The Annals of Statistics
2013, Vol. 41, No. 2, 641–669
DOI: 10.1214/13-AOS1091

MODERATE DEVIATIONS FOR A NONPARAMETRIC ESTIMATOR OF SAMPLE COVERAGE

BY FUQING GAO[1]

Wuhan University

In this paper, we consider moderate deviations for Good's coverage estimator. The moderate deviation principle and the self-normalized moderate deviation principle for Good's coverage estimator are established. The results are also applied to the hypothesis testing problem and the confidence interval for the coverage.

1. Introduction. Let $X_k(n)$ be the frequency of the kth species in a random sample of size n from a multinomial population with a perhaps countably infinite number of species and let P_n be probability measures under which the kth species has probability p_{kn} of being sampled, where $p_n = (p_{kn}; k \geq 1)$ with $\sum_{k=1}^{\infty} p_{kn} = 1$. Let Q_n and $F_j(n)$ denote the sum of the probabilities of the unobserved species, and the total number of species represented j times in the sample, respectively, that is,

$$Q_n = \sum_{k=1}^{\infty} p_{kn}\delta_{k0}(n), \qquad F_j(n) = \sum_{k=1}^{\infty} \delta_{kj}(n), \tag{1.1}$$

where $\delta_{kj}(n) = I_{\{X_k(n)=j\}}$. Then $1 - Q_n$ is called the sample coverage which is the sum of the probabilities of the observed species. Good (1953) proposed the estimator

$$\hat{Q}_n = \frac{F_1(n)}{n} \tag{1.2}$$

for Q_n.

The Good estimator $\hat{Q}_n$ has many applications such as Shakespeare's general vocabulary and authorship of a poem [Efron and Thisted (1976), Thisted and Efron (1987)], genom [Mao and Lindsay (2002)], the probability of discovering new species in a population [Good and Toulmin (1956), Chao (1981)], network species and data confidentiality [Zhang (2005)]. Lladser, Gouet and Reeder (2011) considered the problem of predicting Q_n. They studied prediction and prediction intervals, and gave a real-data example.

Received August 2011; revised January 2013.
[1]Supported by the National Natural Science Foundation of China (10871153, 11171262).
MSC2010 subject classifications. 62G20, 62G10, 60F10.
Key words and phrases. Sample coverage, moderate deviations, Good's estimator.

On the theoretical aspects, many authors studied the asymptotic properties [cf. Esty (1982, 1983), Orlitsky, Santhanam and Zhang (2003), and Zhang and Zhang (2009) and references therein]. Esty (1983) proved the following asymptotic normality:

$$\lim_{n\to\infty} P_n\left(\frac{n(\hat{Q}_n - Q_n)}{\sqrt{b(n)}} \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du, \qquad x \in \mathbb{R}, \tag{1.3}$$

under the condition

$$\lim_{n\to\infty} \frac{E_n(F_1(n))}{n} = c_1 \in (0,1) \quad \text{and} \quad \lim_{n\to\infty} \frac{E_n(F_2(n))}{n} = c_2 \in [0,\infty), \tag{1.4}$$

where

$$b(n) = E_n\big(F_1(n)\big)\big(1 - E_n\big(F_1(n)\big)/n\big) + 2E_n\big(F_2(n)\big). \tag{1.5}$$

Recently, Zhang and Zhang (2009) found a necessary and sufficient condition for the asymptotic normality (1.3) under the condition

$$\limsup_{n\to\infty} \frac{E_n(F_1(n))}{n} < 1, \tag{1.6}$$

that is, under condition (1.6), (1.3) holds if and only if both

$$\lim_{n\to\infty}\big(E_n\big(F_1(n)\big) + E_n\big(F_2(n)\big)\big) = \infty \tag{1.7}$$

and for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{\infty} (np_{kn})^2 e^{-np_{kn}} I_{\{np_{kn} > \varepsilon s_n\}} = 0, \tag{1.8}$$

where for any $\lambda > 0$,

$$s_{\lambda n}^2 = \sum_{k=1}^{\infty} \big(\lambda p_{kn} e^{-\lambda p_{kn}} + (\lambda p_{kn})^2 e^{-\lambda p_{kn}}\big) \quad \text{and} \quad s_n = s_{nn}. \tag{1.9}$$

In this paper, we consider the moderate deviation problem for the Good estimator. It is known that the moderate deviation principle is a basic problem. It provides us with rates of convergence and a useful method for constructing asymptotic confidence intervals. The moderate deviations can be applied to the following nonparameter hypothesis testing problem:

$$H_0 : P_n = P_n^{(0)} \quad \text{and} \quad H_1 : P_n = P_n^{(1)},$$

where $P_n^{(0)}$ and $P_n^{(1)}$ are two probability measures under which the kth species has, respectively, probability $p_{kn}^{(0)}$ and $p_{kn}^{(1)}$ of being sampled, where $p_n^{(i)} = (p_{kn}^{(i)}; k \ge 1)$ with $\sum_{k=1}^{\infty} p_{kn}^{(i)} = 1$, $i = 0, 1$. We can define a rejection region of the hypothesis testing by the moderate deviation principle such that the probabilities of type I and type II errors tend to 0 with an exponential speed. The asymptotic normality provides $\sqrt{b(n)}$ as the asymptotic variance and approximate confidence statements, but it does not prove that the probabilities of type I and type II errors tend to 0 with an exponential speed. The moderate deviations can be applied to a hypothesis testing problem for the expected coverage of the sample.

Gao and Zhao (2011) have established a general delta method on the moderate deviations for estimators. But the method cannot be applied to the Good estimator. In order to study the moderate deviation problem for the Good estimator, we need refined asymptotic analysis techniques and tail probability estimates. The exponential moments inequalities, the truncation method, asymptotic analysis techniques and the Poisson approximation in Zhang and Zhang (2009) play important roles. Our main results are a moderate deviation principle and a self-normalized moderate deviation principle for the Good estimator.

The rest of this paper is organized as follows. The main results are stated in Section 2. Some examples and applications to the hypothesis testing problem and the confidence interval are also given in Section 2. The proofs of the main results are given in Section 3. Some basic concepts for large deviations and the proofs of several technique lemmas are given in the Appendix.

2. Main results and their applications. In this section, we state the main results and give some examples and applications.

2.1. *Main results.* Let $a(t)$, $t \geq 0$, be a function taking values in $[1, +\infty)$ such that

$$\lim_{t\to\infty} \frac{a(t)}{\sqrt{t}} = \infty, \qquad \lim_{t\to\infty} \frac{a(t)}{t} = 0. \tag{2.1}$$

We introduce the following Lindeberg-type condition: for any positive sequence $\{\lambda_n, n \geq 1\}$ with $\lambda_n/n \to 1$ and any $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{\infty} (\lambda_n p_{kn})^2 e^{-\lambda_n p_{kn}} I_{\{\lambda_n p_{kn} > \varepsilon s_n^2/a(s_n^2)\}} = 0. \tag{2.2}$$

REMARK 2.1. For any $L \geq 1$,

$$\begin{aligned}
&\sum_{k=1}^{\infty} (\lambda_n p_{kn})^2 e^{-\lambda_n p_{kn}} I_{\{\lambda_n p_{kn} > L\}} \\
&\quad \leq \sum_{j=0}^{\infty} L 2^{j+1} \exp\{-L 2^j\} \sum_{k=1}^{\infty} \lambda_n p_{kn} I_{\{L2^j \leq \lambda_n p_{kn} < L 2^{j+1}\}} \\
&\quad \leq 8\lambda_n L \exp\{-L\}.
\end{aligned}$$

In particular, take $L = \frac{\varepsilon s_n^2}{a(s_n^2)}$. If $\lim_{n\to\infty} \frac{s_n^2}{a(s_n^2)\log(\lambda_n/s_n^2)} = \infty$, then (2.2) holds.

THEOREM 2.1 (Moderate deviation principle). *Suppose that the conditions* (1.6), (1.7) *and* (2.2) *hold. Then* $\{\frac{n(\hat{Q}_n - Q_n)}{a(b(n))}, n \geq 1\}$ *satisfies a large deviation principle with speed* $\frac{a^2(b(n))}{b(n)}$ *and with rate function* $I(x) = \frac{x^2}{2}$. *In particular, for any* $r > 0$,

$$\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\frac{|n(\hat{Q}_n - Q_n)|}{a(b(n))} \geq r\right) = -\frac{r^2}{2}.$$

THEOREM 2.2 (Self-normalized moderate deviation principle). *Suppose that conditions* (1.6), (1.7) *and* (2.2) *hold. Then*

$$\left\{\frac{\sqrt{b(n)}n(\hat{Q}_n - Q_n)}{a(b(n))\sqrt{F_1(n)(1 - F_1(n)/n) + 2F_2(n)}}, n \geq 1\right\}$$

satisfies a large deviation principle with speed $\frac{a^2(b(n))}{b(n)}$ *and with rate function* $I(x) = \frac{x^2}{2}$.

REMARK 2.2. Let t_n, $n \geq 1$ be a sequence of positive numbers such that

$$t_n \uparrow \infty \quad \text{and} \quad \frac{t_n}{\sqrt{b(n)}} \downarrow 0. \tag{2.3}$$

Then Theorems 2.1 and 2.2 give the following estimates which are much easier to understand and apply:

$$P_n\left(\pm\frac{n(\hat{Q}_n - Q_n)}{\sqrt{b(n)}} \geq t_n\right) = \exp\left\{-(1 + o(1))\frac{t_n^2}{2}\right\}$$

and

$$P_n\left(\pm\frac{n(\hat{Q}_n - Q_n)}{\sqrt{F_1(n)(1 - F_1(n)/n) + 2F_2(n)}} \geq t_n\right) = \exp\left\{-(1 + o(1))\frac{t_n^2}{2}\right\}.$$

Set $u_n = E_n(Q_n) = \sum_{k=1}^{\infty} p_{kn}(1 - p_{kn})^n$. Then $1 - u_n$ is called the expected coverage of the sample in the literature. By Theorems 2.1 and 2.2, and Lemma 3.10, $\hat{Q}_n$ as an estimator of u_n also satisfies moderate deviation principles.

COROLLARY 2.1. *Suppose that conditions* (1.6), (1.7) *and* (2.2) *hold. Then* $\{\frac{n(\hat{Q}_n - u_n)}{a(b(n))}, n \geq 1\}$ *and* $\{\frac{\sqrt{b(n)}n(\hat{Q}_n - u_n)}{a(b(n))\sqrt{F_1(n)(1-F_1(n)/n)+2F_2(n)}}, n \geq 1\}$ *satisfy the large deviation principle with speed* $\frac{a^2(b(n))}{b(n)}$ *and with rate function* $I(x) = \frac{x^2}{2}$.

REMARK 2.3. Lladser, Gouet and Reeder (2011) considered the problem of predicting Q_n, and obtained conditionally unbiased predictors and exact prediction intervals based on a Poissonization argument. The moderate deviations for the predictors are also interesting problems.

2.2. *Application to hypothesis testing and confidence interval.* In this subsection, we apply the moderate deviations to hypothesis testing problems and confidence interval. Let Q_n be the unknown total probability unobserved species, and let $\hat{Q}_n$ be the estimator defined by (1.2).

First, let us consider a nonparametric hypothesis testing problem. Let $P_n^{(0)}$ and $P_n^{(1)}$ be two probability measures under which the kth species has, respectively, probability $p_{kn}^{(0)}$ and $p_{kn}^{(1)}$ of being sampled, where $p_n^{(i)} = (p_{kn}^{(i)}; k \geq 1)$ with $\sum_{k=1}^{\infty} p_{kn}^{(i)} = 1$, $i = 0, 1$. Denote by

$$u_n^{(i)} := \sum_{k=1}^{\infty} p_{kn}^{(i)} (1 - p_{kn}^{(i)})^n, \qquad i = 0, 1,$$

and

$$b^{(i)}(n) := E_n^{(i)}(F_1(n))(1 - E_n^{(i)}(F_1(n))/n) + 2E_n^{(i)}(F_2(n)), \qquad i = 0, 1.$$

Suppose that the conditions (1.6), (1.7) and (2.2) hold for $P_n^{(i)}$, $i = 0, 1$, and that

$$\liminf_{n\to\infty} |u_n^{(0)} - u_n^{(1)}| \neq 0.$$

Consider the nonparameter hypothesis testing

$$H_0 : P_n = P_n^{(0)} \quad \text{and} \quad H_1 : P_n = P_n^{(1)}.$$

We take the statistic $T_n := \hat{Q}_n - u_n^{(0)}$ as test statistic. Suppose that the rejection region for testing the null hypothesis H_0 against H_1 is $\{\frac{n}{a(b^{(0)}(n))}|T_n| \geq c\}$, where c is a positive constant. The probability α_n of type I error and the probability β_n of type II error are

$$\alpha_n = P_n^{(0)}\left(\frac{n}{a(b^{(0)}(n))}|T_n| \geq c\right), \qquad \beta_n = P_n^{(1)}\left(\frac{n}{a(b^{(0)}(n))}|T_n| < c\right),$$

respectively. It follows

$$\beta_n \leq P_n^{(1)}\left(\frac{n}{a(b^{(1)}(n))}|\hat{Q}_n - u_n^{(1)}| \geq \left(|u_n^{(0)} - u_n^{(1)}| - \frac{a(b^{(0)}(n))c}{n}\right)\frac{n}{a(b^{(1)}(n))}\right).$$

Therefore, Corollary 2.1 implies that

$$\lim_{n\to\infty} \frac{b^{(0)}(n)}{a^2(b^{(0)}(n))} \log \alpha_n = -\frac{c^2}{2}, \qquad \lim_{n\to\infty} \frac{b^{(1)}(n)}{a^2(b^{(1)}(n))} \log \beta_n = -\infty.$$

The above result tells us that if the rejection region for the test is $\{\frac{n}{a(b^{(0)}(n))}|T_n| \geq c\}$, then the probability of type I error tends to 0 with exponential decay speed $\exp\{-c^2 a^2(b^{(0)}(n))/(2b^{(0)}(n))\}$, and the probability of type II error tends to 0 with exponential decay speed $\exp\{-ra^2(b^{(1)}(n))/b^{(1)}(n)\}$ for all $r > 0$. But the asymptotic normality does not prove that the probabilities of type I and type II errors tend to 0 with an exponential speed.

We also consider a hypothesis testing problem for the expected coverage of the sample. We denote by $P_n^{u_n}$ the probability measures under which the expected coverage of the sample is $1-u_n$ and set

$$b_{u_n}(n) := E_n^{u_n}(F_1(n))\big(1 - E_n^{u_n}(F_1(n))/n\big) + 2E_n^{u_n}(F_2(n)).$$

Suppose that the conditions (1.6), (1.7) and (2.2) hold for $P_n^{u_n}$ for each $u_n > 0$. Let $0 < u_n^{(0)} \le u_n^{(1)}$ be two real numbers preassigned. Consider the hypothesis testing

$$H_0 : u_n \le u_n^{(0)} \quad \text{and} \quad H_1 : u_n > u_n^{(1)}.$$

We also take the rejection region $D_n := \{\frac{n}{a(b_{u_n^0}(n))}(\hat{Q}_n - u_n^{(0)}) \ge c\}$, where c is a positive constant. When $u_n \le u_n^{(0)}$,

$$\log P_n^{u_n}(D_n) \le \log P_n^{u_n}\left(\frac{n}{a(b_{u_n^0}(n))}(\hat{Q}_n - u_n) \ge c\right) \approx -\frac{c^2 a^2(b_{u_n^0}(n))}{2b_{u_n}(n)}$$

and when $u_n > u_n^{(1)}$,

$$\frac{b_{u_n}(n)}{a^2(b_{u_n}(n))} \log P_n^{u_n}(D_n^c) \to -\infty.$$

Next, we apply the moderate estimates to confidence intervals. For given confidence level $1-\alpha$, set $c_\alpha = \sqrt{-\frac{b(n)}{a^2(b(n))}\log\alpha}$. Then by Theorem 2.1, the $1-\alpha$ confidence interval for Q_n is $(\hat{Q}_n - \frac{a(b(n))}{n}c_\alpha, \hat{Q}_n + \frac{a(b(n))}{n}c_\alpha)$, that is,

$$\left(\hat{Q}_n - \frac{1}{n}\sqrt{-b(n)\log\alpha}, \hat{Q}_n + \frac{1}{n}\sqrt{-b(n)\log\alpha}\right).$$

But the confidence interval contains unknown $b(n)$. We use Theorem 2.2 to obtain another confidence interval with confidence level $1-\alpha$ for Q_n which does not contain unknown $b(n)$,

$$\left(\hat{Q}_n - \frac{\sqrt{-(F_1(n)(1-F_1(n)/n) + 2F_2(n))\log\alpha}}{n},\right.$$
$$\left.\hat{Q}_n + \frac{\sqrt{-(F_1(n)(1-F_1(n)/n) + 2F_2(n))\log\alpha}}{n}\right).$$

2.3. *Examples*. Let us check that some examples in Zhang and Zhang (2009) also satisfy moderate deviation principles if $a(n) = n^\gamma$, where $\gamma \in (1/2, 1)$. For a given decreasing density function $p_n(x)$ on $[0,\infty)$. Define $p_{in} = z_n p_n(i)$, where $z_n = (\sum_{i=1}^\infty p_{in})^{-1}$. Two concrete examples are as follows:

Let $p_n(x) = p(x) = a/(x+1)^b$, where $a > 0$ and $b > 1$. By Example 1 in Zhang and Zhang (2009), $E_n(F_1(n)) \asymp n^{1/b}$ and $\log s_n^2 \asymp \log n$, where

$$c_n \asymp b_n \text{ means } 0 < \liminf_{n\to\infty} \frac{c_n}{b_n} \le \limsup_{n\to\infty} \frac{c_n}{b_n} < \infty.$$

Thus (1.6) and (1.7) hold. By Remark 2.1, (2.2) also holds. Therefore, Theorems 2.1 and 2.2 hold.

Let $p_n(x) = p(x) = r_n^{-1} e^{-x/r_n}$, where $r_n/n \le c$ for some constant $c < \infty$. Then by Example 2 in Zhang and Zhang (2009), $\limsup_{n\to\infty} \frac{E_n(F_1(n))}{n} = \limsup_{n\to\infty} \int_0^1 e^{-ny/r_n}\, dy \le \int_0^1 e^{-y/c}\, dy < 1$ and $s_{\lambda_n n}^2 \asymp r_n \int_0^{\lambda_n/r_n} (1+t)e^{-t}\, dt \asymp r_n$ when $\lambda_n/n \to 1$. Thus, (2.2) is equivalent to

$$o(1) = \frac{1}{r_n} \int_{np_n(x)\ge \varepsilon r_n/a(r_n)} \bigl(\lambda_n p_n(x)\bigr)^2 e^{-\lambda_n p_n(x)}\, dx = \int_{\varepsilon r_n/a(r_n)\le t\le \lambda_n/r_n} t e^{-t}\, dt,$$

which holds if and only if $r_n \to \infty$. Therefore, (1.6), (1.7) and (2.2) hold if and only if $r_n \to \infty$.

3. Proofs of main results. In this section we give proofs of the main results. Let us explain the idea of the proof of Theorem 2.1. First, we divide the proof into two cases: case I and case II, according to the limit $\lim_{n\to\infty} E_n(F_1(n))/n \in (0,1)$ and 0. For case I, by the truncation method and the exponential equivalent method, we simplify our problems to the case which $\{np_{nk}, k\ge 1, n\ge 1\}$ is uniformly bounded. For case II, by the Poisson approximation and the exponential equivalent method, we simplify our problems to the case of independent sums satisfying an analogous Lindeberg condition. For the two cases simplified, we establish moderate deviation principles by the method of the Laplace asymptotic integral (Lemmas 3.7 and 3.8). The exponential moment estimate (Lemma 3.5) plays an important role in the proofs of some exponential equivalence (Lemmas 3.6 and 3.9). The main technique in the estimate of the Laplace asymptotic integral Lemma 3.7 is asymptotic analysis. In particular, we emphasis a transformation defined below (B.3) which plays a crucial role in the proof of Lemma 3.7.

We can assume that the population is sampled sequentially, so that $\mathbf{X}(m) - \mathbf{X}(m-1)$, $m\ge 1$, are i.i.d. multinomial$(1, p_n)$ under P_n, where $\mathbf{X}(n) = (X_k(n); k\ge 1)$ can be viewed as a multinomial $(n; p_n)$ vector under P_n, that is, for all integers $m \ge 1$,

$$P_n\bigl(X_k(n) = x_k; k = 1,\dots,m\bigr) = \frac{n!(1-\sum_{k=1}^m p_{kn})^{n-x_1-\cdots-x_m} \prod_{k=1}^m p_{kn}^{x_k}}{(n-x_1-\cdots-x_m)!x_1!\cdots x_m!}.$$

It is obvious that $E_n(F_1(n))/n \le 1$. Since for any $1\le L<n$,

$$\begin{aligned}\frac{2E_n(F_2(n))}{n-1} &\le L \sum_{np_{kn}\le L} p_{kn}(1-p_{kn})^{n-2} + \sup_{np\ge L} np(1-p)^{n-2} \\ &\le \frac{L}{1-L/n} \frac{E_n(F_1(n))}{n} + Le^{-L}\biggl(1-\frac{L}{n}\biggr)^{-2},\end{aligned}$$

we have that

$$\limsup_{n\to\infty} \frac{2E_n(F_2(n))}{n} \le \limsup_{n\to\infty} \frac{E_n(F_1(n))}{n} + e^{-1} \le 2; \tag{3.1}$$

and if $\limsup_{n\to\infty} \frac{E_n(F_1(n))}{n} = 0$, then $\limsup_{n\to\infty} \frac{E_n(F_2(n))}{n} = 0$. Without loss of generality, we can assume that

$$\lim_{n\to\infty} \frac{E_n(F_1(n))}{n} = c_1 \in [0, 1) \quad \text{and} \quad \lim_{n\to\infty} \frac{E_n(F_2(n))}{n} = c_2 \in [0, 1]. \tag{3.2}$$

Otherwise, we consider subsequence. The proof of Theorem 2.1 will be divided into two cases,

$$\text{case I:} \quad c_1 \in (0, 1); \qquad \text{case II:} \quad c_1 = 0.$$

Now let us introduce the structrue of the proofs of main results. In Section 3.1, we give several moment estimates and exponential moment inequalities which are basic for studying the moderate deviations for the Good estimator. A truncation method and some related estimates are also presented in the subsection. The proofs of cases I and II of Theorem 2.1 are given, respectively, in Sections 3.2 and 3.3. In Section 3.4, we prove Theorem 2.2. The proofs of several technique lemmas are postponed to the Appendix.

3.1. *Several moment estimates and inequalities.* For any $L \ge 1$ and $\varrho > 0$, set

$$M_n^L = \{k \ge 1; np_{kn} \le L\}, \qquad M_n^{Lc} = \{k \ge 1; np_{kn} > L\}$$

and

$$M_{n\varrho} = \{k \ge 1; np_{kn} \le \varrho b(n)/a(b(n))\},$$
$$M_{n\varrho}^c = \{k \ge 1; np_{kn} > \varrho b(n)/a(b(n))\}.$$

LEMMA 3.1. *If $c_1 \in (0, 1)$, then for any positive sequence $\{\lambda_n, n \ge 1\}$ with $\lambda_n/n \to 1$,*

$$\lim_{L\to\infty} \limsup_{n\to\infty} \frac{1}{n} \sum_{k\in M_n^{Lc}} (\lambda_n p_{kn} + (\lambda_n p_{kn})^2) e^{-\lambda_n p_{kn}} = 0. \tag{3.3}$$

In particular, condition (2.2) is valid.

PROOF. Similarly to Remark 2.1, for any $L \ge 1$,

$$\sum_{k\in M_n^{Lc}} \lambda_n p_{kn} e^{-\lambda_n p_{kn}} \le \lambda_n e^{-L}/(1 - e^{-L}),$$
$$\sum_{k\in M_n^{Lc}} (\lambda_n p_{kn})^2 e^{-\lambda_n p_{kn}} \le 8L\lambda_n \exp\{-L\}.$$

Therefore, (3.3) holds. $\square$

REMARK 3.1. From Lemma 1 in Zhang and Zhang (2009), under conditions (1.6) and (1.7),

$$\frac{E_n(F_1(n)) + 2E_n(F_2(n))}{s_n^2} \to 1, \qquad b(n) \asymp s_n^2,$$

and if $c_1 \in (0, 1)$, then $\lim_{n\to\infty} \frac{s_n^2}{n} = c_1 + 2c_2 > 0$.

LEMMA 3.2. *Assume that* (2.2) *holds. If* $0 < \lambda_n \le n$ *and*

$$\limsup_{n\to\infty} \frac{n - \lambda_n}{na(b(n))/b(n)} < \infty,$$

then

$$s_{\lambda_n n}^2 = \big(1 + o(1)\big)s_n^2. \tag{3.4}$$

PROOF. Set $r := \limsup_{n\to\infty} \frac{n-\lambda_n}{na(b(n))/b(n)}$. Then for any $\varepsilon > 0$, for n large enough,

$$\begin{aligned} s_{\lambda_n n}^2 \le e^{\varepsilon} &\sum_{k=1}^{\infty} \big(np_{kn} + (np_{kn})^2\big) e^{-np_{kn}} \\ &+ \sum_{k=1}^{\infty} \big(\lambda_n p_{kn} + (\lambda_n p_{kn})^2\big) e^{-np_{kn}} I_{\{np_{kn} > \varepsilon b(n)/(2ra(b(n)))\}}. \end{aligned}$$

Therefore, by (2.2), the above inequality implies that $\limsup_{n\to\infty} \frac{s_{\lambda_n n}^2}{s_n^2} \le e^{\varepsilon} \to 1$ as $\varepsilon \to 0$. On the other hand, it is clear that for any $\varepsilon > 0$, when n is large enough,

$$s_{\lambda_n n}^2 \ge \sum_{k=1}^{\infty} \big(\lambda_n p_{kn} + (\lambda_n p_{kn})^2\big) e^{-np_{kn}} \ge (1-\varepsilon)^2 s_n^2,$$

which yields that $\liminf_{n\to\infty} \frac{s_{\lambda_n n}^2}{s_n^2} \ge 1$. Thus (3.4) is valid. $\square$

LEMMA 3.3. *For any* $\varrho > 0$,

$$\lim_{n\to\infty} \frac{1}{b(n)} \sum_{k\in M_{n\varrho}} \left| E_n\big(\delta_{kj}(n)\big) - \frac{1}{j!}(np_{kn})^j e^{-np_{kn}} \right| = 0, \qquad j = 1, 2. \tag{3.5}$$

PROOF. Since $(1 - p_{kn})^{n-j} = e^{-np_{kn}}(1 + O(b(n)/a^2(b(n))))$ holds uniformly on $M_{n\varrho}$ for $j = 1, 2$, we obtain that

$$\begin{aligned} &\frac{1}{b(n)} \sum_{k\in M_{n\varrho}} \left| \frac{n!}{(n-j)!} p_{kn}^j (1 - p_{kn})^{n-j} - (np_{kn})^j e^{-np_{kn}} \right| \\ &\quad = \frac{1}{b(n)} \sum_{k\in M_{n\varrho}} (np_{kn})^j e^{-np_{kn}} \left| \frac{n!}{(n-j)!n^j} - \big(1 + O\big(b(n)/a^2\big(b(n)\big)\big)\big) \right| = o(1). \end{aligned}$$

That is, (3.5) holds. □

In order to obtain the exponential moment inequalities, we need some concepts of negative dependence; cf. Joag-Dev and Proschan (1983), Dubhashi and Ranjan (1998). Let $\eta_1, \eta_2, \dots$ be real random variables. $\eta_1, \eta_2, \dots$ are said to be negatively associated if for every two disjoint index finite sets $\Lambda_1, \Lambda_2 \subset \{1, 2, \dots\}$,

$$E\big(f(\eta_k, k \in \Lambda_1)g(\eta_k, k \in \Lambda_2)\big) \le E\big(f(\eta_k, k \in \Lambda_1)\big)E\big(g(\eta_k, k \in \Lambda_2)\big)$$

for all nonnegative functions $f : \mathbb{R}^{\Lambda_1} \to \mathbb{R}$ and $g : \mathbb{R}^{\Lambda_2} \to \mathbb{R}$ that are both nondecreasing or both nonincreasing.

LEMMA 3.4. *$\{X_k(n), k \ge 1\}$ is a sequences of negatively associated random variables, and for each $0 \le j \le n$ $\{\delta_{k0}(n) + \delta_{k1}(n) + \cdots + \delta_{kj}(n), k \ge 1\}$ is also negatively associated.*

PROOF. Let δ_k^m denote the frequency of the kth species in the mth sampling, that is,

$$\delta_k^m = I_{\{X_k(m) - X_k(m-1) = 1\}}.$$

Then $\delta_k^m, k \ge 1$ are zero-one random variables such that $\sum_{k=1}^{\infty} \delta_k^m = 1$. By Lemma 8 in Dubhashi and Ranjan (1998), $\delta_k^m, k \ge 1$, are negative associated. Since $\{\delta_k^m, k \ge 1\}, m = 1, \dots, n$, are i.i.d. under P_n, $\delta_k^m, k \ge 1, m = 1, \dots, n$, are negative associated. Noting that $X_k(n) = \sum_{m=1}^n \delta_k^m$ and

$$\delta_{k0}(n) + \delta_{k1}(n) + \cdots + \delta_{kj}(n) = \psi\big(X_k(n)\big),$$

where $\psi(x) = I_{(-\infty, j]}(x)$ is a decreasing function, we obtain that $\{X_k(n), k \ge 1\}$ and $\{\delta_{k0}(n) + \delta_{k1}(n) + \cdots + \delta_{kj}(n), k \ge 1\}$ are two sequences of negatively associated random variables. □

LEMMA 3.5. *Let M be a subset of the set $\mathbb{N}$ of positive integers. Then for any $r \in \mathbb{R}$,*

$$E_n\left(\exp\left\{r \sum_{k \in M} p_{kn}\delta_{k0}(n)\right\}\right) \le \prod_{k \in M}\left((e^{rp_{kn}} - 1)(1 - p_{kn})^n + 1\right) \tag{3.6}$$

and for any $j \ge 1$,

$$\begin{aligned} &E_n\left(\exp\left\{r \sum_{k \in M}\big(\delta_{k0}(n) + \delta_{k1}(n) + \cdots + \delta_{kj}(n)\big)\right\}\right) \\ &\quad \le \prod_{k \in M}\left((e^r - 1)\sum_{l=0}^{j} \frac{n!}{(n-l)!l!} p_{kn}^l (1 - p_{kn})^{n-l} + 1\right). \end{aligned} \tag{3.7}$$

PROOF. For any $r \in \mathbb{R}$ given, set $\psi_k(x) = e^{rp_{kn}x}$, $x \in \mathbb{R}$. Then, when $r \geq 0$, all ψ_k, $k \geq 1$ are nonnegative and increasing; when $r < 0$, all ψ_k, $k \geq 1$ are nonnegative and decreasing. Therefore, by Lemma 3.4,

$$E_n\left(\exp\left\{r \sum_{k\in M} p_{kn}\delta_{k0}(n)\right\}\right) \leq \prod_{k\in M} E_n\big(\exp\{rp_{kn}\delta_{k0}(n)\}\big)$$
$$\leq \prod_{k\in M} \big((e^{rp_{kn}} - 1)(1 - p_{kn})^n + 1\big).$$

Similarly, we can obtain (3.7). □

As applications of Lemma 3.5, we have the following exponential moment estimates. Its proof is given in Appendix B.

LEMMA 3.6. (1) *For any* $j = 0, 1, 2$ *and* $r \in \mathbb{R}$,

$$\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \times \log E_n\left(\exp\left(\frac{ra^2(b(n))}{b^2(n)} \sum_{k\in M^c_{nQ}} \sum_{l=0}^{j} \big(\delta_{kl}(n) - E_n(\delta_{kl}(n))\big)\right)\right) \leq 0. \tag{3.8}$$

(2) *If* $c_1 \in (0, 1)$, *then for any* $j = 0, 1, 2$ *and* $r \in \mathbb{R}$,

$$\limsup_{L\to\infty}\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \times \log E_n\left(\exp\left(\frac{ra(b(n))}{b(n)} \sum_{k\in M^{Lc}_n} \sum_{l=0}^{j} \big(\delta_{kl}(n) - E_n(\delta_{kl}(n))\big)\right)\right) \leq 0 \tag{3.9}$$

and

$$\limsup_{L\to\infty}\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \times \log E_n\left(\exp\left(\frac{rna(b(n))}{b(n)} \sum_{k\in M^{Lc}_n} p_{kn}\big(\delta_{k0}(n) - E_n(\delta_{k0}(n))\big)\right)\right) \leq 0. \tag{3.10}$$

3.2. *The proof of Theorem* 2.1: *Case* I. In this subsection, we use the Gärtner–Ellis theorem to show Theorem 2.1 under $c_1 \in (0, 1)$. The Laplace asymptotic integral plays a very important role.

By Lemma 3.1, if $c \in (0, 1)$, when L is large enough,

$$b^L(n) := E_n\big(F_1^L(n)\big)\big(1 - E_n\big(F_1^L(n)\big)/n\big) + 2E_n\big(F_2^L(n)\big) \asymp n \tag{3.11}$$

and

$$\lim_{L\to\infty}\limsup_{n\to\infty}\left|\frac{b^L(n)}{b(n)}-1\right|=0, \tag{3.12}$$

where $F_j^L(n)=\sum_{k\in M_n^L}\delta_{kj}(n)$, $j\ge 1$. In this subsection, We assume that L is large enough such that $b^L(n)\asymp n$ and $a(b^L(n))\asymp a(n)$.

The following Laplace asymptotic integral is a key lemma. It will be proved in Appendix B.

LEMMA 3.7. *Suppose that conditions* (1.6) *and* (1.7) *hold. If* $c_1\in(0,1)$, *then for any* $\alpha\in\mathbb{R}$,

$$\begin{aligned}&\lim_{n\to\infty}\frac{b^L(n)}{a^2(b^L(n))}\\ &\quad\times\log E_n\left(\exp\left\{\frac{\alpha a(b^L(n))}{b^L(n)}\sum_{k\in M_n^L}\big(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)\big)\right\}\right)=\frac{\alpha^2}{2}.\end{aligned} \tag{3.13}$$

PROOF OF THEOREM 2.1 UNDER $c_1\in(0,1)$. By the Gärtner–Ellis theorem [cf. Theorem 2.3.6 in Dembo and Zeitouni (1998)] and Lemma 3.7, $\{\frac{1}{a(b^L(n))}\sum_{k\in M_n^L}(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)), n\ge 1\}$ satisfies a large deviation principle with speed $\frac{a^2(b^L(n))}{b^L(n)}$ and with rate function $I(x)=\frac{x^2}{2}$. By Lemma 3.9, we only need to check

$$\begin{aligned}&\limsup_{L\to\infty}\limsup_{n\to\infty}\frac{b(n)}{a^2(b(n))}\\ &\quad\times\log P_n\left(\left|\frac{1}{a(b^L(n))}\sum_{k\in M_n^L}\big(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)\big)-\frac{n(\hat{Q}_n-Q_n)}{a(b(n))}\right|\ge\varepsilon\right)\\ &\quad=-\infty.\end{aligned} \tag{3.14}$$

It is obvious that

$$\begin{aligned}&P_n\left(\left|\frac{1}{a(b^L(n))}\sum_{k\in M_n^L}\big(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)\big)-\frac{n(\hat{Q}_n-Q_n)}{a(b(n))}\right|\ge\varepsilon\right)\\ &\quad\le P_n\left(\left|\frac{a(b^L(n))-a(b(n))}{a(b^L(n))a(b(n))}\sum_{k\in M_n^L}\big(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)\big)\right|\ge\varepsilon/2\right)\\ &\qquad+P_n\left(\frac{1}{a(b(n))}\left|\sum_{k\in M_n^{Lc}}\big(np_{kn}\delta_{k0}(n)-\delta_{k1}(n)\big)\right|\ge\varepsilon/2\right).\end{aligned} \tag{3.15}$$

From (3.12) and $\{\frac{1}{a(b^L(n))}\sum_{k\in M_n^L}(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)), n\geq 1\}$ satisfies the large deviation principle, we obtain that for any $\varepsilon>0$,

$$\begin{aligned}
&\limsup_{L\to\infty}\limsup_{n\to\infty}\frac{b^L(n)}{a^2(b^L(n))}\\
&\qquad\times\log P_n\left(\left|\frac{a(b^L(n))-a(b(n))}{a(b^L(n))a(b(n))}\sum_{k\in M_n^L}(\delta_{k1}(n)-np_{kn}\delta_{k0}(n))\right|\geq\varepsilon\right)\\
&\quad=-\infty.
\end{aligned}\tag{3.16}$$

By Lemma 3.6 and Chebyshev's inequality, we have that for any $\varepsilon>0$,

$$\begin{aligned}
&\limsup_{L\to\infty}\limsup_{n\to\infty}\frac{b(n)}{a^2(b(n))}\log P_n\left(\frac{n}{a(b(n))}\left|\sum_{k\in M_n^{Lc}}p_{kn}\big(\delta_{k0}(n)-E_n(\delta_{k0}(n))\big)\right|\geq\varepsilon\right)\\
&\quad=-\infty
\end{aligned}$$

and for $j=0,1$,

$$\begin{aligned}
&\limsup_{L\to\infty}\limsup_{n\to\infty}\frac{b(n)}{a^2(b(n))}\log P_n\left(\frac{1}{a(b(n))}\left|\sum_{k\in M_n^{Lc}}\sum_{l=0}^{j}\big(\delta_{kl}(n)-E_n(\delta_{kl}(n))\big)\right|\geq\varepsilon\right)\\
&\quad=-\infty,
\end{aligned}$$

which implies that for any $\varepsilon>0$,

$$\begin{aligned}
&\limsup_{L\to\infty}\limsup_{n\to\infty}\frac{b(n)}{a^2(b(n))}\\
&\qquad\times\log P_n\left(\frac{1}{a(b(n))}\left|\sum_{k\in M_n^{Lc}}\big(np_{kn}\delta_{k0}(n)-\delta_{k1}(n)\big)\right|\geq\varepsilon\right)\\
&\quad=-\infty.
\end{aligned}\tag{3.17}$$

Now, (3.14) follows from (3.16) and (3.17). Therefore, the conclusion of Theorem 2.1 holds under $c_1\in(0,1)$. $\square$

3.3. *The proof of Theorem* 2.1: *Case* II. In this subsection, we show Theorem 2.1 under $c_1=0$. In this case, since $\{np_{in}, i\geq 1, n\geq 1\}$ cannot be truncated as a uniformly bounded sequence, the asymptotic analysis techniques in the first case cannot be used. The proof of this case is based on the Poisson approximation [cf. Zhang and Zhang (2009)] and the truncation method.

Let first us introduce the Poissonization defined by Zhang and Zhang (2009). Define

$$\xi_n=\sum_{k=1}^{\infty}\big(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)\big)=n(\hat{Q}_n-Q_n).\tag{3.18}$$

Let N_λ be a Poisson process independent of $\{\mathbf{X}(m), m \geq 1\}$ with $E_n(N_\lambda) = \lambda$. Define the Poissonization $\zeta_{\lambda n}$ of ξ_n as follows:

$$\zeta_{\lambda n} = \sum_{k=1}^{\infty} Y_{k\lambda n} \qquad \text{where } Y_{k\lambda n} = \delta_{k1}(N_\lambda) - \lambda p_{kn}\delta_{k0}(N_\lambda). \tag{3.19}$$

Under probability P_n, $X_k(N_\lambda), k \geq 1$ are independent Poisson variables with means λp_{kn}, so that $Y_{k\lambda n}, k \geq 1$ are independent zero-mean variables with variance $\sigma^2_{k\lambda n} := \lambda p_{kn} e^{-\lambda p_{kn}} + (\lambda p_{kn})^2 e^{-\lambda p_{kn}}$. Then the Poissonization $\{\zeta_{nn}, n \geq 1\}$ satisfies the following moderate deviation principle.

LEMMA 3.8. *Let conditions* (1.6), (1.7) *and* (2.2) *hold. Then* $\{\frac{\zeta_{nn}}{a(s_n^2)}, n \geq 1\}$ *satisfies a large deviation principle with speed* $\frac{a^2(s_n^2)}{s_n^2}$ *and with rate function* $I(x) = \frac{x^2}{2}$.

PROOF. For any $\alpha \in \mathbb{R}$,

$$\begin{aligned}
&E_n\left(\exp\left\{\frac{\alpha a(s_n^2)}{s_n^2}\zeta_{nn}\right\}\right)\\
&\quad = \prod_{k=1}^{\infty} E_n\left(\exp\left\{\frac{\alpha a(s_n^2)}{s_n^2}(I_{\{X_k(N_{\lambda_n})=1\}} - np_{kn} I_{\{X_k(N_{\lambda_n})=0\}})\right\}\right)\\
&\quad = \prod_{k=1}^{\infty}\left((1 - e^{-np_{kn}} - np_{kn}e^{-np_{kn}}) + np_{kn}e^{-np_{kn}}\exp\left\{\frac{\alpha a(s_n^2)}{s_n^2}\right\}\right.\\
&\qquad\qquad \left. + e^{-np_{kn}}\exp\left\{\frac{-\alpha a(s_n^2)}{s_n^2}np_{kn}\right\}\right).
\end{aligned}$$

For any $\varepsilon \in (0, 1/2]$ such that $|\alpha|\varepsilon < 1/2$, for n large enough, we can write

$$\begin{aligned}
&1 - e^{-np_{kn}} - np_{kn}e^{-np_{kn}}\\
&\quad + np_{kn}e^{-np_{kn}}\exp\left\{\frac{\alpha a(s_n^2)}{s_n^2}\right\} + e^{-np_{kn}}\exp\left\{\frac{-\alpha a(s_n^2)}{s_n^2}np_{kn}\right\}\\
&\quad = 1 + \frac{1}{2}\left(\frac{\alpha a(s_n^2)}{s_n^2}\right)^2 (np_{kn} + (np_{kn})^2)e^{-np_{kn}} + o\left(\left(\frac{a(s_n^2)}{s_n^2}\right)^2\right)np_{kn}e^{-np_{kn}}\\
&\qquad + \left(\frac{\alpha a(s_n^2)}{s_n^2}np_{kn} - \frac{1}{2}\left(\frac{\alpha a(s_n^2)}{s_n^2}\right)^2 (np_{kn})^2\right)e^{-np_{kn}} I_{\{np_{kn} > \varepsilon s_n^2/a(s_n^2)\}}\\
&\qquad + O\left(\left(\frac{a(s_n^2)}{s_n^2}\right)^3\right)(np_{kn})^3 e^{-np_{kn}} I_{\{np_{kn} \leq \varepsilon s_n^2/a(s_n^2)\}}\\
&\qquad + e^{-np_{kn}}\left(\exp\left\{\frac{-\alpha a(s_n^2)}{s_n^2}np_{kn}\right\} - 1\right) I_{\{np_{kn} > \varepsilon s_n^2/a(s_n^2)\}}.
\end{aligned}$$

By (2.2),

$$\begin{aligned}\frac{1}{a(s_n^2)}&\sum_{k=1}^{\infty} np_{kn}e^{-np_{kn}} I_{\{np_{kn}>\varepsilon s_n^2/a(s_n^2)\}}\\&\le \frac{1}{\varepsilon s_n^2}\sum_{k=1}^{\infty}(np_{kn})^2 e^{-np_{kn}} I_{\{np_{kn}>\varepsilon s_n^2/a(s_n^2)\}} \to 0\end{aligned}$$

and

$$\begin{aligned}\frac{s_n^2}{a^2(s_n^2)}&\sum_{k=1}^{\infty} e^{-np_{kn}}\left(\exp\left\{\frac{-\alpha a(s_n^2)}{s_n^2}np_{kn}\right\}-1\right) I_{\{np_{kn}>\varepsilon s_n^2/a(s_n^2)\}}\\&\le \frac{2}{\varepsilon^2 s_n^2}\sum_{k=1}^{\infty}(np_{kn})^2 e^{-np_{kn}(1-|\alpha|a(s_n^2)/s_n^2)} I_{\{np_{kn}>\varepsilon s_n^2/a(s_n^2)\}} \to 0.\end{aligned}$$

Therefore, by $\frac{a(s_n^2)}{s_n^2}\frac{1}{s_n^2}\sum_{k=1}^{\infty}(np_{kn})^3 e^{-np_{kn}} I_{\{np_{kn}\le\varepsilon s_n^2/a(s_n^2)\}} \le \varepsilon \to 0$ as $\varepsilon \to 0$, we have that

$$\lim_{n\to\infty}\frac{s_n^2}{a^2(s_n^2)}\log E_n\left(\exp\left\{\frac{\alpha a(s_n^2)}{s_n^2}\zeta_{nn}\right\}\right)=\frac{\alpha^2}{2},$$

which implies the conclusion of the lemma by the Gärtner–Ellis theorem; cf. Theorem 2.3.6 in Dembo and Zeitouni (1998). □

By Lemmas 3.8 and A.1, we need the following exponential approximation: for any $\varepsilon > 0$,

$$\lim_{n\to\infty}\frac{b(n)}{a^2(b(n))}\log P_n\big(|\xi_n-\zeta_{nn}|>\varepsilon a\big(b(n)\big)\big)=-\infty. \tag{3.20}$$

Let us first give a maximal exponential estimate. Its proof is postponed to Appendix B.

LEMMA 3.9. *Let conditions* (1.6), (1.7) *and* (2.2) *hold, and let* $c_1 = 0$. *For any* $M \ge 1$ *fixed, set* $\lambda_n = n - Ma(b(n))\sqrt{\frac{n}{b(n)}}$, $\Delta_n = 2Ma(b(n))\sqrt{\frac{n}{b(n)}}$. *Then for any* $\varepsilon > 0$,

$$\lim_{n\to\infty}\frac{b(n)}{a^2(b(n))}\log P_n\Big(\max_{t\in[\lambda_n,\lambda_n+\Delta_n]}|\zeta_{\lambda_n n}-\zeta_{tn}|\ge\varepsilon a\big(b(n)\big)\Big)=-\infty. \tag{3.21}$$

PROOF OF THEOREM 2.1 UNDER $c_1 = 0$. By Lemmas 3.8 and A.1, we only need to prove (3.20). Set $t_n = \inf\{\lambda; N_\lambda = n\}$. Then t_n has gamma$(n, 1)$ distribution and $\xi_n - \zeta_{t_n n} = (t_n - n)\sum_{k=1}^{\infty} p_{kn}\delta_{k0}(n)$. Therefore, for any $\varepsilon > 0$ and any

$M \geq 1$,

$$
\begin{aligned}
&P_n\big(|\xi_n - \zeta_{nn}| \geq \varepsilon a\big(b(n)\big)\big) \\
&\quad \leq P_n\left(|t_n - n| \geq Ma\big(b(n)\big)\sqrt{\frac{n}{b(n)}}\right) + P_n\left(\sum_{k=1}^{\infty} p_{kn}\delta_{k0}(n) \geq \frac{\varepsilon}{2M}\sqrt{\frac{b(n)}{n}}\right) \\
&\qquad + P_n\left(\max_{t\in[n-\Delta_n/2, n+\Delta_n/2]} |\zeta_n - \zeta_{tn}| \geq \frac{\varepsilon a(b(n))}{2}\right).
\end{aligned}
$$

By Lemma 3.9,

$$
\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\max_{t\in[n-\Delta_n/2, n+\Delta_n/2]} |\zeta_n - \zeta_{tn}| \geq \frac{\varepsilon a(b(n))}{2}\right) = -\infty. \tag{3.22}
$$

By Chebyshev's inequality, it is easy to get that

$$
\lim_{M\to\infty} \limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(|t_n - n| \geq Ma\big(b(n)\big)\sqrt{\frac{n}{b(n)}}\right) = -\infty. \tag{3.23}
$$

Therefore, we only need to prove that

$$
\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\sum_{k=1}^{\infty} p_{kn}\delta_{k0}(n) \geq \frac{\varepsilon}{2M}\sqrt{\frac{b(n)}{n}}\right) = -\infty. \tag{3.24}
$$

It is sufficient that for any $r > 0$,

$$
\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log E_n\left(\exp\left(\frac{ra^2(b(n))}{b(n)}\sqrt{\frac{n}{b(n)}}\sum_{k=1}^{\infty} p_{kn}\delta_{k0}(n)\right)\right) = 0. \tag{3.25}
$$

In fact, by Lemma 3.5, we can get that for any $r > 0$,

$$
\begin{aligned}
&\log E_n\left(\exp\left(\frac{ra^2(b(n))}{b(n)}\sqrt{\frac{n}{b(n)}}\sum_{k=1}^{\infty} p_{kn}\delta_{k0}(n)\right)\right) \\
&\quad \leq \frac{2ra^2(b(n))}{b(n)}\sqrt{\frac{n}{b(n)}} \\
&\qquad \times \sum_{k=1}^{\infty}\left(p_{kn}e^{-np_{kn}} + p_{kn}\exp\left\{-n\left(1 - \frac{ra^2(b(n))}{nb(n)}\sqrt{\frac{n}{b(n)}}\right)p_{kn}\right\}\right) \\
&\quad \leq \frac{ra^2(b(n))}{b(n)}\sqrt{\frac{n}{b(n)}}\left(\frac{2s_n^2}{n} + \frac{s_{\lambda_n n}^2}{\lambda_n}\right),
\end{aligned}
$$

where $\lambda_n = n(1 - \frac{ra^2(b(n))}{nb(n)}\sqrt{\frac{n}{b(n)}})$, which implies that (3.20) holds. $\square$

3.4. *Proof of Theorem* 2.2. By the comparison method in large deviations [cf. Theorem 4.2.13 in Dembo and Zeitouni (1998)], in order to obtain Theorem 2.2, we need the following lemma.

LEMMA 3.10. *For any* $\varepsilon > 0$, *for* $j = 1, 2$,

$$\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\big(\big|F_j(n) - E_n\big(F_j(n)\big)\big| \geq \varepsilon b(n)\big) = -\infty. \tag{3.26}$$

PROOF. By (3.8), for $j = 1, 2$, for any $\varrho > 0$ and $\varepsilon > 0$,

$$\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\frac{1}{b(n)}\left|\sum_{k\in M_{n\varrho}^c} \big(\delta_{kj}(n) - E_n\big(\delta_{kj}(n)\big)\big)\right| \geq \varepsilon\right) = -\infty. \tag{3.27}$$

Therefore, by Lemma 3.3, it suffices to show that

$$\begin{aligned} &\limsup_{\varrho\to 0}\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \\ &\qquad \times \log P_n\left(\frac{1}{b(n)}\left|\sum_{k\in M_{n\varrho}} \left(\delta_{kj}(n) - \frac{1}{j!}(np_{kn})^j e^{-np_{kn}}\right)\right| \geq \varepsilon\right) \\ &\qquad = -\infty. \end{aligned} \tag{3.28}$$

Now, let us show (3.28). Using the partial inversion formula for characteristic function due to Bartlett (1938) [see also Holst (1979), Esty (1983)], for any $r \in \mathbb{R}$,

$$\begin{aligned} &E_n\left(\exp\left\{r\sum_{k=1}^{\infty}\left(\delta_{kj}(n) - \frac{1}{j!}(np_{kn})^j e^{-np_{kn}}\right)\right\}\right) \\ &\quad = \frac{n!}{2\pi n^n e^{-n}} \int_{-\pi}^{\pi} \prod_{k\in M_{n\varrho}^c} E_n\big(\exp\{iu\big(Y_k(n) - np_{kn}\big)\}\big) \\ &\qquad\qquad \times \prod_{k\in M_{n\varrho}} E_n\bigg(\exp\bigg\{iu\big(Y_k(n) - np_{kn}\big) \\ &\qquad\qquad\qquad + r\left(I_{\{Y_k(n)=j\}} - \frac{1}{j!}(np_{kn})^j e^{-np_{kn}}\right)\bigg\}\bigg)\, du, \end{aligned}$$

where $Y_k(n), k \geq 1$ are independent random variables and $Y_k(n)$ is Poisson distributed with mean np_{kn}. Let $\gamma_k(u)$ be defined as in the proof of Lemma 3.7, that is, $\gamma_k(u) = \exp\{np_{kn}(e^{iu} - 1 - iu)\}$. Set

$$\begin{aligned} &\vartheta_k(u,\alpha) \\ &\quad = \left(\exp\{iju - np_{kn}(e^{iu} - 1)\}\left(\exp\left\{\frac{\alpha a^2(b(n))}{b^2(n)}\right\} - 1\right)\frac{1}{j!}(np_{kn})^j e^{-np_{kn}} + 1\right) \\ &\qquad \times \exp\left\{-\frac{\alpha a^2(b(n))}{b^2(n)}\frac{1}{j!}(np_{kn})^j e^{-np_{kn}}\right\}. \end{aligned}$$

Then for any $\alpha \in \mathbb{R}$,

$$E_n\left(\exp\left\{\frac{\alpha a^2(b(n))}{b^2(n)} \sum_{k\in M_{n\varrho}} \left(\delta_{kj}(n) - \frac{1}{j!}(np_{kn})^j e^{-np_{kn}}\right)\right\}\right)$$

$$= \frac{n!}{2\pi n^n e^{-n}} \int_{-\pi}^{\pi} e^{n(e^{iu}-1-iu)} \prod_{k\in M_{n\varrho}} \vartheta_k(u,\alpha)\,du.$$

Set $\tau(n) = \sqrt{a(b(n)/b(n)}$. Then $\frac{\log n}{n\tau^2(n)} = \frac{b(n)\log n}{na(b(n))} \le \frac{\sqrt{b(n)}\log n}{n} \to 0$, and noting that $\sum_{k\in M_{n\varrho}} np_{kn} \le n$, $\sum_{k\in M_{n\varrho}} (np_{kn})^2 \le \varrho nb(n)/a(b(n))$, we obtain that for ϱ small enough,

$$\frac{b(n)}{a^2(b(n))} \log\left(n^{1/2} \sup_{|u|\in[\tau(n),\pi]} \left|e^{n(e^{iu}-1-iu)} \prod_{k\in M_{n\varrho}} \vartheta_k(u,\alpha)\right|\right)$$

$$\le -\frac{b(n)n\tau^2(n)}{a^2(b(n))}\left(1 + O\left(\frac{\log n}{n\tau^2(n)}\right) + O(\varrho)\right) \to -\infty.$$

Since $\sup_{u\in[-\tau(n),\tau(n)]} \sup_{k\in M_{n\varrho}} |np_{kn}(1-\cos u)| \le \varrho$, on $[-\tau(n),\tau(n)]$,

$$\left|e^{n(e^{iu}-1-iu)} \prod_{k\in M_{n\varrho}} \vartheta_k(u,\alpha)\right|$$

$$= \exp\left\{\left(O(\varrho) + O\left(\frac{\alpha^2 a^2(b(n))}{b^2(n)}\right)\right)\frac{a^2(b(n))}{b(n)}\right\} \exp\left\{-\frac{n}{2}u^2\bigl(1 + O(u)o(1)\bigr)\right\}.$$

Thus

$$\limsup_{\varrho\to 0}\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log\left|\int_{-\tau(n)}^{\tau(n)} n^{1/2} e^{n(e^{iu}-1-iu)} \prod_{k\in M_{n\varrho}} \vartheta_k(u,\alpha)\,du\right| = 0$$

and so

$$\limsup_{\varrho\to 0}\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log E_n\left(\exp\left\{r \sum_{k\in M_{n\varrho}} \left(\delta_{kj}(n) - \frac{1}{j!}(np_{kn})^j e^{-np_{kn}}\right)\right\}\right) \le 0.$$

This yields that (3.28) holds. $\square$

PROOF OF THEOREM 2.2. By Lemma 3.10, for any $\varepsilon > 0$,

$$\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\left|\frac{b(n)}{F_1(n)(1-F_1(n)/n) + 2F_2(n)} - 1\right| \ge \varepsilon\right) = -\infty.$$

Now, by

$$\left|\frac{\sqrt{b(n)}n(\hat{Q}_n - Q_n)}{a(b(n))\sqrt{F_1(n)(1-F_1(n)/n) + 2F_2(n)}} - \frac{n(\hat{Q}_n - Q_n)}{a(b(n))}\right|$$

$$= \left|\frac{n(\hat{Q}_n - Q_n)}{a(b(n))}\right|\left|\sqrt{\frac{b(n)}{F_1(n)(1-F_1(n)/n) + 2F_2(n)}} - 1\right|,$$

and the elementary inequality $|x-1| = |x^2-1|/|x+1| \leq |x^2-1|$ for all $x \geq 0$, we obtain that

$$\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\left|\frac{\sqrt{b(n)}n(\hat{Q}_n - Q_n)}{a(b(n))\sqrt{F_1(n)(1-F_1(n)/n)+2F_2(n)}} - \frac{n(\hat{Q}_n - Q_n)}{a(b(n))}\right| \geq \varepsilon\right) = -\infty. \tag{3.29}$$

Therefore, the conclusion of the theorem follows from Lemma A.1 or Theorem 4.2.13 in Dembo and Zeitouni (1998). □

APPENDIX A: SOME CONCEPTS OF LARGE DEVIATIONS

For the sake convenience, let us introduce some notions in large deviations [Dembo and Zeitouni (1998)]. Let $(\mathcal{X}, \rho)$ be a metric space. Let $(\Omega_n, \mathcal{F}_n, P_n)$, $n \geq 1$ be a sequence of probability spaces and let $\{\eta_n, n \geq 1\}$ be a sequence of measurable maps from Ω_n to $\mathcal{X}$. Let $\{\lambda_n, n \geq 1\}$ be a sequence of positive numbers tending to $+\infty$, and let $I : \mathcal{X} \to [0, +\infty]$ be inf-compact; that is, $[I \leq L]$ is compact for any $L \in \mathbb{R}$. Then $\{\eta_n, n \geq 1\}$ is said to satisfy a large deviation principle (LDP) with speed λ_n and with rate function I, if for any open measurable subset G of $\mathcal{X}$,

$$\liminf_{n\to\infty} \frac{1}{\lambda_n} \log P_n(\eta_n \in G) \geq -\inf_{x\in G} I(x) \tag{A.1}$$

and for any closed measurable subset F of $\mathcal{X}$,

$$\limsup_{n\to\infty} \frac{1}{\lambda_n} \log P_n(\eta_n \in F) \leq -\inf_{x\in F} I(x). \tag{A.2}$$

REMARK A.1. Assume that $\{\eta_n, n \geq 1\}$ satisfies $\eta_n \to \mu$ in law and a fluctuation theorem such as central limit theorem, that is, there exists a sequence $l_n \to \infty$ such that $l_n(\eta_n - \mu) \to \eta$ in law, where μ is a constant and η is a nontrivial random variable. Usually, $\{\eta_n, n \geq 1\}$ is said to satisfy a moderate deviation principle (MDP) if $\{r_n(\eta_n - \mu), n \geq 1\}$ satisfies a large deviation principle, where r_n is an intermediate scale between 1 and l_n, that is, $r_n \to \infty$ and $r_n/l_n \to 0$.

In this paper, the following exponential approximation lemma is required. It is slightly different from Theorem 4.2.16 in Dembo and Zeitouni (1998).

LEMMA A.1. *Let $\{\eta_n, n \geq 1\}$ and $\{\eta_n^L, n \geq 1\}$, $L \geq 1$ be sequences of measurable maps from Ω_n to $\mathcal{X}$. Assume that for each $L \geq 1$, $\{\eta_n^L, n \geq 1\}$ satisfies a LDP with speed λ_n^L and with rate function I. If*

$$\lim_{L\to\infty} \limsup_{n\to\infty} \left|\frac{\lambda_n^L}{\lambda_n} - 1\right| = 0 \tag{A.3}$$

and for any $\varepsilon > 0$,

$$\lim_{L\to\infty} \limsup_{n\to\infty} \frac{1}{\lambda_n} \log P_n\big(\rho\big(\eta_n, \eta_n^L\big) \geq \varepsilon\big) = -\infty, \tag{A.4}$$

the $\{\eta_n, n \geq 1\}$ satisfies a LDP with speed λ_n and with rate function I.

PROOF. Set $I(A) = \inf_{x\in A} I(x)$. For any closed subset F,

$$P(\eta_n \in F) \leq P\big(\eta_n^L \in F^\varepsilon\big) + P\big(\rho\big(\eta_n, \eta_n^L\big) \geq \varepsilon\big),$$

where $F^\varepsilon = \{y \in \mathcal{X}; \inf_{x\in F} \rho(y, x) < \varepsilon\}$. By (A.4),

$$P\big(\rho\big(\eta_n, \eta_n^L\big) \geq \varepsilon\big) \leq e^{-\lambda_n(I(F^\varepsilon)+1)}$$

for large n and L. Therefore, for large n and L

$$P\big(\eta_n^L \in F^\varepsilon\big) + P\big(\rho\big(\eta_n, \eta_n^L\big) \geq \varepsilon\big) \leq e^{-\lambda_n(I(F^\varepsilon)+o(1))} + e^{-\lambda_n(I(F^\varepsilon)+1)}$$

and so

$$\limsup_{n\to\infty} \frac{1}{\lambda_n} \log P(\eta_n \in F) \leq -I\big(F^\varepsilon\big) \to -\inf_{x\in F} I(x).$$

The argument for open sets is similar and is omitted. $\square$

APPENDIX B: PROOFS OF LEMMAS 3.6, 3.7 AND 3.9

In this Appendix, we give the proofs of several technique lemmas. The proofs of Lemmas 3.6 and 3.9 are based some exponential moment inequalities for negatively associated random variables and martingales. The refined asymptotic analysis techniques play a basic role in the proof of Lemma 3.7.

PROOF OF LEMMA 3.6. (1) By Lemma 3.5, we have that for any $r \in \mathbb{R}$, and $j = 0, 1, 2$,

$$\begin{aligned}
&\log E_n\left(\exp\left(\frac{ra^2(b(n))}{b^2(n)} \sum_{k\in M_{n\varrho}^c} \sum_{l=0}^{j}\big(\delta_{kl}(n) - E_n\big(\delta_{kl}(n)\big)\big)\right)\right) \\
&\quad\leq \sum_{k\in M_{n\varrho}^c} \left(\log\left(\left(\exp\left\{\frac{ra^2(b(n))}{b^2(n)}\right\} - 1\right) \sum_{l=0}^{j} \frac{n!}{(n-l)!l!} p_{kn}^l (1-p_{kn})^{n-l} + 1\right)\right. \\
&\qquad\qquad\left. - \frac{ra^2(b(n))}{b^2(n)} \sum_{l=0}^{j} \frac{n!}{(n-l)!l!} p_{kn}^l (1-p_{kn})^{n-l}\right) \\
&\quad\leq \frac{a(b(n))}{b(n)} \frac{r^2a^2(b(n))}{b(n)} \\
&\qquad\times \frac{1}{b(n)} \sum_{k\in M_{n\varrho}^c} \left(\frac{2}{\varrho} n p_{kn} e^{-np_{kn}} + \sum_{l=1}^{j} \frac{n!}{(n-l)!l!} p_{kn}^l e^{-(n-l)p_{kn}}\right).
\end{aligned}$$

Therefore, (3.8) holds.

(2) Similarly to the proof of (3.8), we also have that

$$\begin{aligned}
&\log E_n\left(\exp\left(\frac{ra(b(n))}{b(n)}\sum_{k\in M_n^{Lc}}\sum_{l=0}^{j}\left(\delta_{kl}(n)-E_n\left(\delta_{kl}(n)\right)\right)\right)\right)\\
&\quad\le \frac{r^2a^2(b(n))}{b(n)}\frac{1}{b(n)}\sum_{k\in M_n^{Lc}}\left(\frac{2}{L}np_{kn}e^{-np_{kn}}+\sum_{l=1}^{j}\frac{n!}{(n-l)!l!}p_{kn}^{l}e^{-(n-l)p_{kn}}\right)\\
&\quad\to 0.
\end{aligned}$$

Finally, let us prove (3.10). By Lemma 3.5, for any $r\neq 0$,

$$\begin{aligned}
&\log E_n\left(\exp\left(\frac{rna(b(n))}{b(n)}\sum_{k\in M_n^{Lc}}p_{kn}\left(\delta_{k0}(n)-E_n\left(\delta_{k0}(n)\right)\right)\right)\right)\\
&\quad\le\sum_{k\in M_n^{Lc}}\left(\log\left(\left(\exp\left\{\frac{rna(b(n))}{b(n)}p_{kn}\right\}-1\right)(1-p_{kn})^n+1\right)\right.\\
&\qquad\qquad\left.-\frac{rna(b(n))}{b(n)}p_{kn}(1-p_{kn})^n\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\log E_n\left(\exp\left(\frac{rna(b(n))}{b(n)}\sum_{k\in M_n^{Lc}}p_{kn}\left(\delta_{k0}(n)-E_n\left(\delta_{k0}(n)\right)\right)\right)\right)\\
&\quad\le 4\sum_{k\in M_n^{Lc}}\left(\frac{rna(b(n))}{b(n)}p_{kn}\right)^2e^{-np_{kn}}I_{\{|r|na(b(n))p_{kn}/b(n)\le 1\}}\\
&\qquad+12\sum_{k\in M_n^{Lc}}\exp\left\{\frac{2|r|na(b(n))}{b(n)}p_{kn}\right\}e^{-np_{kn}}I_{\{|r|na(b(n))p_{kn}/b(n)\ge 1\}}\\
&\quad\le\frac{4r^2a(b(n))}{b(n)}A_{nL}+24|r|A_n,
\end{aligned}$$

where $A_n := \frac{a(b(n))}{b(n)}\sum_{k=1}^{\infty}p_{kn}e^{-\lambda_n p_{kn}}I_{\{\lambda_n p_{kn}\ge b(n)/(|r|a(b(n)))(1-2|r|a(b(n))/b(n))\}}$, $\lambda_n = n(1-\frac{2|r|a(b(n))}{b(n)})$ and $A_{nL} := \frac{1}{b(n)}\sum_{k\in M_n^{Lc}}n^2p_{kn}^2e^{-np_{kn}}$. By the proof of Lemma 3.1, $A_n\le\frac{a(b(n))}{b(n)}\exp\{-\frac{b(n)}{|r|a(b(n))}(1-\frac{2|r|a(b(n))}{b(n)})\}\to 0$. By (3.3), $\limsup_{L\to\infty}\limsup_{n\to\infty}A_{nL}=0$. Therefore, (3.10) holds. $\square$

PROOF OF LEMMA 3.7. It is known that

$$P_n\left(X_k(n)=x_k;k=1,\dots,m\right)=P_n\left(Y_k(n)=x_k;k=1,\dots,m\,\middle|\,\sum_{k=1}^{m}Y_k(n)=n\right),$$

where $Y_k(n), k \geq 1$ are independent random variables, and $Y_k(n)$ is Poisson distributed with mean np_{kn}. Then, using the partial inversion formula for characteristic function due to Bartlett (1938) [see also Holst (1979), Esty (1983)], for any $\alpha \in \mathbb{R}$,

$$
\begin{aligned}
&E_n\biggl(\exp\biggl\{\frac{\alpha a(b^L(n))}{b^L(n)}\sum_{k\in M_n^L}\bigl(\delta_{k1}(n)-np_{kn}\delta_{k0}(n)\bigr)\biggr\}\biggr)\\
&\quad=\frac{n!}{2\pi n^n e^{-n}}\\
&\qquad\times\int_{-\pi}^{\pi}E_n\biggl(\exp\biggl\{iu\sum_{l=1}^{\infty}(Y_l(n)-np_{ln})\\
&\qquad\qquad\qquad+\frac{\alpha a(b^L(n))}{b^L(n)}\sum_{k\in M_n^L}(I_{\{Y_k(n)=1\}}-np_{kn}I_{\{Y_k(n)=0\}})\biggr\}\biggr)\,du\\
&\quad=\frac{n!}{2\pi n^n e^{-n}}\int_{-\pi}^{\pi}H_n(u,\alpha)\,du,
\end{aligned}
$$

where

$$
\begin{aligned}
H_n(u,\alpha)&=\prod_{k\in M_n^L}\bigl(\theta_k(u,\alpha)+\gamma_k(u)\bigr)\prod_{k\in M_n^{Lc}}\gamma_k(u),\\
\gamma_k(u)&:=E_n\bigl(\exp\{iu(Y_k(n)-np_{kn})\}\bigr)=\exp\{np_{kn}(e^{iu}-1-iu)\}
\end{aligned}
$$

and

$$
\begin{aligned}
\theta_k(u,\alpha)&=\gamma_k(u)+\exp\{-iunp_{kn}\}\biggl(\exp\biggl\{-\frac{\alpha na(b^L(n))}{b^L(n)}p_{kn}\biggr\}-1\biggr)\exp\{-np_{kn}\}\\
&\quad+\exp\{iu(1-np_{kn})\}\biggl(\exp\biggl\{\frac{\alpha a(b^L(n))}{b^L(n)}\biggr\}-1\biggr)np_{kn}\exp\{-np_{kn}\}.
\end{aligned}
$$

It is obvious that $H_n(-u,\alpha)=\overline{H_n(u,\alpha)}$. By Stirling's formula,

$$
\lim_{n\to\infty}\frac{n^n e^{-n}\sqrt{n}}{n!}=\frac{1}{\sqrt{2\pi}},
$$

it suffices to show that for any $\alpha\in\mathbb{R}$,

$$
\lim_{n\to\infty}\frac{b^L(n)}{a^2(b^L(n))}\log\int_{-\pi}^{\pi}n^{1/2}H_n(u,\alpha)\,du=\frac{\alpha^2}{2}. \tag{B.1}
$$

Since $np_{kn}\leq L$ uniformly in $k\in M_n^L$, we can write that for n large enough,

$$
H_n(u,\alpha)=\prod_{k\geq 1}\gamma_k(u)\prod_{k\in M_n^L}\bigl(1+\gamma_k(u)^{-1}\theta_k(u,\alpha)\bigr)=e^{n(e^{iu}-1-iu)}\prod_{k\in M_n^L}h_k(u,\alpha),
$$

where $h_k(u,\alpha) := 1 + \gamma_k(u)^{-1}\theta_k(u,\alpha)$.

Choose a positive function $\kappa(t)$ such that $\kappa(t) \to \infty$ and $a(t)\kappa(t)/t \to 0$, and define $\tau(t) = \sqrt{\frac{a(t)(\kappa(t))^{1/2}}{t}}$, $t \geq 1$ and then $\lim_{t\to\infty} \tau(t) = 0$, $\lim_{t\to\infty} \frac{\tau^2(t)t}{a(t)} = \infty$. Noting that for n large enough, $\sup_{u\in[\tau(n),\pi]}(1-\cos u) \geq \tau^2(n)/4$, we have that

$$\begin{aligned}
&\frac{b^L(n)}{a^2(b^L(n))} \log\Bigl(n^{1/2} \sup_{u\in(\tau(n),\pi]} \bigl|H_n(u,\alpha)\bigr|\Bigr) \\
&\quad \leq \frac{b^L(n)\log n}{2a^2(b^L(n))} - \frac{b^L(n)n\tau^2(n)}{4a^2(b^L(n))} + \frac{b^L(n)}{a^2(b^L(n))} \sum_{k\in M_n^L} \log \sup_{u\in(\tau(n),\pi]} \bigl|h_k(u,\alpha)\bigr| \\
&\quad = -\frac{b^L(n)n\tau^2(n)}{a^2(b^L(n))}\biggl(1 + O\biggl(\frac{\log n}{n\tau^2(n)}\biggr) + O\biggl(\frac{a(n)}{n\tau^2(n)}\biggr)\biggr) \to -\infty,
\end{aligned}$$

which implies that

$$\limsup_{n\to\infty} \frac{b^L(n)}{a^2(b^L(n))} \log\biggl|\int_{|u|\in[\tau(n),\pi]} n^{1/2} H_n(u,\alpha)\,du\biggr| = -\infty. \tag{B.2}$$

Therefore, it suffices to show that

$$\limsup_{n\to\infty} \frac{b^L(n)}{a^2(b^L(n))} \log \int_{-\tau(n)}^{\tau(n)} n^{1/2} H_n(u,\alpha)\,du = \frac{\alpha^2}{2}. \tag{B.3}$$

In order to show (B.3), let us define a transformation as follows. For $\alpha \in \mathbb{R}$ given, set $\rho(n) = \frac{\alpha a(b^L(n))}{b^L(n)} \frac{E_n(F_1^L(n))}{n}$, and define

$$\tilde{H}_n(z) = H_n\bigl(z + i\rho(n), \alpha\bigr), \qquad z \in \mathbb{C},$$

where $\mathbb{C}$ denotes the complex plane. The transformation plays an important role. Let Γ denote the closed path formed by the ordered points $-\tau(n) - i\rho(n)$, $\tau(n) - i\rho(n)$, $\tau(n)$, $-\tau(n)$, $-\tau(n) - i\rho(n)$ on the complex plane. Then by Cauchy's formula,

$$\begin{aligned}
\int_{-\tau(n)}^{\tau(n)} H_n(u,\alpha)\,du &= \int_{-\tau(n)-i\rho(n)}^{\tau(n)-i\rho(n)} \tilde{H}_n(z)\,dz \\
&= -\int_{\tau(n)}^{-\tau(n)} \tilde{H}_n(z)\,dz - \int_{\tau(n)-i\rho(n)}^{\tau(n)} \tilde{H}_n(z)\,dz \\
&\quad - \int_{-\tau(n)}^{-\tau(n)-i\rho(n)} \tilde{H}_n(z)\,dz.
\end{aligned}$$

Noting that $|\int_{\tau(n)-i\rho(n)}^{\tau(n)} \tilde{H}_n(z)\,dz| \leq |\int_0^{\rho(n)} \tilde{H}_n(\tau(n) - iu)\,du|$, by

$$\begin{aligned}
&\sup_{|u|\leq\rho(n)} \bigl|\exp\{n(e^{-u}e^{i\tau(n)} - 1 - i\tau(n) + u)\}\bigr| \\
&\quad \leq \exp\biggl\{-n\biggl(\frac{\tau^2(n)}{4}(1 - |\rho(n)|) - \rho^2(n)\biggr)\biggr\}
\end{aligned}$$

and $\sup_{|u|\le|\rho(n)|}|h_k(\tau(n)+iu,\alpha)| = 1 + O(\frac{a(n)}{n})np_{kn}e^{-np_{kn}}$, similarly to the proof of (B.2), we have that

$$\frac{b^L(n)}{a^2(b^L(n))}\log\biggl(n^{1/2}\biggl|\int_{\tau(n)-i\rho(n)}^{\tau(n)}\tilde{H}_n(z)\,dz\biggr|\biggr) \to -\infty.$$

Similarly, $\frac{b^L(n)}{a^2(b^L(n))}\log(n^{1/2}|\int_{-\tau(n)}^{-\tau(n)-i\rho(n)}\tilde{H}_n(z)\,dz|) \to -\infty$. Therefore, it suffices to prove that

$$\limsup_{n\to\infty}\frac{b^L(n)}{a^2(b^L(n))}\log\int_{-\tau(n)}^{\tau(n)}n^{1/2}\tilde{H}_n(u)\,du = \frac{\alpha^2}{2}. \tag{B.4}$$

Let $\mathfrak{Re}(z)$ and $\mathfrak{Im}(z)$ denote the real part and the imaginary part of a complex number z, respectively. Then

$$\begin{aligned}
&\mathfrak{Re}\bigl(h_k\bigl(u+i\rho(n),\alpha\bigr)\bigr)\\
&\quad = 1 + e^{np_{kn}(1-e^{-\rho(n)}\cos u)}\biggl(\cos\bigl(np_{kn}e^{-\rho(n)}\sin u\bigr)\\
&\qquad\qquad \times\biggl(\exp\biggl\{-\frac{\alpha a(b^L(n))}{b^L(n)}np_{kn}\biggr\}-1\biggr)e^{-np_{kn}}\\
&\qquad\qquad + \bigl(\cos\bigl(np_{kn}e^{-\rho(n)}\sin u\bigr)e^{-\rho(n)}\cos u\\
&\qquad\qquad\quad + \sin\bigl(np_{kn}e^{-\rho(n)}\sin u\bigr)e^{-\rho(n)}\sin u\bigr)\\
&\qquad\qquad\quad \times\biggl(\exp\biggl\{\frac{\alpha a(b^L(n))}{b^L(n)}\biggr\}-1\biggr)np_{kn}e^{-np_{kn}}\biggr)
\end{aligned}$$

and

$$\begin{aligned}
&\mathfrak{Im}\bigl(h_k\bigl(u+i\rho(n),\alpha\bigr)\bigr)\\
&\quad = e^{np_{kn}(1-e^{-\rho(n)}\cos u)}\biggl(-\sin\bigl(np_{kn}e^{-\rho(n)}\sin u\bigr)\\
&\qquad\qquad \times\biggl(\exp\biggl\{-\frac{\alpha a(b^L(n))}{b^L(n)}np_{kn}\biggr\}-1\biggr)e^{-np_{kn}}\\
&\qquad\qquad + \bigl(-\sin\bigl(np_{kn}e^{-\rho(n)}\sin u\bigr)e^{-\rho(n)}\cos u\\
&\qquad\qquad\quad + \cos\bigl(np_{kn}e^{-\rho(n)}\sin u\bigr)e^{-\rho(n)}\sin u\bigr)\\
&\qquad\qquad\quad \times\biggl(\exp\biggl\{\frac{\alpha a(b^L(n))}{b^L(n)}\biggr\}-1\biggr)np_{kn}e^{-np_{kn}}\biggr).
\end{aligned}$$

For convenience, let $O_{jn}(u)$, $j\ge 1$, denote uniformly bounded real functions such that $O_{jn}(u)=0$ for all $|u|>\tau(n)$, and $\lim_{n\to\infty}\sup_{u\in\mathbb{R}}|O_{jn}(u)|=0$. Then for n large enough, for all $u \in [-\tau(n), \tau(n)]$,

$$\begin{aligned}
&\mathfrak{Re}\big(h_k(u+i\rho(n),\alpha)\big)\\
&\quad = 1 + e^{np_{kn}(1-e^{-\rho(n)}\cos u)}\Big(\frac{1}{2}\Big(\frac{\alpha a(b^L(n))}{b^L(n)}\Big)^2\\
&\qquad\qquad \times \Big(n^2 p_{kn}^2 + \Big(1 - \frac{2E_n(F_1^L(n))}{n}\Big)np_{kn}\Big)e^{-np_{kn}}\\
&\qquad\qquad + o\Big(\frac{a(n)}{n}\Big)^2 np_{kn}e^{-np_{kn}}\\
&\qquad\qquad + u^2 O_{1n}(u) O\Big(\frac{a(n)}{n}\Big)np_{kn}e^{-np_{kn}}\Big)
\end{aligned}$$

and

$$\begin{aligned}
&\mathfrak{Im}\big(h_k(u+i\rho(n),\alpha)\big)\\
&\quad = \frac{\alpha a(b^L(n))}{b^L(n)} e^{np_{kn}(1-e^{-\rho(n)}\cos u)} unp_{kn}e^{-np_{kn}}\big(1+u^2 O_{2n}(u)\big).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\big|H_n(u+i\rho(n),\alpha)\big|\\
&\quad = e^{-n(1-e^{-\rho(n)}\cos u-\rho(n))}\exp\Big\{\frac{1}{2}\sum_{k\in M_n^L}\log\big|h_k(u+i\rho(n),\alpha)\big|^2\Big\}\\
&\quad = \exp\Big\{\frac{1}{2}\frac{\alpha^2 a^2(b^L(n))}{b^L(n)} + o\Big(\frac{a^2(n)}{n}\Big)\Big\}\exp\Big\{-\frac{n}{2}u^2\big(1+O_{4n}(u)o(1)\big)\Big\}
\end{aligned}$$

and so

$$\begin{aligned}
&\frac{b^L(n)}{a^2(b^L(n))}\log\int_{\tau(n)}^{-\tau(n)} n^{1/2}\tilde{H}_n(u)\,du\\
&\quad = \frac{\alpha^2}{2} + \frac{b^L(n)}{a^2(b^L(n))}\log\int_{-\tau(n)\sqrt{n}}^{\tau(n)\sqrt{n}}\exp\Big\{-\frac{1}{2}\frac{\alpha^2 a^2(b^L(n))}{b^L(n)}\Big\}\tilde{H}_n\big(un^{-1/2}\big)\,du.
\end{aligned}$$

Now, by

$$\begin{aligned}
&\frac{b^L(n)}{a^2(b^L(n))}\log\int_{-\tau(n)\sqrt{n}}^{\tau(n)\sqrt{n}}\exp\Big\{-\frac{1}{2}\frac{\alpha^2 a^2(b^L(n))}{b^L(n)}\Big\}\big|\tilde{H}_n\big(un^{-1/2}\big)\big|\,du\\
&\quad = o(1) + \frac{b^L(n)}{a^2(b^L(n))}\log\int_{-\tau(n)\sqrt{n}}^{\tau(n)\sqrt{n}}\exp\Big\{-\frac{1}{2}u^2\big(1+O_{4n}\big(un^{-1/2}\big)o(1)\big)\Big\}\,du\\
&\quad \to 0,
\end{aligned}$$

we obtain (B.4). The proof of Lemma 3.7 is complete. □

PROOF OF LEMMA 3.9. For any $t > \lambda_n$, we can write [cf. (A.1) in Zhang and Zhang (2009)]

$$\begin{aligned} Y_{ktn} - Y_{k\lambda_n n} \\ &= -Y_{k\lambda_n n} I_{\{X_k(N_t) > X_k(N_{\lambda_n})\}} \\ &\quad + \delta_{k0}(N_{\lambda_n})\big(\delta_{k1}(N_t) - (t-\lambda_n) p_{kn}\delta_{k0}(N_t)\big). \end{aligned} \tag{B.5}$$

Therefore, it suffices to prove that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \\ &\quad \times \log P_n\Bigg(\sup_{\lambda_n \le t \le \lambda_n + \Delta_n} \sum_{k=1}^{\infty} Y_{k\lambda_n n} I_{\{X_k(N_t) > X_k(N_{\lambda_n})\}} \ge \varepsilon a\big(b(n)\big)\Bigg) \\ &\quad = -\infty \end{aligned} \tag{B.6}$$

and

$$\begin{aligned} &\limsup_{n\to\infty} \frac{b(n)}{a^2(b(n))} \\ &\quad \times \log P_n\Bigg(\sup_{\lambda_n \le t \le \lambda_n + \Delta_n} \Bigg|\sum_{k=1}^{\infty} \delta_{k0}(N_{\lambda_n})\big(\delta_{k1}(N_t) - (t-\lambda_n) p_{kn}\delta_{k0}(N_t)\big)\Bigg| \\ &\qquad\qquad \ge \varepsilon a\big(b(n)\big)\Bigg) \\ &\quad = -\infty. \end{aligned} \tag{B.7}$$

Let us first prove (B.6). Set $T_k = \min\{t \ge 0; X_k(N_t) > X_k(N_{\lambda_n})\}$ and $Z_t^{(n)} = \sum_{T_k \le t} Y_{k\lambda_n n}$. Since $Y_{k\lambda_n n}, k \ge 1$ are independent variables with mean zero and independent of $\mathcal{G} := \sigma(\mathbf{X}(N_t) - \mathbf{X}(N_{\lambda_n}), t \ge \lambda_n)$, $\{Z_t^{(n)}, t \ge \lambda_n\}$ is a martingale, and by the maximal inequality for supermartingales, we have that for any $\varepsilon > 0$, for any $r > 0$,

$$\begin{aligned} &P_n\Big(\sup_{\lambda_n \le t \le \lambda_n + \Delta_n} \big|Z_t^{(n)}\big| \ge \varepsilon a\big(b(n)\big)\Big) \\ &\qquad \le 2e^{-r\varepsilon a(b(n))} \max\big\{E_n\big(\exp\{r Z_{\lambda_n+\Delta_n}^{(n)}\}\big), E_n\big(\exp\{-r Z_{\lambda_n+\Delta_n}^{(n)}\}\big)\big\} \end{aligned}$$

and

$$\begin{aligned} &E_n\big(\exp\{r Z_{\lambda_n+\Delta_n}^{(n)}\}\big) \\ &\quad = E_n\Bigg(E_n\Bigg(\exp\Bigg\{r\sum_{k=1}^{\infty} Y_{k\lambda_n n} I_{\{X_k(N_{\lambda_n+\Delta_n}) > X_k(N_{\lambda_n})\}}\Bigg\}\Bigg|\mathcal{G}\Bigg)\Bigg) \\ &\quad = \prod_{k=1}^{\infty}\big(\big((e^{-r\lambda_n p_{kn}} + \lambda_n p_{kn} e^r - 1 - \lambda_n p_{kn}) e^{-\lambda_n p_{kn}}\big(1 - e^{-\Delta_n p_{kn}}\big) + 1\big)\big). \end{aligned}$$

For any $\alpha \neq 0$, take $r = \frac{\alpha a(b(n))}{b(n)}$. Then for n large enough,

$$\begin{aligned}
&\exp\left\{-\frac{\alpha a(b(n))}{b(n)}\lambda_n p_{kn}\right\} + \lambda_n p_{kn}\exp\left\{\frac{\alpha a(b(n))}{b(n)}\right\} - 1 - \lambda_n p_{kn} \\
&\quad \leq \frac{3\alpha^2 a^2(b(n))}{b^2(n)}\lambda_n^2 p_{kn}^2 I_{\{|\alpha| a(b(n))\lambda_n p_{kn}/b(n)\leq 1\}} \\
&\qquad + \exp\left\{\frac{|\alpha| a(b(n))}{b(n)}\lambda_n p_{kn}\right\} I_{\{|\alpha| a(b(n))\lambda_n p_{kn}/b(n)>1\}} \\
&\qquad + \frac{3\alpha^2 a^2(b(n))}{b^2(n)}\lambda_n p_{kn}.
\end{aligned}$$

Therefore, for n large enough,

$$\begin{aligned}
&\sum_{k=1}^{\infty}\left(\exp\left\{-\frac{\alpha a(b(n))}{b(n)}\lambda_n p_{kn}\right\} + \lambda_n p_{kn}\exp\left\{\frac{\alpha a(b(n))}{b(n)}\right\} - 1 - \lambda_n p_{kn}\right) \\
&\quad \times e^{-\lambda_n p_{kn}}\left(1 - e^{-\Delta_n p_{kn}}\right) \\
&\quad\leq \frac{3\alpha^2 a^2(b(n))}{b(n)} B_{1n} + \frac{\alpha^2 a^2(b(n))}{b(n)} B_{2n} + \frac{3\alpha^2 a^2(b(n))}{b(n)} B_{3n},
\end{aligned}$$

where $B_{1n} := \frac{1}{b(n)}\sum_{k=1}^{\infty}\lambda_n^2 p_{kn}^2 I_{\{|\alpha| a(b(n))\lambda_n p_{kn}/b(n)\leq 1\}} e^{-\lambda_n p_{kn}}(1 - e^{-\Delta_n p_{kn}})$,

$$B_{2n} := \frac{1}{b(n)}\sum_{k=1}^{\infty}\sum_{k=1}^{\infty}\lambda_n^2 p_{kn}^2 \exp\left\{-\left(1 - \frac{|\alpha| a(b(n))}{b(n)}\right)\lambda_n p_{kn}\right\} I_{\{\lambda_n p_{kn} > b(n)/(|\alpha| a(b(n)))\}}$$

and $B_{3n} := \frac{1}{b(n)}\sum_{k=1}^{\infty}\lambda_n p_{kn} e^{-\lambda_n p_{kn}}(1 - e^{-\Delta_n p_{kn}})$. By (2.2), $B_{2n} \to 0$. Then, by $\frac{s_n^2}{n} \to 0$ under $c_1 = 0$, $s_n^2/b(n) \to 1$ and $s_{\lambda_n n}^2/s_n^2 \to 1$,

$$\begin{aligned}
B_{1n} &\leq \frac{4M}{|\alpha|}\sqrt{\frac{1}{\lambda_n b(n)}}\sum_{k=1}^{\infty}\lambda_n^2 p_{kn}^2 I_{\{|\alpha| a(b(n))\lambda_n p_{kn}/b(n)\leq 1\}} e^{-\lambda_n p_{kn}} \\
&\leq \frac{8M}{|\alpha|} s_{\lambda_n n}^2 \sqrt{\frac{1}{\lambda_n b(n)}} \to 0
\end{aligned}$$

and

$$\begin{aligned}
B_{3n} &\leq \frac{4M}{|\alpha|}\sqrt{\frac{1}{\lambda_n b(n)}}\sum_{k=1}^{\infty}\lambda_n p_{kn} I_{\{|\alpha| a(b(n))\lambda_n p_{kn}/b(n)\leq 1\}} e^{-\lambda_n p_{kn}} \\
&\qquad + \frac{2}{b(n)}\sum_{k=1}^{\infty}\lambda_n p_{kn} e^{-\lambda_n p_{kn}} I_{\{|\alpha| a(b(n))\lambda_n p_{kn}/b(n)>1\}} \to 0.
\end{aligned}$$

Thus

$$\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log E_n\left(\exp\left\{\frac{\alpha a(b(n))}{b(n)} Z^{(n)}_{\lambda_n+\Delta_n}\right\}\right) = 0. \tag{B.8}$$

This yields (B.6) by Chebyshev's inequality. Next, we show (B.7). Noting that

$$\begin{aligned}\sup_{\lambda_n\le t\le\lambda_n+\Delta_n} &\bigl|\delta_{k0}(N_{\lambda_n})\bigl(\delta_{k1}(N_t) - (t-\lambda_n)p_{kn}\delta_{k0}(N_t)\bigr)\bigr| \\ &\le \delta_{k0}(N_{\lambda_n})(I_{\{X_k(N_{\lambda_n+\Delta_n})>X_k(N_{\lambda_n})\}} + \Delta_n p_{kn}),\end{aligned}$$

it suffices to show that for any $\varepsilon > 0$,

$$\begin{aligned}&\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\sum_{k=1}^{\infty} \delta_{k0}(N_{\lambda_n}) I_{\{X_k(N_{\lambda_n+\Delta_n})>X_k(N_{\lambda_n})\}} > \varepsilon a\bigl(b(n)\bigr)\right) \\ &\quad = -\infty\end{aligned} \tag{B.9}$$

and

$$\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log P_n\left(\sqrt{\frac{n}{b(n)}} \sum_{k=1}^{\infty} \delta_{k0}(N_{\lambda_n}) p_{kn} > \varepsilon\right) = -\infty. \tag{B.10}$$

Since

$$\begin{aligned}&E_n\left(\exp\left\{\frac{\alpha a(b(n))}{b(n)} \sum_{k=1}^{\infty} \delta_{k0}(N_{\lambda_n}) I_{\{X_k(N_{\lambda_n+\Delta_n})>X_k(N_{\lambda_n})\}}\right\}\right) \\ &\quad = \prod_{k=1}^{\infty}\left(1 + \left(\exp\left\{\frac{\alpha a(b(n))}{b(n)}\right\} - 1\right)\bigl(1 - e^{-\Delta_n p_{kn}}\bigr)e^{-\lambda_n p_{kn}}\right),\end{aligned}$$

a similar argument to the proof of (B.8) gives

$$\lim_{n\to\infty} \frac{b(n)}{a^2(b(n))} \log E_n\left(\exp\left\{\frac{\alpha a(b(n))}{b(n)} \sum_{k=1}^{\infty} \delta_{k0}(N_{\lambda_n}) I_{\{X_k(N_{\lambda_n+\Delta_n})>X_k(N_{\lambda_n})\}}\right\}\right) = 0,$$

which implies (B.9). Similarly, we can obtain (B.10). □

Acknowledgments. The author is very grateful to the Editor, Professor T. Cai, the Associate Editor and an anonymous referee for their helpful comments and suggestions. The author is also thankful to the referee for recommending the reference Lladser, Gouet and Reeder (2011).

REFERENCES

BARTLETT, M. S. (1938). The characteristic function of a conditional statistic. *J. Lond. Math. Soc.* (2) **13** 62–67. MR1574532

CHAO, A. (1981). On estimating the probability of discovering a new species. *Ann. Statist.* **9** 1339–1342. MR0630117

DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*. Springer, New York.

DUBHASHI, D. and RANJAN, D. (1998). Balls and bins: A study in negative dependence. *Random Structures Algorithms* **13** 99–124. MR1642566

EFRON, B. and THISTED, R. (1976). Estimating the number of unseen species: How many words did Shakespeare know? *Biometrika* **63** 435–447.

ESTY, W. W. (1982). Confidence intervals for the coverage of low coverage samples. *Ann. Statist.* **10** 190–196. MR0642730

ESTY, W. W. (1983). A normal limit law for a nonparametric estimator of the coverage of a random sample. *Ann. Statist.* **11** 905–912. MR0707940

GAO, F. and ZHAO, X. (2011). Delta method in large deviations and moderate deviations for estimators. *Ann. Statist.* **39** 1211–1240. MR2816352

GOOD, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* **40** 237–264. MR0061330

GOOD, I. J. and TOULMIN, G. H. (1956). The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika* **43** 45–63. MR0077039

HOLST, L. (1979). A unified approach to limit theorems for urn models. *J. Appl. Probab.* **16** 154–162. MR0520945

JOAG-DEV, K. and PROSCHAN, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11** 286–295. MR0684886

LLADSER, M. E., GOUET, R. and REEDER, J. (2011). Extrapolation of urn models via Poissonization: Accurate measurements of the microbial unknown. *PLoS ONE* **6** e21105.

MAO, C. X. and LINDSAY, B. G. (2002). A Poisson model for the coverage problem with a genomic application. *Biometrika* **89** 669–681. MR1929171

ORLITSKY, A., SANTHANAM, N. P. and ZHANG, J. (2003). Always good Turing: Asymptotically optimal probability estimation. *Science* **302** 427–431. MR2024675

THISTED, R. and EFRON, B. (1987). Did Shakespeare write a newly-discovered poem? *Biometrika* **74** 445–455. MR0909350

ZHANG, C.-H. (2005). Estimation of sums of random variables: Examples and information bounds. *Ann. Statist.* **33** 2022–2041. MR2211078

ZHANG, C.-H. and ZHANG, Z. (2009). Asymptotic normality of a nonparametric estimator of sample coverage. *Ann. Statist.* **37** 2582–2595. MR2543704

SCHOOL OF MATHEMATICS AND STATISTICS
WUHAN UNIVERSITY
WUHAN 430072
P.R. CHINA
E-MAIL: fqgao@whu.edu.cn