal{F}_i)_{1\leq i\leq n}$ be a finite sequence of martingale differences. In the sequel we shall use the following conditions:

**(A1)** There exists a number $\epsilon_n \in (0, \frac{1}{2}]$ such that

$$|\mathbb{E}[\xi_i^k|\mathcal{F}_{i-1}]| \leq \frac{1}{2}k!\epsilon_n^{k-2}\mathbb{E}[\xi_i^2|\mathcal{F}_{i-1}], \qquad \text{for all } k \geq 3 \ \text{ and } \ 1 \leq i \leq n;$$

**(A2)** There exists a number $\delta_n \in [0, \frac{1}{2}]$ such that $\big|\sum_{i=1}^n \mathbb{E}[\xi_i^2|\mathcal{F}_{i-1}] - 1\big| \leq \delta_n^2$.

In the proof of Theorem 2.1, we make use of the following lemma which gives a Cramér moderate deviation result for martingales. See Theorems 2.1 and 2.2 of Fan, Grama and Liu [4].

**Lemma 4.1.** *Assume conditions (A1) and (A2). Then there exits an absolute constant $\alpha \in (0,1)$ such that for all $0 \leq x \leq \alpha\, \epsilon_n^{-1}$ and $\delta_n \leq \alpha$,*

$$\left| \ln \frac{\mathbb{P}(\sum_{i=1}^n \xi_i \geq x)}{1-\Phi\left(x\right)} \right| \leq C_\alpha \bigg( x^3 \epsilon_n + x^2 \delta_n^2 + \, (1+x) \left( \epsilon_n \left| \ln \epsilon_n \right| + \delta_n \right) \bigg)$$

*and*

$$\left| \ln \frac{\mathbb{P}(\sum_{i=1}^n \xi_i \leq -x)}{\Phi\left(-x\right)} \right| \leq C_\alpha \bigg( x^3 \epsilon_n + x^2 \delta_n^2 + \, (1+x) \left( \epsilon_n \left| \ln \epsilon_n \right| + \delta_n \right) \bigg),$$

*where the constant $C_\alpha$ does not depend on $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$, $n$ and $x$.*

Denote

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

$\mathfrak{F}_{n_0} = \{\emptyset, \Omega\}$ and $\mathfrak{F}_{k+1} = \sigma\{Z_i : n_0 \leq i \leq k+1\}$ for all $k > n_0$. Notice that $X_{k,i}$ is independent of $Z_k$. Then it is easy to verify that $\mathbb{E}[\hat{\xi}_{k+1}|\mathfrak{F}_k] = 0$. Thus $(\hat{\xi}_k, \mathfrak{F}_k)_{k=n_0+1,...,n_0+n}$ is a finite sequence of martingale differences. Notice that $X_{k,i} - m, i \geq 1$, are centered and independent random variables. Thus, the following equalities hold

$$\mathbb{E}[\hat{\xi}_{k+1}^2|\mathfrak{F}_k] = v^2 \quad \text{and} \quad \sum_{k=n_0}^{n_0+n-1} \mathbb{E}[\hat{\xi}_{k+1}^2|\mathfrak{F}_k] = nv^2. \tag{4.1}$$

By Rio's inequality (cf. Theorem 2.1 of Rio [19]) and the fact that $X_{k,i}$ is independent to $\mathfrak{F}_k$, we have for any $\rho \geq 1$,

$$\Big( \mathbb{E}[| \sum_{i=1}^{Z_k} (X_{k,i} - m)|^{2\rho} | \mathfrak{F}_k] \Big)^{2/(2\rho)} \quad \leq \quad (2\rho - 1) \sum_{i=1}^{Z_k} \Big( \mathbb{E}|X_{k,i} - m|^{2\rho} \Big)^{2/(2\rho)}.$$

The last inequality implies that for any $\rho \geq 1$,

$$\begin{aligned} \mathbb{E}[| \sum_{i=1}^{Z_k} (X_{k,i} - m)|^{2\rho} | \mathfrak{F}_k] \quad &\leq \quad (2\rho - 1)^\rho \Big( \sum_{i=1}^{Z_k} \Big( \mathbb{E}|X_{k,i} - m|^{2\rho} \Big)^{1/\rho} \Big)^\rho \\ &\leq \quad (2\rho - 1)^\rho Z_k^\rho \mathbb{E}|X_{k,1} - m|^{2\rho}. \end{aligned}$$

Hence, the following inequalities hold for any $\rho \geq 1$,

$$\begin{aligned} \mathbb{E}[|\hat{\xi}_{k+1}|^{2\rho}|\mathfrak{F}_k] \quad &= \quad Z_k^{-\rho} \mathbb{E}[|Z_{k+1} - mZ_k|^{2\rho}|\mathfrak{F}_k] = Z_k^{-\rho} \mathbb{E}[| \sum_{i=1}^{Z_k} (X_{k,i} - m)|^{2\rho} | \mathfrak{F}_k] \\ &\leq \quad (2\rho - 1)^\rho \mathbb{E}|X_{k,1} - m|^{2\rho}. \end{aligned}$$

The last inequality becomes equality when $\rho = 1$. Notice that $X_{k,1}$ has the same distribution as $Z_1$. Thus, by condition (2.2), we get for all $l \geq 2$,

$$\begin{aligned} \mathbb{E}[|\hat{\xi}_{k+1}|^l|\mathfrak{F}_k] \quad &\leq \quad (l-1)^{l/2} \mathbb{E}|X_{k,1} - m|^l \leq (l-1)^{l/2} \frac{1}{2}\, l!\; (l-1)^{-l/2}\, c^{l-2}\, \mathbb{E}(X_{k,1} - m)^2 \\ &= \quad \frac{1}{2}\, l!\, c^{l-2}\, v^2 \;\; = \;\; \frac{1}{2}\, l!\, c^{l-2}\, \mathbb{E}[\hat{\xi}_{k+1}^2|\mathfrak{F}_k]. \end{aligned}$$

Set $\xi_k = \hat{\xi}_{n_0+k}/\sqrt{n}v$ and $\mathcal{F}_k = \mathfrak{F}_{n_0+k}$. It is easy to see that conditions (A1) and (A2) are satisfied with $\epsilon_n = c/\sqrt{n}v$ and $\delta_n = 0$. Applying Lemma 4.1 to $(\xi_k, \mathcal{F}_k)_{1 \leq k \leq n}$, we obtain the desired inequalities.

## 5. Proof of Corollary 2.2

We first show that for any Borel set $B \subset \mathbb{R}$, it holds

$$\limsup_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\Big( \frac{H_{n_0,n}}{a_n} \in B \Big) \leq - \inf_{x \in \overline{B}} \frac{x^2}{2}. \tag{5.1}$$

When $B = \emptyset$, the last inequality holds with $-\inf_{x\in\emptyset} \frac{x^2}{2} = -\infty$. Hence, we only need to consider the case of $B \neq \emptyset$. Let $x_0 = \inf_{x\in B} |x|$, then we have $x_0 = \inf_{x\in\overline{B}} |x|$. Then, from Theorem 2.1, it follows that for $a_n = o(\sqrt{n})$,

$$\begin{aligned} \mathbb{P}\Big( \frac{H_{n_0,n}}{a_n} \in B \Big) &\leq \mathbb{P}\Big( |H_{n_0,n}| \geq a_n x_0 \Big) \\ &\leq 2\Big(1 - \Phi(a_n x_0)\Big) \exp\bigg\{ C\bigg( \frac{(a_n x_0)^3}{\sqrt{n}} + (1 + (a_n x_0)) \frac{\ln n}{\sqrt{n}} \bigg) \bigg\}. \end{aligned}$$

Using the following two-sided bound for the normal distribution function

$$\frac{1}{\sqrt{2\pi}(1+x)} e^{-x^2/2} \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{\pi}(1+x)} e^{-x^2/2}, \quad x \geq 0, \tag{5.2}$$

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

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

which gives (5.1).

Next, we show that the following inequality holds

$$\liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\Big( \frac{H_{n_0,n}}{a_n} \in B \Big) \geq - \inf_{x \in B^o} \frac{x^2}{2}. \tag{5.3}$$

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

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

By the fact that $B^o$ is an open set again, for $x_0 \in B^o$ and any $\varepsilon_2 \in (0, |x_0|]$, it holds $(x_0 - \varepsilon_2, x_0 + \varepsilon_2] \subset B^o$. Without loss of generality, we may assume that $x_0 > 0$. Then, we have

$$\begin{aligned} \mathbb{P}\Big( \frac{H_{n_0,n}}{a_n} \in B \Big) &\geq \mathbb{P}\Big( H_{n_0,n} \in (a_n(x_0 - \varepsilon_2), a_n(x_0 + \varepsilon_2)] \Big) \\ &= \mathbb{P}\Big( H_{n_0,n} \geq a_n(x_0 - \varepsilon_2) \Big) - \mathbb{P}\Big( H_{n_0,n} \geq a_n(x_0 + \varepsilon_2) \Big). \end{aligned} \tag{5.4}$$

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

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

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

$$\begin{aligned}\mathbb{P}\Big(\frac{H_{n_0,n}}{a_n} \in B\Big) &\geq \frac{1}{2}\mathbb{P}\Big(H_{n_0,n} \geq a_n(x_0-\varepsilon_2)\Big) \\ &\geq \frac{1}{2}\Big(1-\Phi\left(a_n(x_0-\varepsilon_2)\right)\Big)\exp\Big\{-C\Big(\frac{(a_n(x_0-\varepsilon_2))^3}{\sqrt{n}}+(1+a_n(x_0-\varepsilon_2))\frac{\ln n}{\sqrt{n}}\Big)\Big\}.\end{aligned}$$

Using (5.2) and the fact that $a_n \to \infty$ and $a_n/\sqrt{n} \to 0$, after some simple calculations, we obtain

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

Letting $\varepsilon_2 \to 0$, we arrive at

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

Since that $\varepsilon_1$ can be arbitrarily small, we get (5.3). Combining (5.1) and (5.3) together, we obtain the desired result. This completes the proof of Corollary 2.2. □

## 6. Proof of Theorem 2.2

*6.1. Preliminary lemmas*

Denote by $f_n(s) = \mathbb{E}s^{Z_n}$, $|s| \leq 1$, the generating function of $Z_n$. In the proof of Theorem 2.2, we shall make use of the following lemma, see Athreya [1].

**Lemma 6.1.** *If $p_1 > 0$, then it holds*

$$\lim_{n\to\infty}\frac{f_n(s)}{p_1^n} = \sum_{k=1}^{\infty} q_k s^k, \tag{6.1}$$

*where $(q_k, k \geq 1)$ is defined by the generating function $Q(s) = \sum_{k=1}^{\infty} q_k s^k, 0 \leq s < 1$, the unique solution of the following functional equation*

$$Q(f(s)) = p_1 Q(s), \qquad \textit{where} \ \ f(s) = \sum_{j=1}^{\infty} p_j s^j, \ \ 0 \leq s < 1,$$

*subject to*

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

By Lemma 6.1, we obtain the following estimation for $Z_n$.

**Lemma 6.2.** *It holds*

$$\mathbb{P}(Z_n \leq n) \leq C_1 \exp\{-nc_0\}. \tag{6.2}$$

*Proof.* When $p_1 > 0$, by Markov's inequality, we deduce that for $s_0 = \frac{1+p_1}{2} \in (0,1)$,

$$\mathbb{P}(Z_n \leq n) = \mathbb{P}(s_0^{Z_n} \geq s_0^n) \leq s_0^{-n} f_n(s_0).$$

Using Lemma 6.1, we have

$$\begin{aligned}\mathbb{P}(Z_n \leq n) &\leq C(\frac{p_1}{s_0})^n Q(s_0) \\ &= C_1 \exp\{-nc_0\},\end{aligned} \tag{6.3}$$

where $C_1 = CQ(s_0)$ and $c_0 = \ln(s_0/p_1)$. Notice that $s_0 \in (p_1, 1)$, which implies that $c_0 > 0$. When $p_1 = p_0 = 0$, we have $Z_n \geq 2^n$, and (6.2) holds obviously for all $n$. □

In the proof of Theorem 2.2, we also make use of the following lemma of Cramér [3].

**Lemma 6.3.** *Let $(X_i)_{i\geq 1}$ be i.i.d. and centered random variables. Assume that $\mathbb{E}\exp\{\lambda|X_1|\}<\infty$ for some constant $\lambda>0$. Set $S_n=\sum_{i=1}^n X_i$ and $v^2=\mathbb{E}X_1^2$. Then*

$$\left|\ln\frac{\mathbb{P}(S_n/(v\sqrt{n})\geq x)}{1-\Phi(x)}\right|\leq C\frac{1+x^3}{\sqrt{n}} \tag{6.4}$$

*uniformly for $0\leq x=o(\sqrt{n})$.*

*6.2. Proof of Theorem 2.2*

By the definition of $R_n$, it is easy to see that $R_n$ can be rewritten as follows:

$$R_n=\frac{1}{v\sqrt{Z_n}}\Big(Z_{n+1}-mZ_n\Big)=\frac{1}{v\sqrt{Z_n}}\sum_{i=1}^{Z_n}\Big(X_{n,i}-m\Big).$$

By the total probability formula, we have

$$\begin{aligned}\mathbb{P}\Big(R_n\geq x\Big) &= \sum_{k=1}^{\infty}\mathbb{P}(Z_n=k)\mathbb{P}\bigg(\frac{1}{\sqrt{k}v}\sum_{i=1}^{k}\Big(X_{n,i}-m\Big)\geq x\bigg)\\ &=: \sum_{k=1}^{\infty}\mathbb{P}(Z_n=k)F_k(x).\end{aligned} \tag{6.5}$$

Notice that $X_{n,i},\ 1\leq i\leq k$, have the same distribution as $Z_1$ and $Z_1\geq 0$. By (2.8), it holds

$$\mathbb{E}e^{\kappa_0|Z_1-m|}\leq e^{\kappa_0 m}+e^{-\kappa_0 m}\mathbb{E}e^{\kappa_0 Z_1}<\infty.$$

When $k\geq n$, by condition (2.8) and Lemma 6.3, we get

$$\left|\ln\frac{F_k(x)}{1-\Phi(x)}\right|\leq C_1\frac{1+x^3}{\sqrt{k}}\leq C_1\frac{1+x^3}{\sqrt{n}} \tag{6.6}$$

uniformly for $0\leq x=o(\sqrt{n})$. Returning to (6.5), by (6.6), we have for all $0\leq x=o(\sqrt{n})$,

$$\begin{aligned}\mathbb{P}\Big(R_n\geq x\Big) &\geq \sum_{k=n}^{\infty}\mathbb{P}(Z_n=k)F_k(x)\geq\Big(1-\Phi(x)\Big)\exp\Big\{-C_1\frac{1+x^3}{\sqrt{n}}\Big\}\sum_{k=n}^{\infty}\mathbb{P}(Z_n=k)\\ &\geq \Big(1-\Phi(x)\Big)\exp\Big\{-C_1\frac{1+x^3}{\sqrt{n}}\Big\}\Big(1-\mathbb{P}(Z_n\leq n)\Big).\end{aligned} \tag{6.7}$$

By Lemma 6.2, we have

$$\mathbb{P}(Z_n\leq n)\leq C_2\exp\{-C_3 n\}. \tag{6.8}$$

Applying the last inequality to (6.7), we obtain for all $0\leq x=o(\sqrt{n})$,

$$\begin{aligned}\mathbb{P}\Big(R_n\geq x\Big) &\geq \Big(1-\Phi(x)\Big)\exp\Big\{-C_1\frac{1+x^3}{\sqrt{n}}\Big\}\Big(1-C_2\exp\{-C_3 n\}\Big)\\ &\geq \Big(1-\Phi(x)\Big)\exp\Big\{-C_4\frac{1+x^3}{\sqrt{n}}\Big\}.\end{aligned} \tag{6.9}$$

Returning to (6.5), by (6.6) and (6.8), we deduce that for all $0 \leq x = o(\sqrt{n})$,

$$\begin{aligned}
\mathbb{P}\Big(R_n \geq x\Big) &\leq \sum_{k=1}^{n-1} \mathbb{P}(Z_n = k) F_k(x) + \sum_{k=n}^{\infty} \mathbb{P}(Z_n = k) F_k(x) \\
&\leq \mathbb{P}(Z_n \leq n-1) + \Big(1 - \Phi(x)\Big) \exp\Big\{ C_1 \frac{1+x^3}{\sqrt{n}} \Big\} \sum_{k=n}^{\infty} \mathbb{P}(Z_n = k) \\
&\leq C_2 \exp\{-C_3 n\} + \Big(1 - \Phi(x)\Big) \exp\Big\{ C_1 \frac{1+x^3}{\sqrt{n}} \Big\} \\
&\leq \Big(1 - \Phi(x)\Big) \exp\Big\{ C_4 \frac{1+x^3}{\sqrt{n}} \Big\}. \qquad (6.10)
\end{aligned}$$

Combining (6.9) and (6.10) together, we obtain the desired inequality, that is (2.9).

## 7. Proofs of Theorems 2.3 and 2.4

The proof of Theorem 2.3 is similar to the proof of Theorem 2.2. However, instead of using Lemma 6.3, we should make use of the following lemma of Fan [6].

**Lemma 7.1.** *Let $(X_i)_{i\geq 1}$ be i.i.d. and centered random variables. Assume that $X_1 \geq -C$ and $\mathbb{E}|X_1|^{2+\rho} < \infty$ for some constants $C > 0$ and $\rho \in (0, 1]$. Let $S_n = \sum_{i=1}^n X_i$ and $\upsilon^2 = \mathbb{E}X_1^2$. Then*

$$\Big| \ln \frac{\mathbb{P}(S_n/(\upsilon\sqrt{n}) \leq -x)}{\Phi(-x)} \Big| \leq C \frac{1 + x^{2+\rho}}{n^{\rho/2}}$$

*holds uniformly for $0 \leq x = o(\sqrt{n})$.*

The proof of Theorem 2.4 is analogous to the proof of Theorem 2.2. However, instead of using Lemma 6.3, we should make use of the following lemma of Linnik [11].

**Lemma 7.2.** *Let $(X_i)_{i\geq 1}$ be i.i.d. and centered random variables. Assume that $\mathbb{E}\exp\{\iota_0 |X_1|^{\frac{4\tau}{2\tau+1}}\} < \infty$ for two constants $\iota_0 > 0$ and $\tau \in (0, \frac{1}{6}]$. Let $S_n = \sum_{i=1}^n X_i$ and $\upsilon^2 = \mathbb{E}X_1^2$. Then*

$$\frac{\mathbb{P}(S_n/(\upsilon\sqrt{n}) \geq x)}{1 - \Phi(x)} = 1 + o(1)$$

*holds uniformly for $0 \leq x = o(n^\tau)$.*

*Acknowledgements*

We would like to thank the two referees for their helpful remarks and suggestions. This work has been partially supported by the National Natural Science Foundation of China (Grant No. 11971063), by CY (Grant "Investissements d'Avenir" ANR-16-IDEX-0008) "EcoDep" PSI-AAP2020-0000000013) and Labex MME-DII (ANR-11-LABEX-0023-01).

## References

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

[2] Beknazaryan, A., Sang, H., Xiao, Y. (2019) Cramér type moderate deviations for random fields. *J. Appl. Probab.* **56**(1): 223–245.

[3] Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualite's Sci. Indust.* **736**: 5–23.

[4] Fan, X., Grama, I., Liu, Q. (2013). Cramér large deviation expansions for martingales under Bernstein's condition. *Stochastic Process. Appl.* **123**(11): 3919–3942.

[5] Fan, X., Hu, H., Liu, Q. (2020). Uniform Cramér moderate deviations and Berry-Esseen bounds for a supercritical branching process in a random environment. *Front. Math. China* **15**(5): 891–914

[6] Fan, X. (2017). Sharp large deviation results for sums of bounded from above random variables. *Sci. China Math.* **60**(12), 2465–2480.

[7] Fleischmann, K., Wachtel, V. (2008). Large deviations for sums indexed by the generations of a Galton-Watson process. *Probab. Theory and Related Fields* **141**: 445–470.

[8] Grama, I., Liu, Q., Miqueu, M. (2017). Berry-Esseen's bound and Cramér's large deviations for a supercritical branching process in a random environment. *Stochastic Process. Appl.* **127**, 1255–1281.

[9] He, H. (2016). On large deviation rates for sums associated with Galton-Watson processes, *Adv. Appl. Probab.* **48**(2): 672–690.

[10] Huang, C., Liu, Q. (2012). Moments, moderate and large deviations for a branching process in a random environment. *Stochastic Process. Appl.*, **122**, 522–545.

[11] Linnik, Y.V. (1961). On the probability of large deviations for the sums of independent variables. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 2, pp. 289-306). Univ of California Press.

[12] Li, L., Li, J. (2021). Large deviation rates for supercritical branching processes with immigration. *J. Theoret. Probab.* **34**(1): 162–172.

[13] Liu, J.N., Zhang, M. (2016). Large deviation for supercritical branching processes with immigration. *Acta Math. Sinica, English Series* **32**(8): 893–900.

[14] Lotka, A. (1939). Theorie analytique des assiciation biologiques. *Actualités Sci. Ind.* **780**: 123–136.

[15] Nagaev, S.V. (1967). On estimating the expected number of direct descendants of a particle in a branching process. *Theory Probab. Appl.* **12**: 314–320.

[16] Ney, P.E., Vidyashankar A.N. (2003). Harmonic moments and large deviation rates for supercritical branching processes. *Ann. Appl. Probab.* **13**: 475–489.

[17] Ney, P.E., Vidyashankar A.N. (2004). Local limit theory and large deviations for supercritical branching processes. *Ann. Appl. Probab.* **14**: 1135–1166.

[18] Puhalskii, A. (1997). Large deviations of semimartingales: a maxingale problem approach I. Limits as solutions to a maxingale problem. *Stochastics* **61**(3-4): 141–243.

[19] Rio, E. (2009). Moment inequalities for sums of dependent random variables under projective conditions, *J. Theoret. Probab.* **22** (1): 146–163.