

1-7-2019

Bootstrap inference for the finite population total under complex sampling designs

Zhonglei Wang
Xiamen University

Jae Kwang Kim
Iowa State University, jkim@iastate.edu

Lihua Peng
The University of Melbourne

Follow this and additional works at: https://lib.dr.iastate.edu/stat_las_pubs



Part of the [Statistical Methodology Commons](#), and the [Statistical Theory Commons](#)

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/stat_las_pubs/255. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

Bootstrap inference for the finite population total under complex sampling designs

Abstract

Bootstrap is a useful tool for making statistical inference, but it may provide erroneous results under complex survey sampling. Most studies about bootstrap-based inference are developed under simple random sampling and stratified random sampling. In this paper, we propose a unified bootstrap method applicable to some complex sampling designs, including Poisson sampling and probability-proportional-to-size sampling. Two main features of the proposed bootstrap method are that studentization is used to make inference, and the finite population is bootstrapped based on a multinomial distribution by incorporating the sampling information. We show that the proposed bootstrap method is second-order accurate using the Edgeworth expansion. Two simulation studies are conducted to compare the proposed bootstrap method with the Wald-type method, which is widely used in survey sampling. Results show that the proposed bootstrap method is better in terms of coverage rate especially when sample size is limited.

Disciplines

Statistical Methodology | Statistical Theory

Comments

This pre-print is made available through arxiv: <https://arxiv.org/abs/1901.01645>.

Bootstrap inference for the finite population total under complex sampling designs

Zhonglei Wang* Jae Kwang Kim[†] Liuhua Peng[‡]

Abstract

Bootstrap is a useful tool for making statistical inference, but it may provide erroneous results under complex survey sampling. Most studies about bootstrap-based inference are developed under simple random sampling and stratified random sampling. In this paper, we propose a unified bootstrap method applicable to some complex sampling designs, including Poisson sampling and probability-proportional-to-size sampling. Two main features of the proposed bootstrap method are that studentization is used to make inference, and the finite population is bootstrapped based on a multinomial distribution by incorporating the sampling information. We show that the proposed bootstrap method is second-order accurate using the Edgeworth expansion. Two simulation studies are conducted to compare the proposed bootstrap method with the Wald-type method, which is widely used in survey sampling. Results show that the proposed bootstrap method is better in terms of coverage rate especially when sample size is limited.

Keywords: Confidence interval, Edgeworth expansion, Multinomial distribution, Second-order accurate.

*Wang Yanan Institute for Studies in Economics and School of Economics, Xiamen University, Xiamen, Fujian 361005, P.R.C.

[†]Department of Statistics, Iowa State University, Ames, IA 50011, U.S.A.; Email: jkim@iastate.edu

[‡]School of Mathematics and Statistics, the University of Melbourne, Victoria 3010, Australia

1 Introduction

Bootstrap, first proposed by [Efron \(1979\)](#), is a simulation-based approach for accessing uncertainty of estimates and for constructing confidence intervals. Bootstrap is widely used in that it is easy to implement and is second-order accurate under mild conditions ([Hall; 1992](#), §3.3). However, classical bootstrap methods are not applicable under most sampling designs since the independent or identical distributed assumption may fail.

Under complex sampling, bootstrap methods have been proposed to handle variance estimation. In survey sampling, one of the most popular bootstrap approaches is the rescaling bootstrap method proposed by [Rao and Wu \(1988\)](#) under stratified random sampling, and they demonstrated that their bootstrap- t intervals are second-order accurate if the variance component is *known*. Such a variance, however, is seldom known in practice. [Rao et al. \(1992\)](#) generalized the rescaling bootstrap method to cover the non-smooth statistics, but they did not discuss the second-order accuracy. [Sitter \(1992a\)](#) considered a mirror-match bootstrap method for sampling designs without replacement and discussed the second-order accuracy based on the *known* population variance as [Rao and Wu \(1988\)](#). [Sitter \(1992b\)](#) extended the without-replacement bootstrap method ([Gross; 1980](#)) to complex sampling designs and compared the proposed method with the rescaling bootstrap method ([Rao and Wu; 1988](#)) and the mirror-match bootstrap method ([Sitter; 1992a](#)). [Shao and Sitter \(1996\)](#) proposed a bootstrap method for the case when survey data are subject to missingness. [Sverchkov and Pfeffermann \(2004\)](#) proposed to use a multinomial distribution to reconstruct the finite population to estimate the mean square error. [Beaumont and Patak \(2012\)](#) proposed a generalized bootstrap method for variance estimation under Poisson sampling. [Antal and Tillé \(2011\)](#) proposed one-one resampling methods to estimate the variance for some complex sampling designs. [Mashreghi et al. \(2016\)](#) gave a comprehensive overview of the bootstrap methods in survey sampling for variance estimation.

In survey sampling, the literature on bootstrap-based approaches for interval estimation is very limited. [Bickel and Freedman \(1984\)](#) first considered interval estimation under strat-

ified random sampling. [Booth et al. \(1994\)](#) generalized the method of [Bickel and Freedman \(1984\)](#) to show that the constructed confidence interval for a smooth function of the finite population mean is second-order accurate. However, all of the theoretical results, including that of [Rao and Wu \(1988\)](#) are restricted to stratified random sampling. Although [Beaumont and Patak \(2012\)](#) discussed a generalized bootstrap method for survey sampling with special attention to Poisson sampling, they did not provide rigorous results for the second-order accuracy of their methods.

In this paper, we focus on interval estimation under complex sampling. The goal of this study is to develop a unified bootstrap method to approximate the sampling distribution of the design-based estimator under some popular sampling designs, including Poisson sampling, simple random sampling (SRS) and probability-proportional-to-size (PPS) sampling. The proposed bootstrap methods apply multinomial distributions to generate the bootstrap finite populations by incorporating the sampling information, and the same sampling design is conducted to obtain a bootstrap sample from each bootstrap finite population. A similar idea has been successfully applied to SRS by [Gross \(1980\)](#) and [Chao and Lo \(1985\)](#). Our bootstrap methods differ from that proposed by [Sverchkov and Pfeffermann \(2004\)](#) in the sense that the finite population is iteratively bootstrapped, and an asymptotically pivotal statistic is used to make statistical inference for the finite population total. We also study the theoretical properties of the proposed bootstrap methods for different sampling designs using the Edgeworth expansion. We summarize our contributions in this paper below:

1. We have proposed a unified bootstrap method for interval estimation under some popular complex sampling designs, including Poisson sampling, SRS and PPS sampling. A simulation study also confirms that the proposed method works even under two-stage cluster sampling.
2. For three commonly used sampling designs, we have provided a rigorous proof for the second-order accuracy of the proposed bootstrap methods and shown that the estimation error is $o_p(n^{-1/2})$ ([DiCiccio and Romano; 1995](#)) under mild conditions. Wald-type

method is widely used in survey sampling, so the proposed bootstrap method is an important contribution since it provides more accurate inference compared with the Wald-type method under mild conditions. Besides, to our knowledge, we are the first to provide the Edgeworth expansion of a *studentized* estimator under Poisson sampling.

The remaining part of the paper is organized as follows. Sampling designs and design-based estimators under consideration are briefly reviewed in Section 2. In the following three sections, we propose bootstrap methods for Poisson sampling, SRS and PPS sampling, respectively, and theoretical properties are also investigated. Two simulation studies are conducted to compare the proposed bootstrap method with the Wald-type method in Section 6. Some concluding remarks are made in Section 7.

2 Sampling designs and estimates

In survey sampling, the finite population is often assumed to be fixed, and the randomness is due to the sampling design. Let $\mathcal{F}_N = \{y_1, \dots, y_N\}$ be the finite population of size N , and we are interested in making inference for the finite population total $Y = \sum_{i=1}^N y_i$. For simplicity, we assume that the elements in \mathcal{F}_N are scalars. To avoid unnecessary details, we also assume that the population size N is known, so it is equivalent to make statistical inference for the finite population mean $\bar{Y} = N^{-1}Y$.

We consider three commonly used sampling designs, including Poisson sampling, SRS and PPS sampling. For without-replacement sampling designs, such as Poisson sampling and SRS, I_i is the sampling indicator with $I_i = 1$ indicating that the i -th element is in the sample and 0 otherwise, and $\pi_i = E(I_i)$ is the first-order inclusion probability of the i -th element for $i = 1, \dots, N$, where the expectation is taken with respect to the sampling design. Let $\Pi_N = \{\pi_1, \dots, \pi_N\}$ be the set of first-order inclusion probabilities, and it is assumed to be known. Poisson sampling generates a sample based on N independent Bernoulli experiments, one for each element in the finite population. That is, $I_i \sim \text{Ber}(\pi_i)$ for $i = 1, \dots, N$,

where $\text{Ber}(\pi_i)$ is a Bernoulli distribution with success probability $\pi_i \in (0, 1)$, and a sample is $\{y_i : I_i = 1, i = 1, \dots, N\}$. Let $n = \sum_{i=1}^N I_i$ be a realized sample size and $n_0 = E(n) = \sum_{i=1}^N \pi_i$ be the expected sample size under Poisson sampling. For SRS, a without-replacement sample of size n is selected with equal probabilities, and we can show $\pi_i = nN^{-1}$ for $i = 1, \dots, N$ under SRS. Denote $\hat{Y}_{Poi} = \sum_{i=1}^N y_i \pi_i^{-1} I_i$ to be the Horvitz-Thompson estimator (Horvitz and Thompson; 1952) of Y under Poisson sampling, and the corresponding one is $\hat{Y}_{SRS} = \sum_{i=1}^N y_i \pi_i^{-1} I_i = Nn^{-1} \sum_{i=1}^N I_i y_i$ under SRS. The sample size n is random under Poisson sampling, but it is fixed under SRS. Without loss of generality, assume that the first n elements are sampled under Poisson sampling or SRS, and the design-unbiased variance estimators are $\hat{V}_{Poi} = \sum_{i=1}^n y_i^2 (1 - \pi_i) \pi_i^{-2}$ and $\hat{V}_{SRS} = N(N - n)n^{-1} s_{SRS}^2$, respectively, where $s_{SRS}^2 = n^{-1} \sum_{i=1}^n (y_i - \bar{y})^2$ is the sample variance of $\{y_1, \dots, y_n\}$, and $\bar{y} = n^{-1} \sum_{i=1}^n y_i$.

PPS sampling generates a sample of size n by independently and identically selecting an element from \mathcal{F}_N n times with selection probabilities $\{p_i : i = 1, \dots, N\}$, where $p_i \in (0, 1)$ is the *known* selection probability of y_i for $i = 1, \dots, N$ and $\sum_{i=1}^N p_i = 1$. Replicates may occur in the sample under PPS sampling, and the population total Y is estimated by the Hansen–Hurwitz estimator (Hansen and Hurwitz; 1943), which is denoted as $\hat{Y}_{PPS} = n^{-1} \sum_{i=1}^n Z_i$, where $Z_i = p_{a,i}^{-1} y_{a,i}$, $p_{a,i} = p_k$ and $y_{a,i} = y_k$ if $a_i = k$, and a_i is the index of the selected element for the i -th draw. A design-unbiased variance estimator is $\hat{V}_{PPS} = n^{-2} \sum_{i=1}^n (Z_i - \hat{Y}_{PPS})^2$.

Throughout the paper, assume that the (expected) sample size is less than the population size. Since we study a sequence of finite populations and inclusion probabilities in the following three sections, assume that y_i and π_i are indexed by N implicitly, and samples are generated independently for different finite populations. We use the notation “ $a_n \asymp b_n$ ” to indicate that a_n and b_n have the same asymptotic order. That is, $a_n \asymp b_n$ is equivalent to $a_n = O(b_n)$ and $b_n = O(a_n)$.

3 Bootstrap method for Poisson sampling

We propose the following bootstrap method to approximate the sampling distribution of

$$T_{Poi} = \hat{V}_{Poi}^{-1/2}(\hat{Y}_{Poi} - Y) \text{ under Poisson sampling.}$$

Step 1. Based on the sample $\{y_1, \dots, y_n\}$, generate (N_1^*, \dots, N_n^*) from a multinomial distribution $MN(N; \rho)$ with N trials and a probability vector ρ , where $\rho = (\rho_1, \dots, \rho_n)$ and

$$\rho_i = \frac{\pi_i^{-1}}{\sum_{j=1}^n \pi_j^{-1}}$$

for $i = 1, \dots, n$. Denote $\mathcal{F}_N^* = \{y_1^*, \dots, y_N^*\}$ and $\Pi_N^* = \{\pi_1^*, \dots, \pi_N^*\}$, and they consist of N_i^* copies of y_i and π_i , respectively. Let the bootstrap finite population total be $Y^* = \sum_{i=1}^N y_i^* = \sum_{i=1}^n N_i^* y_i$.

Step 2. For $i = 1, \dots, n$, generate m_i^* independently from a binomial distribution $\text{Bin}(N_i^*, \pi_i)$ with N_i^* trials and a success probability π_i . The bootstrap sample consists of m_i^* replicates of y_i under Poisson sampling. Denote $\hat{Y}_{Poi}^* = \sum_{i=1}^n m_i^* y_i \pi_i^{-1}$ and $T_{Poi}^* = (\hat{V}_{Poi}^*)^{-1/2}(\hat{Y}_{Poi}^* - Y^*)$, where $\hat{V}_{Poi}^* = \sum_{i=1}^n m_i^* y_i^2 (1 - \pi_i) \pi_i^{-2}$ is the bootstrap variance estimator.

Step 3. Repeat the two steps above independently M times.

Step 1 corresponds to generating a bootstrap finite population \mathcal{F}_N^* and bootstrap first-order inclusion probabilities Π_N^* by incorporating the sampling information. Based on \mathcal{F}_N^* and Π_N^* , Step 2 is used to generate a bootstrap sample, from which a bootstrap replicate of T_{Poi} is obtained. Instead of sampling from the bootstrap finite population \mathcal{F}_N^* directly, Step 2 provides a more efficient way to generate a sample using N_1^*, \dots, N_n^* under Poisson sampling. In Step 2, we center T_{Poi}^* by the bootstrap population total Y^* not by \hat{Y}_{Poi} . The reason is that the finite population should be fixed, and the randomness is due to Poisson sampling. Thus, the statistic should be centered using the corresponding population total Y^* . If we center T_{Poi}^* by \hat{Y}_{Poi} , it causes additional variability due to generating different bootstrap

finite populations. The same argument applies for the other two sampling designs. We use the empirical distribution of T_{Poi}^* to approximate that of T_{Poi} and make inference for Y .

Before discussing the theoretical properties of the proposed bootstrap method, we introduce some mild conditions on \mathcal{F}_N and Π_N .

(C1) There exist constants $\alpha \in (2^{-1}, 1]$ and $0 < C_1 \leq C_2$ such that $n_0 \asymp N^\alpha$, and π_i satisfies

$$C_1 \leq n_0^{-1} N \pi_i \leq C_2$$

for $i = 1, \dots, N$.

(C2) The sequence of finite populations and first-order inclusion probabilities satisfy

$$\lim_{N \rightarrow \infty} (n_0 N^{-2} V_{Poi}) = \sigma_1^2,$$

where $V_{Poi} = \sum_{i=1}^N y_i^2 (1 - \pi_i) \pi_i^{-1}$, and σ_1^2 is a positive constant.

(C3) The following condition holds for finite populations, that is,

$$\lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N y_i^8 = C_3,$$

where C_3 is a positive constant.

(C4) Denote $X_i = V_{Poi}^{-1/2} y_i \pi_i^{-1} (I_i - \pi_i)$ for $i = 1, \dots, N$, and let $m = \lfloor n_0^{-1/2} N / (\log n_0) \rfloor$ be the integer part of $n_0^{-1/2} N / (\log n_0)$. Then, there exist constants $t_0 > 0$ and $a > 2$ such that, for any subset $\{X_{\ell_1}, \dots, X_{\ell_m}\}$ of $\{X_1, \dots, X_N\}$,

$$\left| \prod_{i=1}^m E\{\exp(\iota t X_{\ell_i}) \mid \mathcal{F}_N\} \right| = O(m^{-a})$$

uniformly in $|t| > t_0 > 0$, where ι is the imaginary unit.

We briefly comment on these conditions. Condition (C1) is commonly used in survey

sampling (Fuller; 2009); the first part of (C1) is a mild restriction on the expected sample size, and the second part regulates the first-order inclusion probabilities. Condition (C2) rules out the degenerate case of the Horvitz–Thompson estimator under Poisson sampling. The moment condition in (C3) guarantees the convergence of the variance estimators and other quantities, and it is also required for SRS and PPS sampling that we will discuss in the following two sections. To illustrate the existence of \mathcal{F}_N and Π_N satisfying (C1) and (C2) simultaneously, consider $\pi_i = n_0 N^{-1}$, so (C1) holds, where $n_0 = \lfloor N^{2/3} \rfloor$ and $C_1 = C_2 = 1$, for example. Then, we have $\lim_{N \rightarrow \infty} (n_0 N^{-2} V_{Poi}) = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N y_i^2 (1 - n_0 N^{-1})$. If $N^{-1} \sum_{i=1}^N y_i^2$ converges as $N \rightarrow \infty$, then (C2) holds. Condition (C4) is a counterpart of non-lattice assumption and is useful in deriving Edgeworth expansions. Specifically, for any subset $\{X_{\ell_1}, \dots, X_{\ell_m}\}$ of $\{X_1, \dots, X_N\}$, condition (C4) ensures that $\{X_{\ell_i}\}_{i=1}^m$ have subsequences of length $O(\log m)$ with different spans; see Feller (2008, §16.6) for more discussion on a similar assumption.

Denote $(\mathcal{F}_N, \mathcal{B}_N, P_{N, Poi})$ to be a probability space, where \mathcal{B}_N and $P_{N, Poi}(\cdot)$ are the σ -algebra and the probability measure on \mathcal{F}_N associated with Poisson sampling, respectively. That is, $\mathcal{F}_N = \times_{i=1}^N \Omega_i$, $\mathcal{B}_N = \otimes_{i=1}^N \mathcal{A}_i$ and $P_{N, Poi}(A_1 \times A_2 \times \dots \times A_N) = \prod_{i=1}^N \mu_i(A_i)$, where $\Omega_i = \{0, 1\}$, \mathcal{A}_i is the power set of Ω_i , and $\mu_i(\{1\}) = 1 - \mu_i(\{0\}) = \pi_i$ for $i = 1, \dots, N$. Let $\mathcal{F} = \times_{N=1}^{\infty} \mathcal{F}_N$ be the product space and $\mathcal{B} = \otimes_{N=1}^{\infty} \mathcal{B}_N$ be the product- σ -algebra; see Klenke (2014, §14.1) for details about the notations. By Corollary 14.33 of Klenke (2014), there exists a uniquely determined probability measure \mathbb{P}_{Poi} on $(\mathcal{F}, \mathcal{B})$ such that $\mathbb{P}_{Poi}(F_1 \times F_2 \times \dots \times F_n \times \times_{N=n+1}^{\infty} \mathcal{F}_N) = \prod_{i=1}^n P_{i, Poi}$, where $F_i \in \mathcal{F}_i$ for $i = 1, \dots, n$ and $n \in \mathbb{N}$.

Lemma 3.1. *Suppose that (C1)–(C3) hold. Then,*

$$n_0 N^{-2} (\hat{V}_{Poi} - V_{Poi}) \rightarrow 0 \tag{1}$$

as $N \rightarrow \infty$ almost surely (\mathbb{P}_{Poi}).

Let $\mu_{Poi}^{(3)} = \sum_{i=1}^N y_i^3 (1 - \pi_i) \{(1 - \pi_i)^2 \pi_i^{-2} - 1\}$ and $\hat{\mu}_{Poi}^{(3)} = \sum_{i=1}^n y_i^3 (1 - \pi_i) \pi_i^{-1} \{(1 - \pi_i)^2 \pi_i^{-2} - 1\}$.

Then,

$$n_0^2 N^{-3} \mu_{Poi}^{(3)} = O(1) \quad \text{and} \quad \frac{n_0^2}{N^3} (\hat{\mu}_{Poi}^{(3)} - \mu_{Poi}^{(3)}) = O_p(n_0^{-1/2}). \quad (2)$$

In addition, denote $\tau_{Poi}^{(3)} = \sum_{i=1}^N y_i^3 (1 - \pi_i)^2 \pi_i^{-2}$ and $\hat{\tau}_{Poi}^{(3)} = \sum_{i=1}^n y_i^3 (1 - \pi_i)^2 \pi_i^{-3}$. Then,

$$\frac{n_0^2}{N^3} \tau_{Poi}^{(3)} = O(1) \quad \text{and} \quad \frac{n_0^2}{N^3} (\hat{\tau}_{Poi}^{(3)} - \tau_{Poi}^{(3)}) = O_p(n_0^{-1/2}). \quad (3)$$

Lemma 3.1 shows some basic properties of the finite population quantities and their design-based estimators. Specifically, the ratio $\hat{V}_{Poi}^{-1} V_{Poi} \rightarrow 1$ almost surely (\mathbb{P}_{Poi}) by (C2). Under Poisson sampling, $\mu_{Poi}^{(3)}$ is the third central moment of \hat{Y}_{Poi} , and $\tau_{Poi}^{(3)}$ is a quantity involved in the Edgeworth expansion of the distribution of T_{Poi} .

Theorem 3.1. *Assume that conditions (C1)–(C4) hold. Let $\hat{F}_{Poi}(z) = \mathbb{P}_{Poi}(T_{Poi} \leq z)$ be the cumulative distribution function of T_{Poi} under Poisson sampling. Then,*

$$\frac{\hat{\mu}_{Poi}^{(3)}}{\hat{V}_{Poi}^{3/2}} = O_p(n_0^{-1/2}) \quad \text{and} \quad \frac{\hat{\tau}_{N,Poi}^{(3)}}{\hat{V}_{N,Poi}^{3/2}} = O_p(n_0^{-1/2}). \quad (4)$$

Furthermore,

$$\hat{F}_{Poi}(z) = \Phi(z) + \left\{ \frac{\hat{\mu}_{N,Poi}^{(3)}}{6\hat{V}_{N,Poi}^{3/2}} (1 - z^2) + \frac{\hat{\tau}_{N,Poi}^{(3)}}{2\hat{V}_{N,Poi}^{3/2}} z^2 \right\} \phi(z) + o_p(n_0^{-1/2}) \quad (5)$$

uniformly in $z \in \mathbb{R}$, where $\Phi(z)$ is the cumulative distribution function of the standard normal distribution with the probability density function $\phi(z)$.

we make brief comments on the $o_p(\cdot)$ notation in (5) of Theorem 3.1. The probability $\hat{F}_{Poi}(z)$ on the left side of (5) is not random. However, we use estimators in the Edgeworth expansion to make it easier to compare (5) with the result in the following theorem, so instead of $o(\cdot)$, it is reasonable to use $o_p(\cdot)$ on the right side of (5). Similar argument can be made for Edgeworth expansions under the other two sampling designs.

In order to establish the Edgeworth expansion for the conditional distribution of T_{Poi}^* , we need the following assumption, which is similar to condition (C4) but with m replaced by n_0 . We isolate (C4) and (C5) since (C5) is not needed for Theorem 3.1.

(C5) There exist constants $t_0 > 0$ and $a > 2$ such that, for any subset $\{X_{\ell_1}, \dots, X_{\ell_{n_0}}\}$ of $\{X_1, \dots, X_N\}$ with cardinality n_0 ,

$$\left| \prod_{i=1}^{n_0} E\{\exp(itX_{\ell_i}) \mid \mathcal{F}_N\} \right| = O(n_0^{-a})$$

uniformly in $|t| > t_0 > 0$.

The next theorem presents the Edgeworth expansion for the distribution of T_{Poi}^* based on the proposed bootstrap method.

Theorem 3.2. *Suppose that conditions (C1)–(C5) hold. Let $\hat{F}_{Poi}^*(z)$ be the cumulative distribution function of T_{Poi}^* conditional on the bootstrap finite population \mathcal{F}_{Poi}^* . Then,*

$$\hat{F}_{Poi}^*(z) = \Phi(z) + \left\{ \frac{\hat{\mu}_{N,Poi}^{(3)}}{6\hat{V}_{N,Poi}^{3/2}}(1 - z^2) + \frac{\hat{\tau}_{N,Poi}^{(3)}}{2\hat{V}_{N,Poi}^{3/2}}z^2 \right\} \phi(z) + o_p(n_0^{-1/2}) \quad (6)$$

uniformly in $z \in \mathbb{R}$.

By comparing (5) in Theorem 3.1 with (6) in Theorem 3.2, we show that the proposed bootstrap method is second-order accurate, but the Wald-type method, which is based on the asymptotic normality of T_{Poi} , is not if $\mu_{Poi}^{(3)}$ and $\tau_{Poi}^{(3)}$ are nonzero by noting the fact that $\hat{\mu}_{Poi}^{(3)}$ and $\hat{\tau}_{Poi}^{(3)}$ are design-unbiased estimators of $\mu_{Poi}^{(3)}$ and $\tau_{Poi}^{(3)}$, respectively. Typically, the cumulative distribution function $\hat{F}_{Poi}^*(z)$ is hard to study analytically, so we use an empirical distribution to approximate it.

Now, consider establishing confidence intervals for the population total Y . An approximate two-sided confidence interval at significance level α based on the Wald-type method

can be constructed as

$$\left(\hat{Y}_{Poi} - q_{1-\alpha/2} \hat{V}_{Poi}^{1/2}, \hat{Y}_{Poi} - q_{\alpha/2} \hat{V}_{Poi}^{1/2} \right), \quad (7)$$

where $q_{\alpha/2}$ and $q_{1-\alpha/2}$ are the $(\alpha/2)$ and $(1 - \alpha/2)$ quantiles of the standard normal distribution, respectively. According to Theorem 3.1, though the upper and lower confidence limits of (7) have error rates of order $O_p(n_0^{-1/2})$, this two-sided confidence interval has error rate of order $o_p(n^{-1/2})$ since $\hat{\mu}_{N,Poi}^{(3)}/(6\hat{V}_{N,Poi}^{3/2})(1 - z^2) + \hat{\tau}_{N,Poi}^{(3)}(2\hat{V}_{N,Poi}^{3/2})z^2$ is an even function of z , and the $n_0^{-1/2}$ order term in the Edgeworth expansion of T_{Poi} cancel in the error rate. However, the $n_0^{-1/2}$ order term leads to an error rates of order $O_p(n_0^{-1/2})$ for one-sided confidence intervals based on the normal approximation.

The confidence interval of Y based on the proposed bootstrap methods is

$$\left(\hat{Y}_{Poi} - q_{1-\alpha/2}^* \hat{V}_{Poi}^{1/2}, \hat{Y}_{Poi} - q_{\alpha/2}^* \hat{V}_{Poi}^{1/2} \right), \quad (8)$$

where $q_{\alpha/2}^*$ and $q_{1-\alpha/2}^*$ are the $(\alpha/2)$ and $(1 - \alpha/2)$ quantiles of $\hat{F}_{Poi}^*(z)$. By Theorem 3.2, the coverage error of (8) is of order $o_p(n_0^{-1/2})$. Moreover, the upper and lower limits of (8) have error rates $o_p(n_0^{-1/2})$, which outperforms the confidence interval (7) based on Wald-type method. In addition, the one-sided confidence interval by the proposed bootstrap method is more accurate than the one-sided confidence interval obtained by the Wald-type method. Furthermore, as discussed in Section 3.6 of Hall (1992), an asymmetric equal-tailed confidence interval may convey important information. The same arguments can be used for the other two sampling designs.

4 Bootstrap method for SRS

We propose the following procedure to make statistical inference for $T_{SRS} = \hat{V}_{SRS}^{-1/2}(\hat{Y}_{SRS} - Y)$ under SRS.

Step 1. Generate (N_1^*, \dots, N_n^*) from $MN(N; \rho)$, where $\rho_i = n^{-1}$ for $i = 1, \dots, n$. Then, \mathcal{F}_N^* contains N_i^* copies of y_i for $i = 1, \dots, n$, and the bootstrap finite population total is $Y^* = \sum_{i=1}^n N_i^* y_i$.

Step 2. Generate a bootstrap sample of size n , denoted as $\{y_1^*, \dots, y_n^*\}$, from \mathcal{F}_N^* using SRS. Then, we can obtain $T_{SRS}^* = (\hat{V}_{SRS}^*)^{-1/2}(\hat{Y}_{SRS}^* - Y^*)$, where $\hat{Y}_{SRS}^* = Nn^{-1} \sum_{i=1}^n y_i^*$, $\hat{V}_{SRS}^* = N(N-n)n^{-1} s_{SRS}^{*2}$, $s_{SRS}^{*2} = n^{-1} \sum_{i=1}^n (y_i^* - \bar{y}^*)^2$, and $\bar{y}^* = n^{-1} \sum_{i=1}^n y_i^*$.

Step 3. Repeat the two steps above independently M times.

The three steps for SRS are similar to those under Poisson sampling, but we do not need Π_N^* since $\pi_i^* = nN^{-1}$ for $i = 1, \dots, N$. Different from that under Poisson sampling, the bootstrap sample is generated directly from \mathcal{F}_N^* . One commonly used algorithm to generate a sample of size n under SRS is to select elements sequentially from the finite population without replacement. If $n = o(N)$, the computational complexity of selecting each element is $O(N)$. Besides the above bootstrap procedure, we propose the following one. It can be shown that these two procedures are equivalent under SRS, but the computational complexity of the latter is $O(n)$ for selecting each element.

Step 1'. The same as Step 1 above.

Step 2'. Initialize $N_i^{*(0)} = N_i^*$ and $m_i^* = 0$ for $i = 1, \dots, n$.

Step 3'. Generate a bootstrap sample of size n from \mathcal{F}_N^* under SRS.

Step 3.1'. Initialize $k = 1$.

Step 3.2'. Select an index, say $l^{(k)}$, from $\{1, \dots, n\}$ with selection probability $p_i^{(k)} = N_i^{*(k-1)} / \sum_{j=1}^n N_j^{*(k-1)}$ for $i = 1, \dots, n$.

Step 3.3'. Update $m_i^* = m_i^* + 1$ if $i = l^{(k)}$. Set $N_i^{*(k)} = N_i^{*(k-1)}$ if $i \in \{1, \dots, n\} \setminus \{l^{(k)}\}$, and $N_i^{*(k)} = N_i^{*(k-1)} - 1$ if $i = l^{(k)}$, where $A \setminus B = \{x \in A : x \notin B\}$ for two sets A and B .

Step 3.4'. Set $k = k + 1$, and go back to Step 3.2' until $k > n$.

Step 3.5'. Obtain $T_{SRS}^* = (\hat{V}_{SRS}^*)^{-1/2}(\hat{Y}_{SRS}^* - Y^*)$, where $\hat{Y}_{SRS}^* = Nn^{-1} \sum_{i=1}^n m_i^* y_i$, $\hat{V}_{SRS}^* = N(N-n)n^{-1} s_{SRS}^{*2}$, $s_{SRS}^{*2} = n^{-1} \sum_{i=1}^n m_i^* (y_i - \bar{y}^*)^2$, and $\bar{y}^* = n^{-1} \sum_{i=1}^n m_i^* y_i$.

Step 4'. Repeat the above three steps independently M times.

We list some necessary conditions for studying the theoretical properties of the proposed bootstrap method under SRS.

(C6) There exist $\beta \in (2^{-1}, 1]$ and $\kappa \in (0, 1)$ such that $n \asymp N^\beta$ and $nN^{-1} \leq 1 - \kappa$ as $N \rightarrow \infty$.

(C7) The finite population satisfies

$$\lim_{N \rightarrow \infty} \sigma_{SRS}^2 = \sigma_2^2,$$

where $\sigma_{SRS}^2 = N^{-1} \sum_{i=1}^N (y_i - N^{-1}Y)^2$, and σ_2^2 is a positive constant.

(C8) The distribution $G_{N,SRS}$ converges weakly to a strongly non-lattice distribution G_{SRS} , where $G_{N,SRS}$ assigns probability $1/N$ to y_1, \dots, y_N .

Condition (C6) is a counterpart of (C1), and it is used to rule out the trivial case when the sample size equals to that of the finite population. Condition (C7) regulates the variance of \mathcal{F}_N with respect to the distribution $G_{N,SRS}$, and it concentrates our discussion on the non-degenerate case under SRS. The non-latticed assumption in (C8) is used to make the discussion easier, and a distribution $G(x)$ is strongly non-latticed if $|\int \exp(tx) dG(x)| \neq 1$ for all $t \neq 0$; see [Babu and Singh \(1984\)](#) for details.

We can use a similar argument made in Section 3 to show that there exists a probability measure \mathbb{P}_{SRS} on the product space $\mathcal{F} = \times_{N=1}^{\infty} \mathcal{F}_N$ equipped with the product σ -algebra \mathcal{B} .

Lemma 4.1. *Suppose that (C3), (C6) and (C7) hold. Then,*

$$\mu_{SRS}^{(3)} = O(1), \tag{9}$$

where $\mu_{SRS}^{(3)} = N^{-1} \sum_{i=1}^N (y_i - N^{-1}Y)^3$ is the third central moment of \mathcal{F}_N with respect to the distribution $G_{N,SRS}$. Besides,

$$s_{SRS}^2 - \sigma_{SRS}^2 \rightarrow 0 \quad (10)$$

as $N \rightarrow \infty$ almost surely (\mathbb{P}_{SRS}). In addition,

$$\hat{\mu}_{SRS}^{(3)} - \mu_{SRS}^{(3)} = o_p(1), \quad (11)$$

where $\hat{\mu}_{SRS}^{(3)} = n^{-1} \sum_{i=1}^n y_i^3 + 2\bar{y}_n^3 - 3\bar{y}_n n^{-1} \sum_{i=1}^n y_i^2$, and $\bar{y}_n = n^{-1} \sum_{i=1}^n y_i$ is the sample mean.

Lemma 4.1 is the counterpart of Lemma 3.1 under SRS, and it shows the convergence properties of the sample variance and third central moment under mild conditions. We do not use scaling factors in (9)–(11) since the quantities involved are with respect to the distribution $G_{N,SRS}$.

Theorem 4.1. *Suppose that (C3) and (C6)–(C8) hold. Then,*

$$\hat{F}_{SRS}(z) = \Phi(z) + \frac{(1 - n/N)^{1/2} \hat{\mu}_{SRS}^{(3)}}{6n^{1/2} s_{SRS}^3} \left\{ 3z^2 - \frac{1 - 2n/N}{1 - n/N} (z^2 - 1) \right\} \phi(z) + o_p(n^{-1/2}) \quad (12)$$

uniformly in $z \in \mathbb{R}$, where $\hat{F}_{SRS}(z) = \mathbb{P}_{SRS}(T_{SRS} \leq z)$ is the cumulative distribution function of T_{SRS} .

Theorem 4.1 shows the Edgeworth expansion for the distribution of T_{SRS} , and this result is obtained by one result in Section 2 of Babu and Singh (1985). Instead of using $\mu_{SRS}^{(3)}$ and σ_{SRS} as done by Babu and Singh (1985), we use their estimators in (12) based on Lemma 4.1.

Theorem 4.2. *Suppose that (C3) and (C6)–(C8) hold. Then, we have*

$$\hat{F}_{SRS}^*(z) = \Phi(z) + \frac{(1 - n/N)^{1/2} \hat{\mu}_{SRS}^{(3)}}{6n^{1/2} s_{SRS}^3} \left\{ 3z^2 - \frac{1 - 2n/N}{1 - n/N} (z^2 - 1) \right\} \phi(z) + o_p(n^{-1/2}) \quad (13)$$

uniformly in $z \in \mathbb{R}$, where $\hat{F}_{SRS}^*(z)$ is the cumulative distribution function of T_{SRS}^* conditional on the bootstrap finite population \mathcal{F}_N^* .

Theorem 4.2 shows the Edgeworth expansion for the distribution of T_{SRS}^* obtained by the proposed bootstrap method. By comparing (12) in Theorem 4.1 with (13) in Theorem 4.2, we have shown the second-order accuracy of the proposed bootstrap method.

5 Bootstrap method for PPS sampling

We consider PPS sampling in this section and propose the following bootstrap method to approximate the sampling distribution of $T_{PPS} = \hat{V}_{PPS}^{-1/2}(\hat{Y}_{PPS} - Y)$.

Step 1. Obtain $(N_{a,1}^*, \dots, N_{a,n}^*)$ from a multinomial distribution $MN(N; \rho)$, where $\rho = (\rho_1, \dots, \rho_n)$ and $\rho_i = p_{a,i}^{-1}(\sum_{j=1}^n p_{a,j}^{-1})^{-1}$ for $i = 1, \dots, n$. Then, $\mathcal{F}_N^* = \{y_1^*, \dots, y_N^*\}$ consists of $N_{a,i}^*$ copies of $y_{a,i}$, and the bootstrap finite population total is $Y^* = \sum_{i=1}^N y_i^* = \sum_{i=1}^n N_{a,i}^* y_{a,i}$. The bootstrap selection probabilities are $\{(C_N^*)^{-1} p_1^*, \dots, (C_N^*)^{-1} p_N^*\}$, where $C_N^* = \sum_{i=1}^N p_i^* = \sum_{i=1}^n N_{a,i}^* p_{a,i}$, and $\{p_1^*, \dots, p_N^*\}$ consists of $N_{a,i}^*$ copies of $p_{a,i}$ for $i = 1, \dots, n$.

Step 2. Based on \mathcal{F}_N^* , generate a sample of size n by independently and identically selecting an element from \mathcal{F}_N^* n times with selection probabilities $\{(C_N^*)^{-1} p_i^* : i = 1, \dots, N\}$. Then, we have $T_{PPS}^* = (\hat{V}_{PPS}^*)^{-1/2}(\hat{Y}_{PPS}^* - Y^*)$, where $\hat{Y}_{PPS}^* = n^{-1} \sum_{i=1}^n C_N^* (p_{b,i}^*)^{-1} y_{b,i}^*$, $y_{b,i}^* = y_k^*$ and $p_{b,i}^* = p_k^*$ if the index of the i -th draw is k , and $\hat{V}_{PPS}^* = n^{-2} \sum_{i=1}^n \{C_N^* (p_{b,i}^*)^{-1} y_{b,i}^* - \hat{Y}_{PPS}^*\}^2$ is the counterpart of \hat{V}_{PPS} based on the bootstrap sample.

Step 3. Repeat the two steps above independently M times.

To implement the proposed bootstrap method for PPS sampling, the bootstrap selection probability should be standardized before drawing a sample. Similarly to the previous two sections, we use the empirical distribution of T_{PPS}^* to make statistical inference for T_{PPS} .

The computational complexity of selecting an element in Step 2 is $O(N)$. An equivalent way of carrying out the proposed bootstrap method under PPS sampling is described below, and its computational complexity is $O(n)$ for selecting an element.

Step 1'. The same as Step 1 above.

Step 2'. Obtain an independent and identical sample of size n from $\{1, \dots, n\}$, and the selection probability of i is $p_i^\dagger = (C_N^*)^{-1} N_i^* p_{a,i}$ for $i = 1, \dots, n$. Denote m_i^* to be the number of i 's in the sample. Then, we have $T_{PPS}^* = (\hat{V}_{PPS}^*)^{-1/2} (\hat{Y}_{PPS}^* - Y^*)$, where $\hat{V}_{PPS}^* = n^{-2} \sum_{i=1}^n m_i^* (C_N^* p_{a,i}^{-1} y_{a,i} - \hat{Y}_{PPS}^*)^2$ and $\hat{Y}_{PPS}^* = n^{-1} \sum_{i=1}^n m_i^* C_N^* p_{a,i}^{-1} y_{a,i}$.

Step 3'. Repeat the above three steps independently M times.

The following regularity conditions are required to validate the proposed bootstrap method under PPS sampling.

(C9) There exists $\gamma \in (2^{-1}, 1]$ such that $n \asymp N^\gamma$, and the selection probabilities satisfy

$$C_4 \leq N p_i \leq C_5$$

for $i = 1, \dots, N$, where C_4 and C_5 are positive constants.

(C10) The sequence of finite populations and selection probabilities satisfy

$$\lim_{N \rightarrow \infty} (N^{-2} \sigma_{PPS}^2) = \sigma_3^2,$$

where $\sigma_{PPS}^2 = \sum_{i=1}^N p_i (p_i^{-1} y_i - Y)^2$, and σ_3^2 is a positive number.

(C11) The distribution $G_{N,PPS}$ is non-lattice, where $G_{N,PPS}$ assigns probability p_i to $p_i^{-1} y_i$ for $i = 1, \dots, N$.

Condition (C9) regulates the sample size and selection probabilities, and (C10) rules out the degenerate case under PPS sampling. To show (C9) and (C10) can be satisfied

simultaneously, take $p_i = N^{-1}$ for $i = 1, \dots, N$. Then, (C9) holds with $0 < C_4 < 1 < C_5$, and $\sigma_{PPS}^2 = \sum_{i=1}^N p_i (p_i^{-1} y_i - Y)^2 = N \sum_{i=1}^N y_i^2 - N^2 \bar{Y}^2$, where $\bar{Y} = N^{-1} Y$. Thus, $N^{-2} \sigma_{PPS}^2 = N^{-1} \sum_{i=1}^N y_i^2 - \bar{Y}^2$ converges if both $N^{-1} \sum_{i=1}^N y_i^2$ and \bar{Y} converge as $N \rightarrow \infty$. Since $G_{N,PPS}$ corresponds to the PPS sampling procedure, condition (C11) focuses our attention to the non-lattice case.

Based on a similar argument made under Poisson sampling, there exists a probability measure \mathbb{P}_{PPS} on $\mathcal{F} = \times_{N=1}^{\infty} \mathcal{F}_N$ equipped with the product σ -algebra \mathcal{B} under PPS sampling.

Lemma 5.1. *Suppose that (C3), (C9) and (C10) hold. Then,*

$$N^{-2}(s_{PPS}^2 - \sigma_{PPS}^2) \rightarrow 0 \quad (14)$$

as $N \rightarrow \infty$ almost surely (\mathbb{P}_{PPS}), where $s_{PPS}^2 = n^{-1} \sum_{i=1}^n (Z_i - \bar{Z}_n)^2$ is the sample variance of $\{Z_1, \dots, Z_n\}$. Let $\mu_{PPS}^{(3)} = \sum_{i=1}^N p_i (p_i^{-1} y_i - Y)^3$ and $\hat{\mu}_{PPS}^{(3)} = n^{-1} \sum_{i=1}^n Z_i^3 + 2\bar{Z}_n^3 - 3\bar{Z}_n n^{-1} \sum_{i=1}^n Z_i^2$, then

$$N^{-3} \mu_{PPS}^{(3)} = O(1) \quad \text{and} \quad N^{-3} (\hat{\mu}_{PPS}^{(3)} - \mu_{PPS}^{(3)}) = O_p(n^{-1/2}). \quad (15)$$

Lemma 5.1 shows convergence properties of estimators of the variance and third central moment. The next theorem deals with the Edgeworth expansion for the distribution of T_{PPS} .

Theorem 5.1. *Suppose that (C3), (C9)–(C11) hold. Then,*

$$\hat{F}_{PPS}(z) = \Phi(z) + \frac{\hat{\mu}_{PPS}^{(3)}}{6\sqrt{n}s_{PPS}^3} (2z^2 + 1)\phi(z) + o_p(n^{-1/2}) \quad (16)$$

uniformly in $z \in \mathbb{R}$, where $\hat{F}_{PPS} = \mathbb{P}_{PPS}(T_{PPS} \leq z)$ is the cumulative distribution function of T_{PPS} under PPS sampling.

Based on the result in Theorem 5.1, the Wald-type method may provide inefficient inference results for Y compared with the proposed bootstrap method if the sample size is small

and $\mu_{PPS}^{(3)} \neq 0$.

Theorem 5.2. *Suppose that (C3), (C9)–(C11) hold. Then, we have*

$$\hat{F}_{PPS}^*(z) = \Phi(z) + \frac{\hat{\mu}_{PPS}^{(3)}}{6\sqrt{n}s_{PPS}^3}(2z^2 + 1)\phi(z) + o_p(n^{-1/2}) \quad (17)$$

uniformly in $z \in \mathbb{R}$, where $\hat{F}_{PPS}^*(z)$ is the cumulative distribution function of T_{PPS}^* conditional on the bootstrap finite population \mathcal{F}_N^* .

Theorem 5.2 shows the Edgeworth expansion for the cumulative distribution function of T_{PPS}^* based on the proposed bootstrap method. By comparing (16) in Theorem 5.1 with (17) in Theorem 5.2, we have shown that the proposed bootstrap method is second-order accurate under PPS sampling.

6 Simulation study

6.1 Single-stage sampling designs

We conduct a simulation study based on single-stage sampling designs in this section. A finite population $\mathcal{F}_N = \{y_1, \dots, y_N\}$ is generated by

$$y_i \sim \text{Exp}(10)$$

for $i = 1, \dots, N$, where $\text{Exp}(\lambda)$ is an exponential distribution with a scale parameter λ , and the population size is $N = 500$, which is assumed to be known. The size measure is simulated by $z_i = \log(3 + s_i)$ for $i = 1, \dots, N$, where $s_i \mid y_i \sim \text{Exp}(y_i)$. The expected sample size is $n_0 \in \{10, 100\}$. We are interested in constructing a 90% confidence interval for the finite population mean \bar{Y} by survey data under the following sampling designs, and its true value is around 9.7.

1. Poisson sampling. The first-order inclusion probability is $\pi_i = n_0 z_i \left(\sum_{j=1}^N z_j \right)^{-1}$ for $i = 1, \dots, N$, and its expected sample size is n_0 .
2. SRS with sample size n_0 .
3. PPS sampling. The selection probability for this design is $p_i = z_i \left(\sum_{j=1}^N z_j \right)^{-1}$ for $i = 1, \dots, N$, and the sample size is n_0 .

Based on a sample, denote \tilde{V} to be the design-unbiased variance estimator of \tilde{Y} , where \tilde{Y} is the design-unbiased estimate of \bar{Y} under a specific sampling design. We consider the following methods to construct the 90% confidence interval.

Method I. Proposed bootstrap method by setting $M = 1\,000$. Denote $q_{B,0.05}$ and $q_{B,0.95}$ to be the 5%-th and 95%-th sample quantiles of $\{(\tilde{V}^{*(m)})^{-1/2}(\tilde{Y}^{*(m)} - \bar{Y}^{*(m)}) : m = 1, \dots, M\}$ obtained by the proposed bootstrap method, where $\tilde{V}^{*(m)}$, $\tilde{Y}^{*(m)}$ and $\bar{Y}^{*(m)}$ are the bootstrap counterparts of \tilde{V} , \tilde{Y} and \bar{Y} in the m -th repetition. Then, a 90% confidence interval for \bar{Y} can be constructed by

$$(\tilde{Y} - q_{B,0.95}\tilde{V}^{1/2}, \tilde{Y} - q_{B,0.05}\tilde{V}^{1/2}).$$

Method II. Wald-type method. A Wald-type 90% confidence interval for \bar{Y} is obtained by

$$(\tilde{Y} - q_{0.95}\tilde{V}^{1/2}, \tilde{Y} - q_{0.05}\tilde{V}^{1/2}),$$

where $q_{0.05}$ and $q_{0.95}$ are the 5%-th and 95%-th quantiles of the standard normal distribution.

We conduct 1 000 Monte Carlo simulations for each sampling design, and the two methods are compared in terms of the coverage rate and the length of the constructed 90% confidence interval. Table 1 summarizes the simulation results. When the sample size is small, the proposed bootstrap method is more preferable in the sense that its coverage rates are closer to

Table 1: Coverage rate and length of the constructed 90% confidence interval for the proposed bootstrap method (Bootstrap) and the Wald-type method (Wald-type) under single-stage sampling designs, including Poisson sampling (Poisson), SRS and PPS sampling (PPS). “C.R.” stands for the coverage rate, and “C.L.” presents the Monte Carlo mean of the lengths of the constructed confidence interval.

Design	Method	$n_0 = 10$		$n_0 = 100$	
		C.R.	C.L.	C.R.	C.L.
Poisson	Bootstrap	0.90	15.5	0.90	3.6
	Wald-type	0.84	12.1	0.88	3.6
SRS	Bootstrap	0.90	13.0	0.89	2.8
	Wald-type	0.83	9.1	0.89	2.8
PPS	Bootstrap	0.88	10.3	0.90	2.6
	Wald-type	0.83	7.5	0.89	2.6

0.9 compared with the Wald-type method under the three sampling designs. The confidence interval constructed by the proposed bootstrap method is wider compared with that by the Wald-type method. As the sample size increases to $n_0 = 100$, the performance of the two methods is approximately the same in the sense that the coverage rates of both methods are close to 0.9, and confidence interval lengths are approximately the same.

In addition, we also compare the two methods in terms of approximating the probability $\mathbb{P}\{\tilde{V}^{-1/2}(\tilde{Y} - \bar{Y}) \leq z\}$, which is obtained by 10 000 Monte Carlo simulations. We set $z \in \{-0.5, -0.25, -0.1, 0, 0.1, 0.25, 0.5\}$ as done by [Lai and Wang \(1993\)](#). [Table 2](#) summarizes the simulation results. For both sample sizes, the proposed bootstrap method can approximate the target distribution well, but the performance of the Wald-type method is not as good as the proposed one when sample size is small.

Table 2: Values of $P_z = \mathbb{P}\{\tilde{V}^{-1/2}(\tilde{Y} - \bar{Y}) \leq z\}$, the normal approximation $\Phi(z)$ and the bootstrap approximation Boot_z for three sampling designs including Poisson sampling (Poisson), SRS and PPS sampling (PPS). For convenience, we include the values $\Phi(z)$ for both sample sizes.

Design	z	$n_0 = 10$			$n_0 = 100$		
		P_z	$\Phi(z)$	Boot_z	P_z	$\Phi(z)$	Boot_z
Poisson	-0.5	0.37	0.31	0.36	0.32	0.31	0.32
	-0.25	0.45	0.40	0.44	0.41	0.40	0.41
	-0.1	0.50	0.46	0.49	0.47	0.46	0.47
	0	0.54	0.50	0.53	0.51	0.50	0.51
	0.1	0.58	0.54	0.57	0.55	0.54	0.55
	0.25	0.64	0.60	0.63	0.61	0.60	0.61
	0.5	0.73	0.69	0.73	0.70	0.69	0.70
SRS	-0.5	0.37	0.31	0.34	0.32	0.31	0.32
	-0.25	0.45	0.40	0.42	0.41	0.40	0.41
	-0.1	0.50	0.46	0.48	0.47	0.46	0.47
	0	0.54	0.50	0.52	0.51	0.50	0.51
	0.1	0.58	0.54	0.56	0.55	0.54	0.55
	0.25	0.63	0.60	0.62	0.61	0.60	0.61
	0.5	0.73	0.69	0.71	0.70	0.69	0.70
PPS	-0.5	0.37	0.31	0.34	0.33	0.31	0.32
	-0.25	0.45	0.40	0.42	0.42	0.40	0.41
	-0.1	0.50	0.46	0.48	0.47	0.46	0.47
	0	0.54	0.50	0.52	0.51	0.50	0.51
	0.1	0.57	0.54	0.56	0.55	0.54	0.55
	0.25	0.63	0.60	0.62	0.61	0.60	0.61
	0.5	0.72	0.69	0.71	0.70	0.69	0.70

6.2 Two-stage sampling designs

In this section, we test the performance of the proposed method under two-stage sampling designs. A finite population $\mathcal{F}_N = \{y_{i,j} : i = 1, \dots, H; j = 1, \dots, N_i\}$ is generated by

$$\begin{aligned}
 y_{i,j} &= 50 + a_i + e_{i,j}, \\
 a_i &\sim N(0, 50), \\
 e_{i,j} &\sim \text{Exp}(20), \\
 N_i \mid a_i &\sim \text{Poisson}(q_i) + c_0
 \end{aligned}$$

for $i = 1, \dots, H$ and $j = 1, \dots, N_i$, where $\text{Poisson}(\lambda)$ is a Poisson distribution with a rate parameter λ , $q_i = (a_i - 25)^2/20$, $c_0 = 40$ is the minimum cluster size, and $H = 100$ is the number of clusters in the finite population. The finite population size is $N = 7\,129$, and the cluster sizes range from 43 to 129. We assume that the finite population size N and cluster sizes N_1, \dots, N_H are known. We are interested in constructing a 90% confidence interval for the finite population mean $\bar{Y} = N^{-1} \sum_{i=1}^H \sum_{j=1}^{N_i} y_{i,j}$, where the true value of \bar{Y} is approximately 70.5.

We consider two different sampling designs for the first stage; one is Poisson sampling, and the other one is PPS sampling. The first-order inclusion probability (selection probability) of the i -th cluster is proportional to its cluster size N_i under Poisson (PPS) sampling for $i = 1, \dots, H$. SRS is conducted within each selected cluster independently in the second stage. The expected sample size of the first-stage sampling is n_1 , and that of the second-stage sampling is n_2 . In this simulation, we consider two scenarios for the sample sizes, that is, $(n_1, n_2) = (5, 10)$ and $(n_1, n_2) = (10, 30)$.

The derivations of the design-unbiased estimator \tilde{Y} and its variance estimator \tilde{V} under the two-stage sampling designs in this simulation study are presented in Appendix 9.2. We consider the following methods to construct the 90% confidence intervals for the parameters of interest.

Method I. The proposed method extended to a two-stage sampling design. This method is approximately the same as that mentioned in Section 6.1 with the following two steps to bootstrap the finite population, and we set $M = 1\,000$ for this method.

Step 1. Use the proposed method to bootstrap the H clusters by treating them as “elements”, and the original sample within each selected cluster are replicated accordingly.

Step 2. For each bootstrap cluster, apply the proposed method to bootstrap the cluster finite population independently.

Table 3: Coverage rate and length of the 90% confidence interval for \bar{Y} by the proposed bootstrap method (Bootstrap) and the Wald-type method (Wald-type) under two-stage sampling designs. The first column show the first-stage sample designs, that is, Poisson sampling (Poisson) and PPS sampling (PPS), and SRS is used in the second stage. “C.R.” shows the coverage rate, and “C.L.” presents the Monte Carlo mean of the length for the 90% confidence interval.

Design	(n_1, n_2)	Method	C.R.	C.L.
Poisson	(5,10)	Bootstrap	0.90	114.08
		Wald-type	0.85	98.58
	(10,30)	Bootstrap	0.90	73.92
		Wald-type	0.88	68.66
PPS	(5,10)	Bootstrap	0.89	17.56
		Wald-type	0.85	14.57
	(10,30)	Bootstrap	0.90	9.40
		Wald-type	0.86	8.24

Method II. Wald-type method, and it is the same as the one discussed in Section 6.1.

We conduct 1 000 Monte Carlo simulations for each scenario. Table 3 summarizes the coverage rate and average length of the constructed 90% confidence interval for the finite population mean. The coverage rates of the proposed bootstrap method are closer to 0.9 even when the sample size is limited. However, the coverage rates of the commonly used Wald-type method are not as good as the proposed bootstrap method. Specifically, the coverage rates of the Wald-type method are only around 0.86 for three scenarios, and it improves to 0.88 when sample size is large under Poisson sampling. The confidence intervals of the proposed bootstrap method are wider than those of the Wald-type method when sample size is small.

As in Section 6.1, we also compare those two methods in terms of approximating $\mathbb{P}\{\tilde{V}^{-1/2}(\tilde{Y} - \bar{Y}) \leq z\}$, which is obtained by 10 000 Monte Carlo simulations. We set $z \in \{-0.5, -0.25, -0.1, 0, 0.1, 0.25, 0.5\}$. Table 4 summarizes the simulation results. For both sample sizes, the proposed bootstrap method can approximate the target distribution well, but the performance of the Wald-type method is not as good as the proposed one especially when the sample size is small.

Table 4: Values of $P_z = \mathbb{P}\{\tilde{V}^{-1/2}(\tilde{Y} - \bar{Y}) \leq z\}$, the normal approximation $\Phi(z)$ and the bootstrap approximation Boot_z under two-stage sampling designs. The first column show the first-stage sample designs, that is, Poisson sampling (Poisson) and PPS sampling (PPS), and SRS is used in the second stage.

Design	(n_1, n_2)	z	P_z	$\Phi(z)$	Boot_z
Poisson	(5,10)	-0.5	0.35	0.31	0.35
		-0.25	0.43	0.40	0.43
		-0.1	0.48	0.46	0.49
		0	0.53	0.50	0.52
		0.1	0.58	0.54	0.56
		0.25	0.62	0.60	0.62
		0.5	0.73	0.69	0.72
	(10,30)	-0.5	0.34	0.31	0.33
		-0.25	0.43	0.40	0.42
		-0.1	0.48	0.46	0.48
		0	0.52	0.50	0.52
		0.1	0.56	0.54	0.56
		0.25	0.61	0.60	0.62
		0.5	0.72	0.69	0.71
PPS	(5,10)	-0.5	0.32	0.31	0.32
		-0.25	0.41	0.40	0.41
		-0.1	0.47	0.46	0.47
		0	0.51	0.50	0.51
		0.1	0.55	0.54	0.55
		0.25	0.61	0.60	0.62
		0.5	0.71	0.69	0.71
	(10,30)	-0.5	0.32	0.31	0.32
		-0.25	0.41	0.40	0.41
		-0.1	0.46	0.46	0.47
		0	0.50	0.50	0.50
		0.1	0.54	0.54	0.54
		0.25	0.60	0.60	0.60
		0.5	0.69	0.69	0.69

7 Conclusion

In this paper, we propose bootstrap methods for Poisson sampling, SRS and PPS sampling, and we show that the proposed bootstrap methods are second-order accurate. The first step of the proposed bootstrap methods corresponds to an inverse sampling procedure by

incorporating the sampling information. Since the proposed bootstrap method is based on an asymptotically pivotal statistic, it is necessary to estimate the variance of the design-unbiased estimator. Simulation results show that the proposed bootstrap method provides more conservative confidence interval than the Wald-type method when the sample size is small, and the 90% confidence interval constructed by the proposed bootstrap method has a better coverage rate. Although the proposed bootstrap method is discussed under the single-stage sampling designs, simulation shows that it works well under some two-stage sampling designs, and Edgeworth expansion for two-stage sampling designs are under investigation. It may be extended to other complex sampling designs when the asymptotic distribution of the design-unbiased estimator exists, but the second-order accuracy may not be guaranteed. Besides, the proposed bootstrap method can be easily parallelized in practice.

8 Acknowledgment

We would like to thank Dr. J. N. K. Rao for the suggestion to discuss the simple random sampling and the two anonymous reviewers for the detailed and constructive comments.

9 Supplement

9.1 Proofs

For the purpose of clarity, we explicitly express $y_{N,i}$, Y_N , $I_{N,i}$, $\pi_{N,i}$ and $p_{N,i}$ for y_i , Y , I_i , π_i and p_i to highlight that they are indexed by N , and the same notation is used for other quantities without further mentioning. Denote $E(\cdot | \mathcal{F}_N)$ and $\text{var}(\cdot | \mathcal{F}_N)$ to be the expectation and variance with respect to the probability measure of a specific sampling design, say \mathbb{P}_{Poi} under Poisson sampling, $E_*(\cdot)$ and $\text{var}_*(\cdot)$ to be the conditional mean and variance with respect to the multinomial distribution in the first steps of the proposed bootstrap method conditional on the realized sample $\{y_{N,1}, \dots, y_{N,n}\}$, and $E_{**}(\cdot)$ and $\text{var}_{**}(\cdot)$ to be the expectation and

variance with respect to the sampling design in the second step conditional on the bootstrap finite population \mathcal{F}_N^* .

Proof of Lemma 3.1. Denote $X_{N,i}^{(1)} = n_0 N^{-2} y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} (I_{N,i} - \pi_{N,i})$, then $n_0 N^{-2} (\hat{V}_{N,Poi} - V_{N,Poi}) = \sum_{i=1}^N X_{N,i}^{(1)}$. Let $D_N^{(1)}$ be the event $\{|\sum_{i=1}^N X_{N,i}^{(1)}| > \epsilon\}$ for $N \in \mathbb{N}_+$, where $\epsilon \in (0, \infty)$ and \mathbb{N}_+ is the set of positive integers.

By the Borel-Cantelli Lemma (Athreya and Lahiri; 2006, Theorem 7.2.2), to show (1), it is enough to prove

$$\sum_{N=1}^{\infty} \mathbb{P}_{Poi}(D_N^{(1)}) < \infty \quad (\text{A.1})$$

for $\epsilon > 0$. By the Markov's inequality (Athreya and Lahiri; 2006, Proposition 6.2.4), we have

$$\begin{aligned} \mathbb{P}_{Poi}(D_N^{(1)}) &\leq \epsilon^{-4} E \left\{ \left(\sum_{i=1}^N X_{N,i}^{(1)} \right)^4 \mid \mathcal{F}_N \right\} \\ &= \epsilon^{-4} \left[\sum_{i=1}^N E \left\{ \left(X_{N,i}^{(1)} \right)^4 \mid \mathcal{F}_N \right\} \right. \\ &\quad \left. + \sum_{(i,j) \in \Gamma_N} E \left\{ \left(X_{N,i}^{(1)} \right)^2 \mid \mathcal{F}_N \right\} E \left\{ \left(X_{N,j}^{(1)} \right)^2 \mid \mathcal{F}_N \right\} \right], \end{aligned}$$

where the last equality holds since $E(X_{N,i}^{(1)} \mid \mathcal{F}_N) = 0$ for $i = 1, \dots, N$, and $X_{N,i}^{(1)}$ is independent of $X_{N,j}^{(1)}$ for $(i, j) \in \Gamma_N$ with $\Gamma_N = \{(i, j) : i, j = 1, \dots, N \text{ and } i \neq j\}$.

Consider

$$\begin{aligned} &E \left\{ \left(X_{N,i}^{(1)} \right)^4 \mid \mathcal{F}_N \right\} \\ &= n_0^4 N^{-8} y_{N,i}^8 (1 - \pi_{N,i})^5 \pi_{N,i}^{-4} \{(1 - \pi_{N,i})^3 \pi_{N,i}^{-3} + 1\} \\ &\leq C_{1,1} n_0^{-3} N^{-1} y_{N,i}^8, \end{aligned} \quad (\text{A.2})$$

where $C_{1,1}$ is a positive constant determined by (C1).

Next, consider

$$\begin{aligned} E \left\{ \left(X_{N,i}^{(1)} \right)^2 \mid \mathcal{F}_N \right\} &= n_0^2 N^{-4} y_{N,i}^4 \pi_{N,i}^{-3} (1 - \pi_{N,i})^3 \\ &\leq C_{1,2} n_0^{-1} N^{-1} y_{N,i}^4, \end{aligned} \quad (\text{A.3})$$

where $C_{1,2}$ is a positive constant.

Based on some algebra and (C3), we have

$$\sum_{(i,j) \in \Gamma_N} y_{N,i}^4 y_{N,j}^4 = O(N^2). \quad (\text{A.4})$$

By (A.2), (A.3) and (A.4), we have

$$\begin{aligned} &\mathbb{P}_{Poi}(D_N^{(1)}) \\ &\leq \epsilon^{-4} C_{1,1} n_0^{-3} N^{-1} \sum_{i=1}^N y_{N,i}^8 + \epsilon^{-4} C_{1,2}^2 n_0^{-2} N^{-2} \sum_{(i,j) \in \Gamma_N} y_{N,i}^4 y_{N,j}^4 \\ &= O(N^{-2\alpha}) \end{aligned}$$

for any fixed $\epsilon > 0$, where the last inequality holds by (C3). Since $\alpha \in (2^{-1}, 1]$ by (C1), we have proved (1) based on (A.1).

For $\mu_{N,Poi}^{(3)} = \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1\}$, we have

$$\begin{aligned} \left| n_0^2 N^{-3} \mu_{Poi}^{(3)} \right| &= \left| n_0^2 N^{-3} \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1\} \right| \\ &\leq 2n_0^2 N^{-3} \sum_{i=1}^N |y_{N,i}|^3 \pi_{N,i}^{-2} \\ &\leq 2C_1^{-2} N^{-1} \sum_{i=1}^N |y_{N,i}|^3 = O(1), \end{aligned}$$

where the first inequality holds by $0 < \pi_{N,i} < 1$ and $0 < 1 - \pi_{N,i} < 1$, the second inequality holds by (C1), and the last equality holds by (C3).

Mentioned that

$$\begin{aligned}
& E \left(n_0^2 N^{-3} \hat{\mu}_{N, Poi}^{(3)} \mid \mathcal{F}_N \right) \\
&= E \left[n_0^2 N^{-3} \sum_{i=1}^n y_{N,i}^3 (1 - \pi_{N,i}) \pi_{N,i}^{-1} \{ (1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1 \} \mid \mathcal{F}_N \right] \\
&= n_0^2 N^{-3} \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i}) \{ (1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1 \} \\
&= n_0^2 N^{-3} \mu_{N, Poi}^{(3)}
\end{aligned}$$

and

$$\begin{aligned}
& \text{var} \left(\hat{\mu}_{N, Poi}^{(3)} \mid \mathcal{F}_N \right) \\
&= \text{var} \left[n_0^2 N^{-3} \sum_{i=1}^n y_{N,i}^3 (1 - \pi_{N,i}) \pi_{N,i}^{-1} \{ (1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1 \} \mid \mathcal{F}_N \right] \\
&\leq 4n_0^4 N^{-6} \sum_{i=1}^N y_{N,i}^6 \pi_{N,i}^{-5} \\
&= O(n_0^{-1}),
\end{aligned}$$

we can obtain that $n_0^2 N^{-3} (\hat{\mu}_{N, Poi}^{(3)} - \mu_{N, Poi}^{(3)}) = O_p(n_0^{-1/2})$ by the Markov's inequality. The results concerning $\tau_{N, Poi}^{(3)}$ and $\hat{\tau}_{N, Poi}^{(3)}$ can be proved similarly and is omitted here. Thus, we finalize the proof of Lemma 3.1. \square

The following lemmas are useful in establishing Theorem 3.1 and 3.2.

Lemma 9.1. Denote $X_{N,i} = V_{N, Poi}^{-1/2} y_{N,i} \pi_{N,i}^{-1} (I_{N,i} - \pi_{N,i})$ for $i = 1, \dots, N$. Let $\Delta_{N,1} = \sum_{i=1}^N X_{N,i}$ and $\phi_{\Delta_{N,1}}(t) = E \{ \exp(it\Delta_{N,1}) \mid \mathcal{F}_N \}$ be the characteristic function (c.f.) of $\Delta_{N,1}$, where ι is the imaginary unit. Then under conditions (C1)–(C3),

$$|\phi_{\Delta_{N,1}}(t)| \leq \exp(-t^2/3), \quad (\text{A.5})$$

$$|\phi_{\Delta_{N,1}}(t) - \exp(-t^2/2)| \leq 16|t|^3 V_{N, Poi}^{-3/2} \nu_{N, Poi}^{(3)} \exp(-t^2/3) \quad (\text{A.6})$$

for all $|t| \leq V_{N,Poi}^{3/2} / \left(4\nu_{N,Poi}^{(3)}\right)$, where $\nu_{N,Poi}^{(3)} = \sum_{i=1}^N |y_{N,i}|^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} + 1\}$.

Furthermore,

$$\begin{aligned} & \left| \phi_{\Delta_{N,1}}(t) - \exp(-t^2/2) - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \exp(-t^2/2) \right| \\ & \leq C_{2,1} \exp(-19t^2/48) (t^4 n_0^{-1} + t^6 n_0^{-1}) \end{aligned} \quad (\text{A.7})$$

for all $|t| \leq \min \left(\left\{ \max_{1 \leq i \leq N} E(X_{N,i}^2 | \mathcal{F}_N) \right\}^{-1/2}, V_{N,Poi}^{3/2} / \left(4\nu_{N,Poi}^{(3)}\right) \right)$, where $C_{2,1}$ is a positive constant and recall that $\mu_{N,Poi}^{(3)} = \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1\}$.

Proof. As $I_{N,i} \sim \text{Ber}(\pi_{N,i})$, $E(X_{N,i} | \mathcal{F}_N) = 0$ and $E(X_{N,i}^2 | \mathcal{F}_N) = V_{N,Poi}^{-1} y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-1}$ for $i = 1, \dots, N$. In addition,

$$E(X_{N,i}^3 | \mathcal{F}_N) = V_{N,Poi}^{-3/2} y_{N,i}^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1\}$$

and

$$E\{|X_{N,i}|^3 | \mathcal{F}_N\} = V_{N,Poi}^{-3/2} |y_{N,i}|^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} + 1\}.$$

Thus, $\sum_{i=1}^N E(X_{N,i}^2 | \mathcal{F}_N) = 1$, $\sum_{i=1}^N E(X_{N,i}^3 | \mathcal{F}_N) = V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} = O(n_0^{-1/2})$ by (C2) and Lemma 3.1. In addition, $\nu_{N,Poi}^{(3)} = \sum_{i=1}^N |y_{N,i}|^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} + 1\} = O(n_0^{-2} N^{-3})$, which implies $\sum_{i=1}^N E\{|X_{N,i}|^3 | \mathcal{F}_N\} = V_{N,Poi}^{-3/2} \nu_{N,Poi}^{(3)} = O(n_0^{-1/2})$. By the fact that $\max_{1 \leq i \leq N} E\{|X_{N,i}|^3 | \mathcal{F}_N\} \leq \sum_{i=1}^N E\{|X_{N,i}|^3 | \mathcal{F}_N\} = O(n_0^{-1/2})$, $E\{|X_{N,i}|^3 | \mathcal{F}_N\} < \infty$ for $i = 1, \dots, N$. By Lemma 5.1 of Petrov (1995),

$$\begin{aligned} |\phi_{\Delta_{N,1}}(t)| & \leq \exp(-t^2/3), \\ |\phi_{\Delta_{N,1}}(t) - \exp(-t^2/2)| & \leq 16|t|^3 V_{N,Poi}^{-3/2} \nu_{N,Poi}^{(3)} \exp(-t^2/3) \end{aligned}$$

for all $|t| \leq V_{N,Poi}^{3/2} / \left(4\nu_{N,Poi}^{(3)}\right)$.

It remains to show (A.7). Denote $\phi_{X_{N,i}}(t) = E\{\exp(tX_{N,i}) | \mathcal{F}_N\}$ as the characteristic

function of $X_{N,i}$. Note that for any complex numbers z, w , $|\exp(z) - 1 - w| \leq (|z - w| + |w|^2/2) \exp\{\max(|z|, |w|)\}$, it follows that

$$\begin{aligned}
& \left| \phi_{\Delta_{N,1}}(t) - \exp(-t^2/2) - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \exp(-t^2/2) \right| \\
&= \left| \prod_{i=1}^N \phi_{X_{N,i}}(t) - \exp(-t^2/2) - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \exp(-t^2/2) \right| \\
&= \exp(-t^2/2) \left| \exp \left[\sum_{i=1}^N \log\{\phi_{X_{N,i}}(t)\} + t^2/2 \right] - 1 - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \right| \\
&\leq \exp(-t^2/2) \left[\left| \sum_{i=1}^N \log\{\phi_{X_{N,i}}(t)\} + t^2/2 - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \right| \right. \\
&\quad \left. + 2^{-1} \left| 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \right|^2 \right] \\
&\quad \times \exp \left\{ \max \left(\left| \sum_{i=1}^N \log\{\phi_{X_{N,i}}(t)\} + t^2/2 \right|, \left| 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \right| \right) \right\}.
\end{aligned}$$

By Lemma 11.4.3 of [Athreya and Lahiri \(2006\)](#),

$$\begin{aligned}
& \left| \sum_{i=1}^N \log\{\phi_{X_{N,i}}(t)\} + t^2/2 - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \right| \tag{A.8} \\
&\leq \sum_{i=1}^N \left| \log\{\phi_{X_{N,i}}(t)\} - 2^{-1}(t)^2 E(X_{N,i}^2 | \mathcal{F}_N) - 6^{-1}(t)^3 E(X_{N,i}^3 | \mathcal{F}_N) \right| \\
&\leq \sum_{i=1}^N (E[\min\{|tX_{N,i}|^3/3, (tX_{N,i})^4/24\} | \mathcal{F}_N] + t^4 \{E(X_{N,i}^2 | \mathcal{F}_N)\}^2/4) \\
&\leq C_{2,2} t^4 n_0^{-1}
\end{aligned}$$

for all $|t| \leq \{\max_{1 \leq i \leq N} E(X_{N,i}^2 | \mathcal{F}_N)\}^{-1/2}$, where $C_{2,2}$ is a positive constant and the last inequality is by the fact that

$$\begin{aligned}
& \sum_{i=1}^N E(X_{N,i}^4 | \mathcal{F}_N) \\
&= V_{N,Poi}^{-2} \sum_{i=1}^N y_{N,i}^4 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^3 \pi_{N,i}^{-3} + 1\} \\
&= O(n_0^{-1}).
\end{aligned}$$

Similarly, by Lemma 11.4.3 of [Athreya and Lahiri \(2006\)](#),

$$\begin{aligned}
& \left| \sum_{i=1}^N \log\{\phi_{X_{N,i}}(t)\} + t^2/2 \right| \\
& \leq \sum_{i=1}^N \left| \log\{\phi_{X_{N,i}}(t)\} - 2^{-1}(t)^2 E(X_{N,i}^2 \mid \mathcal{F}_N) \right| \\
& \leq 5|t|^3 \sum_{i=1}^N E\{|X_{N,i}|^3 \mid \mathcal{F}_N\}/12 = 5|t|^3 V_{N,Poi}^{-3/2} \nu_{N,Poi}^{(3)}/12
\end{aligned}$$

for all $|t| \leq \{\max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N)\}^{-1/2}$.

Thus, if $|t| \leq \min\left(\{\max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N)\}^{-1/2}, V_{N,Poi}^{3/2}/\left(4\nu_{N,Poi}^{(3)}\right)\right)$,

$$\begin{aligned}
& \max\left(\left|\sum_{i=1}^N \log\{\phi_{X_{N,i}}(t)\} + t^2/2\right|, \left|6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)}\right|\right) \\
& \leq 5t^2/48.
\end{aligned} \tag{A.9}$$

Mentioned that

$$2^{-1} \left|6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)}\right|^2 \leq C_{2,3} t^6 n_0^{-1} \tag{A.10}$$

for some positive constant $C_{2,3}$.

Finally, by [\(A.8\)](#), [\(A.9\)](#) and [\(A.10\)](#), it follows that

$$\begin{aligned}
& \left| \phi_{\Delta_{N,1}}(t) - \exp(-t^2/2) - 6^{-1}(t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \exp(-t^2/2) \right| \\
& \leq \exp(-t^2/2) (C_{2,2} t^4 n_0^{-1} + C_{2,3} t^6 n_0^{-1}) \exp(5t^2/48) \\
& \leq C_{2,1} \exp(-19t^2/48) (t^4 n_0^{-1} + t^6 n_0^{-1})
\end{aligned}$$

for all $|t| \leq \min\left(\{\max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N)\}^{-1/2}, V_{N,Poi}^{3/2}/\left(4\nu_{N,Poi}^{(3)}\right)\right)$. This concludes the proof of this Lemma. \square

Lemma 9.2. *Denote*

$$\Delta_{N,2} = -2^{-1}V_{N,Poi}^{-3/2} \sum_{(i,j) \in \Gamma_N} y_{N,i}y_{N,j}^2(1 - \pi_{N,j})\pi_{N,i}^{-1}\pi_{N,j}^{-2}(I_{N,i} - \pi_{N,i})(I_{N,j} - \pi_{N,j}),$$

then under conditions (C1)–(C3),

$$\begin{aligned} & E \{ \Delta_{N,2} \exp(ut\Delta_{N,1}) \mid \mathcal{F}_N \} \\ &= 2^{-1}t^2 \exp(-t^2/2)V_{N,Poi}^{-5/2}\Theta_{N,Poi}^{(2,3)} + \varpi_{N,1}(t), \end{aligned}$$

for all $|t| \leq \{ \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \}^{-1/2}$, where $X_{N,i} = V_{N,Poi}^{-1/2}y_{N,i}\pi_{N,i}^{-1}(I_{N,i} - \pi_{N,i})$, $\Delta_{N,1} = \sum_{i=1}^N X_{N,i}$, $\Gamma_N = \{(i, j) : i, j = 1, \dots, N \text{ and } i \neq j\}$,

$$\Theta_{N,Poi}^{(2,3)} = \sum_{(i,j) \in \Gamma_N} y_{N,i}^2 y_{N,j}^3 (1 - \pi_{N,i})(1 - \pi_{N,j})^2 \pi_{N,i}^{-1} \pi_{N,j}^{-2},$$

and $\varpi_{N,1}(t)$ satisfies

$$\begin{aligned} & |\varpi_{N,1}(t)| \\ & \leq C_{3,1} \exp \left\{ -t^2/2 + 2|t|^3 V_{N,Poi}^{-3/2} \nu_{N,Poi}^{(3)} / 3 + t^2 \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\} \\ & \quad \times \left(|t|^3 n_0^{-1} + t^4 n_0^{-3/2} + |t|^5 n_0^{-1} \right), \end{aligned}$$

with $C_{3,1}$ being a positive constant and $\nu_{N,Poi}^{(3)} = \sum_{i=1}^N |y_{N,i}|^3 (1 - \pi_{N,i}) \{ (1 - \pi_{N,i})^2 \pi_{N,i}^{-2} + 1 \}$.

Proof. First, write

$$E \{ (I_{N,i} - \pi_{N,i}) \exp(utX_{N,i}) \mid \mathcal{F}_N \} = utV_{N,Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i}) + \varpi_{N1,i}(t),$$

for $i = 1, \dots, N$, where

$$\begin{aligned}
& |\varpi_{N1,i}(t)| \\
&= |E \{(I_{N,i} - \pi_{N,i}) \exp(itX_{N,i}) \mid \mathcal{F}_N\} - itV_{N, Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i})| \\
&= (1 - \pi_{N,i})\pi_{N,i} \left| \left[\exp \left\{ itV_{N, Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i})\pi_{N,i}^{-1} \right\} - 1 \right. \right. \\
&\quad \left. \left. - itV_{N, Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i})\pi_{N,i}^{-1} \right] \right. \\
&\quad \left. - \left\{ \exp \left(-itV_{N, Poi}^{-1/2}y_{N,i} \right) - 1 + itV_{N, Poi}^{-1/2}y_{N,i} \right\} \right| \\
&\leq 2^{-1}(1 - \pi_{N,i})\pi_{N,i} \left\{ \left| itV_{N, Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i})\pi_{N,i}^{-1} \right|^2 + \left| itV_{N, Poi}^{-1/2}y_{N,i} \right|^2 \right\} \\
&\leq C_{3,2}t^2N^{-1}y_{N,i}^2,
\end{aligned}$$

and $C_{3,2}$ is a positive constant. The last but one inequality is due to the fact that for any real number x , $|\exp(\iota x) - 1 - \iota x| \leq |x|^2/2$. As a consequence, for any $(i, j) \in \Gamma_N$,

$$\begin{aligned}
& E \{(I_{N,i} - \pi_{N,i}) \exp(itX_{N,i}) \mid \mathcal{F}_N\} E \{(I_{N,j} - \pi_{N,j}) \exp(itX_{N,j}) \mid \mathcal{F}_N\} \\
&= \left\{ itV_{N, Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i}) + \varpi_{N1,i}(t) \right\} \left\{ itV_{N, Poi}^{-1/2}y_{N,j}(1 - \pi_{N,j}) + \varpi_{N1,j}(t) \right\} \\
&= -t^2V_{N, Poi}^{-1}y_{N,i}y_{N,j}(1 - \pi_{N,i})(1 - \pi_{N,j}) + \varpi_{N2,ij}(t),
\end{aligned}$$

where

$$\begin{aligned}
& |\varpi_{N2,ij}(t)| \\
&\leq \left| itV_{N, Poi}^{-1/2}y_{N,i}(1 - \pi_{N,i})\varpi_{N1,j}(t) \right| + \left| itV_{N, Poi}^{-1/2}y_{N,j}(1 - \pi_{N,j})\varpi_{N1,i}(t) \right| \\
&\quad + |\varpi_{N1,i}(t)\varpi_{N1,j}(t)| \\
&\leq C_{3,3} \left\{ |t|^3n_0^{1/2}N^{-2}(y_{N,i}^2|y_{N,j}| + |y_{N,i}|y_{N,j}^2) + t^4N^{-2}y_{N,i}^2y_{N,j}^2 \right\}
\end{aligned}$$

with $C_{3,3}$ being a positive constant.

Denote $\phi_{X_{N,i}}(t) = E \{\exp(itX_{N,i}) \mid \mathcal{F}_N\}$ for $i = 1, \dots, N$. By the same technique as in

the proof of Lemma 5.1 of [Petrov \(1995\)](#), we can show that

$$\begin{aligned} & \left| \prod_{k \neq i, j} \phi_{X_{N,k}}(t) \right| \\ & \leq \exp \left\{ -t^2 \sum_{k \neq i, j} E(X_{N,k}^2 | \mathcal{F}_N)/2 + 2|t|^3 \sum_{k \neq i, j} E\{|X_{N,k}|^3 | \mathcal{F}_N\}/3 \right\}. \end{aligned}$$

Using the inequality $|\exp(z) - 1| \leq |z| \exp(|z|)$ for all complex number z , we can obtain that

$$\begin{aligned} & \left| \prod_{k \neq i, j} \phi_{X_{N,k}}(t) - \exp(-t^2/2) \right| \\ & = \exp(-t^2/2) \left| \exp \left[\sum_{k \neq i, j} \log\{\phi_{X_{N,k}}(t)\} + t^2/2 \right] - 1 \right| \\ & \leq \exp(-t^2/2) \left| \sum_{k \neq i, j} \log\{\phi_{X_{N,k}}(t)\} + t^2/2 \right| \exp \left[\left| \sum_{k \neq i, j} \log\{\phi_{X_{N,k}}(t)\} + t^2/2 \right| \right] \end{aligned}$$

By Lemma 11.4.3 of [Athreya and Lahiri \(2006\)](#),

$$\begin{aligned} & \left| \sum_{k \neq i, j} \log\{\phi_{X_{N,k}}(t)\} + t^2/2 \right| \\ & \leq \sum_{k \neq i, j} \left| \log\{\phi_{X_{N,k}}(t)\} - 2^{-1}(it)^2 E(X_{N,k}^2 | \mathcal{F}_N) \right| \\ & \quad + 2^{-1}t^2 E(X_{N,i}^2 | \mathcal{F}_N) + 2^{-1}t^2 E(X_{N,j}^2 | \mathcal{F}_N) \\ & \leq 5|t|^3 \sum_{k \neq i, j} E\{|X_{N,k}|^3 | \mathcal{F}_N\}/12 + 2^{-1}t^2 \{E(X_{N,i}^2 | \mathcal{F}_N) + E(X_{N,j}^2 | \mathcal{F}_N)\}. \end{aligned}$$

for all $|t| \leq \left\{ \max_{1 \leq i \leq N} E(X_{N,i}^2 | \mathcal{F}_N) \right\}^{-1/2}$. Thus, we obtain

$$\begin{aligned} & \left| \prod_{k \neq i,j} \phi_{X_{N,k}}(t) - \exp(-t^2/2) \right| \\ & \leq \exp \left\{ -t^2 \sum_{k \neq i,j} E(X_{N,k}^2 | \mathcal{F}_N)/2 + 5|t|^3 \sum_{k \neq i,j} E\{|X_{N,k}|^3 | \mathcal{F}_N\}/12 \right\} \\ & \quad \times \left[5|t|^3 \sum_{k \neq i,j} E\{|X_{N,k}|^3 | \mathcal{F}_N\}/12 \right. \\ & \quad \left. + 2^{-1}t^2 \{E(X_{N,i}^2 | \mathcal{F}_N) + E(X_{N,j}^2 | \mathcal{F}_N)\} \right] \end{aligned}$$

for all $|t| \leq \left\{ \max_{1 \leq i \leq N} E(X_{N,i}^2 | \mathcal{F}_N) \right\}^{-1/2}$.

Finally, we have

$$\begin{aligned} & E \{ \Delta_{N,2} \exp(it\Delta_{N,1}) | \mathcal{F}_N \} \\ & = -2^{-1}V_{N,Poi}^{-3/2} \sum_{(i,j) \in \Gamma_N} y_{N,i}y_{N,j}^2(1 - \pi_{N,j})\pi_{N,i}^{-1}\pi_{N,j}^{-2} \\ & \quad \times E \{ (I_{N,i} - \pi_{N,i})(I_{N,j} - \pi_{N,j}) \exp(it\Delta_{N,1}) | \mathcal{F}_N \} \\ & = -2^{-1}V_{N,Poi}^{-3/2} \sum_{(i,j) \in \Gamma_N} y_{N,i}y_{N,j}^2(1 - \pi_{N,j})\pi_{N,i}^{-1}\pi_{N,j}^{-2} \prod_{k \neq i,j} \phi_{X_{N,k}}(t) \\ & \quad \times E \{ (I_{N,i} - \pi_{N,i}) \exp(itX_{N,i}) | \mathcal{F}_N \} \\ & \quad \times E \{ (I_{N,j} - \pi_{N,j}) \exp(itX_{N,j}) | \mathcal{F}_N \} \\ & = -2^{-1}V_{N,Poi}^{-3/2} \sum_{(i,j) \in \Gamma_N} y_{N,i}y_{N,j}^2(1 - \pi_{N,j})\pi_{N,i}^{-1}\pi_{N,j}^{-2} \prod_{k \neq i,j} \phi_{X_{N,k}}(t) \\ & \quad \times \{ -t^2V_{N,Poi}^{-1}y_{N,i}y_{N,j}(1 - \pi_{N,i})(1 - \pi_{N,j}) + \varpi_{N2,ij}(t) \} \\ & = 2^{-1}t^2V_{N,Poi}^{-5/2} \sum_{(i,j) \in \Gamma_N} y_{N,i}^2y_{N,j}^3(1 - \pi_{N,i})(1 - \pi_{N,j})^2\pi_{N,i}^{-1}\pi_{N,j}^{-2} \prod_{k \neq i,j} \phi_{X_{N,k}}(t) \\ & \quad - 2^{-1}V_{N,Poi}^{-3/2} \sum_{i \neq j} y_{N,i}y_{N,j}^2(1 - \pi_{N,j})\pi_{N,i}^{-1}\pi_{N,j}^{-2} \prod_{k \neq i,j} \phi_{X_{N,k}}(t)\varpi_{N2,ij}(t) \\ & = 2^{-1}t^2 \exp(-t^2/2)V_{N,Poi}^{-5/2}\Theta_{N,Poi}^{(2,3)} + \varpi_{N,1}(t) \end{aligned}$$

for all $|t| \leq \left\{ \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\}^{-1/2}$ and $\varpi_{N,1}(t)$ satisfies

$$\begin{aligned}
& |\varpi_{N,1}(t)| \\
\leq & C_{3,4} \exp \left\{ -t^2/2 + 2|t|^3 \sum_{i=1}^N E\{|X_{N,i}|^3 \mid \mathcal{F}_N\}/3 + t^2 \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\} \\
& \times \left\{ |t|^3 n_0^{-1} N^{-2} \sum_{i \neq j} (|y_{N,i}|^3 |y_{N,j}|^3 + y_{N,i}^2 y_{N,j}^4) \right. \\
& \quad + t^4 n_0^{-3/2} N^{-2} \sum_{i \neq j} |y_{N,i}|^3 y_{N,j}^4 \\
& \quad + t^4 n_0^{-1/2} N^{-3} \sum_{i \neq j} (y_{N,i}^4 |y_{N,j}|^3 + y_{N,i}^2 |y_{N,j}|^5) \\
& \quad \left. + |t|^5 n_0^{-1} \left(N^{-2} \sum_{i \neq j} y_{N,i}^2 |y_{N,j}|^3 \right) \left(N^{-1} \sum_{i=1}^N |y_{N,i}|^3 \right) \right\} \\
\leq & C_{3,1} \exp \left\{ -t^2/2 + 2|t|^3 V_{N,Poi}^{-3/2} \mathcal{V}_{N,Poi}^{(3)}/3 + t^2 \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\} \\
& \times \left(|t|^3 n_0^{-1} + t^4 n_0^{-3/2} + |t|^5 n_0^{-1} \right),
\end{aligned}$$

where $C_{3,4}$ is a positive constant. □

Lemma 9.3. Denote $\hat{Y}_{N,Poi} = \sum_{i=1}^N y_{N,i} \pi_{N,i}^{-1} I_{N,i}$, $Y_N = \sum_{i=1}^N y_{N,i}$ and $\hat{V}_{N,Poi} = \sum_{i=1}^N y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} I_{N,i}$, then under conditions (C1)–(C3),

$$\hat{V}_{N,Poi}^{-1/2} \left(\hat{Y}_{N,Poi} - Y_N \right) = \Delta_{N,1} + \Delta_{N,2} + \Delta_{N,3} + \Delta_{N,4},$$

where

$$\begin{aligned}
\Delta_{N,1} &= V_{N,Poi}^{-1/2} \sum_{i=1}^N y_{N,i} \pi_{N,i}^{-1} (I_{N,i} - \pi_{N,i}), \\
\Delta_{N,2} &= -2^{-1} V_{N,Poi}^{-3/2} \sum_{(i,j) \in \Gamma_N} y_{N,i} y_{N,j}^2 (1 - \pi_{N,j}) \pi_{N,i}^{-1} \pi_{N,j}^{-2} (I_{N,i} - \pi_{N,i}) (I_{N,j} - \pi_{N,j}), \\
\Delta_{N,3} &= -2^{-1} V_{N,Poi}^{-3/2} \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i})^2 \pi_{N,i}^{-2},
\end{aligned}$$

and recall that $\Gamma_N = \{(i, j) : i, j = 1, \dots, N \text{ and } i \neq j\}$. In addition, $\Delta_{N,4}$ satisfies

$$\mathbb{P}_{Poi} \left(|\Delta_{N,4}| \geq n_0^{-1/2} (\log n_0)^{-1} \right) = o(n_0^{-1/2}).$$

Proof. Denote $\Lambda_{N,1} = V_{N,Poi}^{-1} \sum_{i=1}^N y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} (I_{N,i} - \pi_{N,i})$. Mentioned that $E(\Lambda_{N,1} | \mathcal{F}_N) = 0$ and

$$E(\Lambda_{N,1}^2 | \mathcal{F}_N) = V_{N,Poi}^{-2} \sum_{i=1}^N y_{N,i}^4 (1 - \pi_{N,i})^3 \pi_{N,i}^{-3} = O(n_0^{-1}),$$

we have that $\Lambda_{N,1} = O_p(n_0^{-1/2})$. In addition,

$$\begin{aligned} & E(\Lambda_{N,1}^4 | \mathcal{F}_N) \\ &= E \left[\left\{ V_{N,Poi}^{-1} \sum_{i=1}^N y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} (I_{N,i} - \pi_{N,i}) \right\}^4 \mid \mathcal{F}_N \right] \\ &= V_{N,Poi}^{-4} \left[\sum_{i=1}^N y_{N,i}^8 (1 - \pi_{N,i})^5 \pi_{N,i}^{-4} \{ (1 - \pi_{N,i})^3 \pi_{N,i}^{-3} + 1 \} \right. \\ &\quad \left. + \sum_{(i,j) \in \Gamma_N} y_{N,i}^4 y_{N,j}^4 (1 - \pi_{N,i})^3 (1 - \pi_{N,j})^3 \pi_{N,i}^{-3} \pi_{N,j}^{-3} \right] \\ &= O(n_0^{-2}). \end{aligned}$$

By some algebra,

$$\begin{aligned} & \hat{V}_{N,Poi}^{-1/2} \left(\hat{Y}_{N,Poi} - Y_N \right) \\ &= \left\{ V_{N,Poi}^{-1/2} \sum_{i=1}^N y_{N,i} \pi_{N,i}^{-1} (I_{N,i} - \pi_{N,i}) \right\} (1 + \Lambda_{N,1})^{-1/2} \\ &= \left\{ V_{N,Poi}^{-1/2} \sum_{i=1}^N y_{N,i} \pi_{N,i}^{-1} (I_{N,i} - \pi_{N,i}) \right\} \{ 1 - 2^{-1} \Lambda_{N,1} + O_p(\Lambda_{N,1}^2) \}. \end{aligned}$$

Use the notations of $\Delta_{N,1}$, $\Delta_{N,2}$ and $\Delta_{N,3}$, we have

$$\begin{aligned}
& \hat{V}_{N,Poi}^{-1/2} \left(\hat{Y}_{N,Poi} - Y_N \right) \\
&= \Delta_{N,1} \left\{ 1 - 2^{-1} \Lambda_{N,1} + O_p(\Lambda_{N,1}^2) \right\} \\
&= \Delta_{N,1} - 2^{-1} \Delta_{N,1} \Lambda_{N,1} + O_p(\Delta_{N,1} \Lambda_{N,1}^2) \\
&= \Delta_{N,1} + \Delta_{N,2} + \Delta_{N,3} + \Lambda_{N,2} + O_p(\Delta_{N,1} \Lambda_{N,1}^2)
\end{aligned}$$

where

$$\Lambda_{N,2} = -2^{-1} V_{N,Poi}^{-3/2} \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i}) \pi_{N,i}^{-3} \left\{ (I_{N,i} - \pi_{N,i})^2 - E(I_{N,i} - \pi_{N,i})^2 \right\}.$$

It remains to show

$$\mathbb{P}_{Poi} \left(|\Lambda_{N,2}| \geq n_0^{-1/2} (\log n_0)^{-1} \right) = o(n_0^{-1/2}) \tag{A.11}$$

and

$$\mathbb{P}_{Poi} \left(|\Delta_{N,1} \Lambda_{N,1}^2| \geq n_0^{-1/2} (\log n_0)^{-1} \right) = o(n_0^{-1/2}). \tag{A.12}$$

For (A.11), it is a directly consequence of

$$E(\Lambda_{N,2} | \mathcal{F}_N) = 0$$

and

$$E(\Lambda_{N,2}^2 | \mathcal{F}_N) = O(n_0^{-2}).$$

For (A.12), as

$$\begin{aligned}
& E(\Delta_{N,1}^4 \mid \mathcal{F}_N) \\
&= \sum_{i=1}^N E(X_{N,i}^4 \mid \mathcal{F}_N) + \sum_{(i,j) \in \Gamma_N} E(X_{N,i}^2 \mid \mathcal{F}_N) E(X_{N,j}^2 \mid \mathcal{F}_N) \\
&= O(1),
\end{aligned}$$

where $X_{N,i} = V_{N,Poi}^{-1/2} y_{N,i} \pi_{N,i}^{-1} (I_{N,i} - \pi_{N,i})$, we have that

$$\begin{aligned}
& \mathbb{P}_{Poi} \left(|\Delta_{N,1} \Lambda_{N,1}^2| \geq n_0^{-1/2} (\log n_0)^{-1} \right) \\
&\leq \mathbb{P}_{Poi} \left(|\Delta_{N,1}| \geq n_0^{1/8} (\log n_0) \right) + \mathbb{P}_{Poi} \left(|\Lambda_{N,1}^2| \geq n_0^{-5/8} (\log n_0)^{-2} \right) \\
&\leq n_0^{-1/2} (\log n_0)^{-4} E(\Delta_{N,1}^4 \mid \mathcal{F}_N) + n_0^{5/4} (\log n_0)^4 E(\Lambda_{N,1}^4 \mid \mathcal{F}_N) \\
&= o(n_0^{-1/2}).
\end{aligned}$$

Thus, we finish the proof of this lemma. \square

Lemma 9.4. *Assume condition (C3) holds. Then for any positive integer s satisfies $s \rightarrow \infty$ as $N \rightarrow \infty$ and $s = o(N)$, there exists a subset $\{y_{N,\ell_1}, \dots, y_{N,\ell_s}\} \subset \{y_{N,1}, \dots, y_{N,N}\}$ such that*

$$\lim_{s \rightarrow \infty} s^{-1} \sum_{i=1}^s y_{N,\ell_i}^4 < \infty. \tag{A.13}$$

Proof. We prove this lemma by contradiction.

First, we split the population $\{y_{N,1}, \dots, y_{N,N}\}$ into $\lfloor N/s \rfloor$ subsets, with the first $\lfloor N/s \rfloor - 1$ subsets as $\{y_{(j-1)s+1}, \dots, y_{N,js}\}_{j=1}^{\lfloor N/s \rfloor - 1}$ and the last subset as $\{y_{(\lfloor N/s \rfloor - 1)s+1}, \dots, y_{N,N}\}$. Here $\lfloor x \rfloor$ denotes the integer part of $x \in \mathbb{R}$. Assume that (A.13) is not satisfied by any subset of $\{y_{N,1}, \dots, y_{N,N}\}$ with cardinality s . Then

$$\lim_{s \rightarrow \infty} s^{-1} \sum_{i=1}^s y_{N,(j-1)s+i}^4 = \infty,$$

for all $j = 1, \dots, \lfloor N/s \rfloor - 1$. This implies

$$\begin{aligned}
& N^{-1} \sum_{i=1}^N y_{N,i}^4 \\
&= N^{-1} s \sum_{j=1}^{\lfloor N/s \rfloor - 1} s^{-1} \sum_{i=1}^s y_{N,(j-1)s+i}^4 + N^{-1} \sum_{i=(\lfloor N/s \rfloor - 1)s+1}^N y_{N,i}^4 \\
&\rightarrow \infty
\end{aligned}$$

as $N \rightarrow \infty$.

This result is contradicted with condition (C3). Thus, there exists at least one subset of $\{y_{N,1}, \dots, y_{N,N}\}$ with cardinality s satisfies (A.13). \square

Lemma 9.5. Denote $\hat{Y}_{N,Poi} = \sum_{i=1}^N y_{N,i} \pi_{N,i}^{-1} I_{N,i}$, $Y_N = \sum_{i=1}^N y_{N,i}$ and $\hat{V}_{N,Poi} = \sum_{i=1}^N y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} I_{N,i}$. Let $\hat{F}_{N,Poi}(z) = \mathbb{P}_{Poi}(T_{N,Poi} \leq z)$ be the cumulative distribution function (cdf) of $T_{N,Poi}$ under Poisson sampling, where $T_{N,Poi} = \hat{V}_{N,Poi}^{-1/2} (\hat{Y}_{N,Poi} - Y_N)$. Then, under conditions (C1)–(C4),

$$\hat{F}_{N,Poi}(z) = \Phi(z) + \left(\frac{\mu_{N,Poi}^{(3)}}{6V_{N,Poi}^{3/2}} (1 - z^2) + \frac{\tau_{N,Poi}^{(3)}}{2V_{N,Poi}^{3/2}} z^2 \right) \phi(z) + o(n_0^{-1/2})$$

uniformly in $z \in \mathbb{R}$, where

$$\mu_{N,Poi}^{(3)} = \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1\}$$

and

$$\tau_{N,Poi}^{(3)} = \sum_{i=1}^N y_{N,i}^3 (1 - \pi_{N,i})^2 \pi_{N,i}^{-2}.$$

Proof. According to Lemma 9.3, $T_{N,Poi} = \Delta_{N,1} + \Delta_{N,2} + \Delta_{N,3} + \Delta_{N,4}$, where $\Delta_{N,1}$, $\Delta_{N,2}$,

$\Delta_{N,3}$ are defined in Lemma 9.3 and $\Delta_{N,4}$ satisfies

$$\mathbb{P}_{Poi} \left(|\Delta_{N,4}| \geq n_0^{-1/2} (\log n_0)^{-1} \right) = o(n_0^{-1/2}).$$

Thus, it suffices to show that

$$\sup_{z \in \mathbb{R}} |\mathbb{P}_{Poi}(\Delta_{N,1} + \Delta_{N,2} + \Delta_{N,3} \leq z) - F_{E,N}(z)| = o(n_0^{-1/2}),$$

where

$$F_{E,N}(z) = \Phi(z) + \left(6^{-1} V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} (1 - z^2) + 2^{-1} V_{N,Poi}^{-3/2} \tau_{N,Poi}^{(3)} z^2 \right) \phi(z).$$

Define

$$F_{E,N1}(z) = \Phi(z) + 6^{-1} \left(V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} - 3 V_{N,Poi}^{-3/2} \tau_{N,Poi}^{(3)} \right) (1 - z^2) \phi(z).$$

As $\Delta_{N,3}$ is nonrandom and $\Delta_{N,3} = O(n_0^{-1/2})$, by the fact that

$$\sup_{z \in \mathbb{R}} |F_{E,N1}(z - \Delta_{N,3}) - F_{E,N}(z)| = o(n_0^{-1/2}),$$

it is enough to prove that

$$\sup_{z \in \mathbb{R}} |\mathbb{P}_{Poi}(\Delta_{N,1} + \Delta_{N,2} \leq z) - F_{E,N1}(z)| = o(n_0^{-1/2}). \quad (\text{A.14})$$

Denote $W_N = \Delta_{N,1} + \Delta_{N,2}$ and let $\phi_{W_N}(t)$ be the characteristic function (c.f.) of W_N , that is

$$\phi_{W_N}(t) = E\{\exp(itW_N) \mid \mathcal{F}_N\} = E[\exp\{it(\Delta_{N,1} + \Delta_{N,2})\} \mid \mathcal{F}_N].$$

In addition, denote

$$\begin{aligned}
& \phi_{E,N1}(t) \\
&= \int \exp(tz) dF_{E,N1}(z) \\
&= \exp(-t^2/2) \left\{ 1 + 6^{-1}(it)^3 \left(V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} - 3V_{N,Poi}^{-3/2} \tau_{N,Poi}^{(3)} \right) \right\}.
\end{aligned}$$

By Esseen's smoothing lemma ([Petrov; 1995](#), Theorem 5.1), for any arbitrary $\varepsilon \in (0, 1)$,

$$\begin{aligned}
& \sup_{z \in \mathbb{R}} |\mathbb{P}_{Poi}(W_N \leq z) - F_{E,N1}(z)| \\
&\leq \frac{1}{\pi} \int_{|t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N}(t) - \phi_{E,N1}(t)| dt + n_0^{-1/2} \varepsilon,
\end{aligned}$$

where a_ε is chosen to satisfy $|dF_{E,N1}(z)/dx| \leq a_\varepsilon \varepsilon$. Thus, it suffices to prove

$$\frac{1}{\pi} \int_{|t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N}(t) - \phi_{E,N1}(t)| dt = o(n_0^{-1/2}).$$

Recall $\Theta_{N,Poi}^{(2,3)} = \sum_{(i,j) \in \Gamma_N} y_{N,i}^2 y_{N,j}^3 (1 - \pi_{N,i})(1 - \pi_{N,j})^2 \pi_{N,i}^{-1} \pi_{N,j}^{-2}$, where $\Gamma_N = \{(i, j) : i, j = 1, \dots, N \text{ and } i \neq j\}$, then

$$V_{N,Poi}^{-5/2} \Theta_{N,Poi}^{(2,3)} = V_{N,Poi}^{-3/2} \tau_{N,Poi}^{(3)} - V_{N,Poi}^{-5/2} \sum_{i=1}^N y_{N,i}^5 (1 - \pi_{N,i})^3 \pi_{N,i}^{-3},$$

where $V_{N,Poi}^{-3/2} \tau_{N,Poi}^{(3)} = O(n_0^{-1/2})$ and

$$V_{N,Poi}^{-5/2} \sum_{i=1}^N y_{N,i}^5 (1 - \pi_{N,i})^3 \pi_{N,i}^{-3} = O(n_0^{-1/2} N^{-1}).$$

Denote

$$\phi_{E,N2}(t) = \exp(-t^2/2) \left\{ 1 + 6^{-1}(it)^3 \left(V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} - 3V_{N,Poi}^{-5/2} \Theta_{N,Poi}^{(2,3)} \right) \right\},$$

then

$$\frac{1}{\pi} \int_{|t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{E,N1}(t) - \phi_{E,N2}(t)| dt = o(n_0^{-1/2}).$$

So it is sufficient to show that

$$\frac{1}{\pi} \int_{|t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N}(t) - \phi_{E,N2}(t)| dt = o(n_0^{-1/2}). \quad (\text{A.15})$$

A simple calculation yields $E(\Delta_{N,1} | \mathcal{F}_N) = E(\Delta_{N,2} | \mathcal{F}_N) = 0$, $E(\Delta_{N,1}^2 | \mathcal{F}_N) = 1$ and

$$\begin{aligned} & E(\Delta_{N,2}^2 | \mathcal{F}_N) \\ &= 4^{-1} V_{N,Poi}^{-3} \sum_{(i,j) \in \Gamma_N} y_{N,i}^2 y_{N,j}^4 (1 - \pi_{N,i})(1 - \pi_{N,j})^3 \pi_{N,i}^{-1} \pi_{N,j}^{-3} = O(n_0^{-1}) \end{aligned}$$

This implies that for $|t| \leq b_N n_0^{1/2}$ where $b_N \rightarrow 0$ as $N \rightarrow \infty$, $|t\Delta_{N,2}| = o_p(1)$. By the inequality that $|\exp(\iota x) - 1 - \iota x| \leq |x|^2/2$ for any real number x , we write

$$\begin{aligned} \phi_{W_N}(t) &= E\{\exp(\iota t W_N) | \mathcal{F}_N\} = E[\exp\{\iota t(\Delta_{N,1} + \Delta_{N,2})\} | \mathcal{F}_N] \\ &= E[\exp(\iota t \Delta_{N,1}) \{1 + \iota t \Delta_{N,2} + O_p(|t\Delta_{N,2}|^2)\} | \mathcal{F}_N] \\ &= E\{\exp(\iota t \Delta_{N,1}) | \mathcal{F}_N\} + \iota t E\{\Delta_{N,2} \exp(\iota t \Delta_{N,1}) | \mathcal{F}_N\} \\ &\quad + O(E\{|t\Delta_{N,2}|^2 | \mathcal{F}_N\}). \end{aligned}$$

According to Lemma 9.1 and Lemma 9.2,

$$\begin{aligned} & \left| \phi_{\Delta_{N,1}}(t) - \exp(-t^2/2) - 6^{-1} (\iota t)^3 V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} \exp(-t^2/2) \right| \\ & \leq C_{2,1} \exp(-19t^2/48) (t^4 n_0^{-1} + t^6 n_0^{-1}) \end{aligned}$$

and

$$\begin{aligned} & E \{ \Delta_{N,2} \exp(it\Delta_{N,1}) \mid \mathcal{F}_N \} \\ &= 2^{-1} t^2 \exp(-t^2/2) V_{N,Poi}^{-5/2} \Theta_{N,Poi}^{(2,3)} + \varpi_{N,1}(t), \end{aligned}$$

for all $|t| \leq \min \left(\left\{ \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\}^{-1/2}, V_{N,Poi}^{3/2} / \left(4\nu_{N,Poi}^{(3)} \right) \right)$, where $\varpi_{N,1}(t)$ satisfies

$$\begin{aligned} & |\varpi_{N,1}(t)| \\ & \leq C_{3,1} \exp \left\{ -t^2/2 + 2|t|^3 V_{N,Poi}^{-3/2} \nu_{N,Poi}^{(3)} / 3 + t^2 \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\} \\ & \quad \times \left(|t|^3 n_0^{-1} + t^4 n_0^{-3/2} + |t|^5 n_0^{-1} \right). \end{aligned}$$

Recall that $E(\Delta_{N,2}^2 \mid \mathcal{F}_N) = O(n_0^{-1})$, it can be easily verified that

$$\phi_{W_N}(t) = \phi_{E,N2}(t) + \varpi_N(t)$$

for $|t| \leq \min \left(b_N n_0^{1/2}, \left\{ \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\}^{-1/2}, V_{N,Poi}^{3/2} / \left(4\nu_{N,Poi}^{(3)} \right) \right)$, where

$$\begin{aligned} |\varpi_N(t)| & \leq C_{4,1} \left[t^2 n_0^{-1} + \exp(-19t^2/48) (t^4 n_0^{-1} + t^6 n_0^{-1}) \right. \\ & \quad \left. + \exp \left\{ -t^2/2 + 2|t|^3 V_{N,Poi}^{-3/2} \nu_{N,Poi}^{(3)} / 3 + t^2 \max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) \right\} \right. \\ & \quad \left. \times \left(t^4 n_0^{-1} + |t|^5 n_0^{-3/2} + t^6 n_0^{-1} \right) \right] \end{aligned}$$

and $C_{4,1}$ is a positive constant.

Under the assumption that $\lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N y_{N,i}^8 = C_3$ for a positive constant C_3 , we have $\max_{1 \leq i \leq N} |y_{N,i}| = O(N^{1/8})$. Then, $\max_{1 \leq i \leq N} E(X_{N,i}^2 \mid \mathcal{F}_N) = O(N^{-3/4})$. It follows that

for $|t| \leq n_0^{1/4}(\log n_0)^{-1}$,

$$\begin{aligned} & \frac{1}{\pi} \int_{|t| \leq n_0^{1/4}(\log n_0)^{-1}} \frac{1}{|t|} |\phi_{W_N}(t) - \phi_{E,N_2}(t)| dt \\ & \leq \frac{1}{\pi} \int_{|t| \leq n_0^{1/4}(\log n_0)^{-1}} \frac{1}{|t|} |\varpi_N(t)| dt = o(n_0^{-1/2}). \end{aligned}$$

It is obvious that

$$\frac{1}{\pi} \int_{n_0^{1/4}(\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{E,N_2}(t)| dt = o(n_0^{-1/2}),$$

it remains to establish

$$\frac{1}{\pi} \int_{n_0^{1/4}(\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N}(t)| dt = o(n_0^{-1/2}). \quad (\text{A.16})$$

Denote

$$\begin{aligned} U_{N,i,j} &= -2^{-1} V_{N,Poi}^{-3/2} \{y_{N,i}(1 - \pi_{N,i})\pi_{N,i}^{-1} + y_{N,j}(1 - \pi_{N,j})\pi_{N,j}^{-1}\} \\ &\quad \times y_{N,i}y_{N,j}\pi_{N,i}^{-1}\pi_{N,j}^{-1}(I_{N,i} - \pi_{N,i})(I_{N,j} - \pi_{N,j}), \end{aligned}$$

then $\Delta_{N,2} = \sum_{1 \leq i < j \leq N} U_{N,i,j}$ and $W_N = \sum_{i=1}^N X_{N,i} + \sum_{1 \leq i < j \leq N} U_{N,i,j}$.

Take $m = \lfloor n_0^{-1/2} N / (\log n_0) \rfloor$. According to Lemma 9.4, we assume that $m^{-1} \sum_{i=1}^m y_{N,i}^4 = O(1)$ for sufficient large N without loss of generality. Define

$$\Delta_{N,2}(m) = \sum_{i=1}^m \sum_{j=i+1}^N U_{N,i,j}.$$

By simple algebra,

$$\begin{aligned}
& E \left[\{\Delta_{N,2}(m)\}^2 \mid \mathcal{F}_N \right] \\
&= E \left\{ \left(\sum_{i=1}^m \sum_{j=i+1}^N U_{N,i,j} \right)^2 \mid \mathcal{F}_N \right\} \\
&= 4^{-1} V_{N, Poi}^{-3} \sum_{i=1}^m \sum_{j=i+1}^N \{ y_{N,i}(1 - \pi_{N,i})\pi_{N,i}^{-1} + y_{N,j}(1 - \pi_{N,j})\pi_{N,j}^{-1} \}^2 \\
&\quad \times y_{N,i}^2 y_{N,j}^2 (1 - \pi_{N,i})(1 - \pi_{N,j})\pi_{N,i}^{-1}\pi_{N,j}^{-1} \\
&= O \left(n_0^{-1} N^{-2} \sum_{i=1}^m \sum_{j=i+1}^N (y_{N,i}^4 y_{N,j}^2 + y_{N,i}^2 y_{N,j}^4) \right) = O(n_0^{-1} N^{-1} m).
\end{aligned}$$

By the inequality that $|\exp(\iota x) - 1 - \iota x| \leq 2^{-1}|x|^2$ for all real x , we have

$$\begin{aligned}
& |E\{\exp(\iota t W_N) \mid \mathcal{F}_N\} - E(\exp[\iota t \{W_N - \Delta_{N,2}(m)\}] \mid \mathcal{F}_N) \\
&\quad - \iota t E(\Delta_{N,2}(m) \exp[\iota t \{W_N - \Delta_{N,2}(m)\}] \mid \mathcal{F}_N)| \\
&\leq E(|\exp[\iota t \{W_N - \Delta_{N,2}(m)\}]| |\exp\{\iota t \Delta_{N,2}(m)\} - 1 - \iota t \Delta_{N,2}(m)| \mid \mathcal{F}_N) \\
&\leq 2^{-1} E\{|\iota t \Delta_{N,2}(m)|^2 \mid \mathcal{F}_N\} \\
&\leq C_{4,2} t^2 n_0^{-1} N^{-1} m,
\end{aligned}$$

where $C_{4,2}$ is a positive constant. This clearly indicates that

$$\begin{aligned}
|E\{\exp(\iota t W_N) \mid \mathcal{F}_N\}| &\leq |E(\exp[\iota t \{W_N - \Delta_{N,2}(m)\}] \mid \mathcal{F}_N)| \\
&\quad + |t| |E(\Delta_{N,2}(m) \exp[\iota t \{W_N - \Delta_{N,2}(m)\}] \mid \mathcal{F}_N)| \\
&\quad + C_{4,2} t^2 n_0^{-1} N^{-1} m.
\end{aligned}$$

In view of the fact that $X_{N,1}, \dots, X_{N,m}$ are the only terms in $W_N - \Delta_{N,2}(m)$ that depend on

$I_{N,1}, \dots, I_{N,m}$, for a positive constant $C_{4,3}$,

$$\begin{aligned}
& |E(\exp[it\{W_N - \Delta_{N,2}(m)\}] | \mathcal{F}_N)| \\
&= \left| E \left(\exp \left[it \left\{ \sum_{i=1}^m X_{N,i} + \sum_{i=m+1}^N X_{N,i} + \Delta_{N,2} - \Delta_{N,2}(m) \right\} \right] | \mathcal{F}_N \right) \right| \\
&\leq \left| \prod_{i=1}^m E\{\exp(itX_{N,i}) | \mathcal{F}_N\} \right| \left| E \left(\exp \left[it \left\{ \sum_{i=m+1}^N X_{N,i} + \Delta_{N,2} - \Delta_{N,2}(m) \right\} \right] | \mathcal{F}_N \right) \right| \\
&\leq \left| \prod_{i=1}^m E\{\exp(itX_{N,i}) | \mathcal{F}_N\} \right| \leq C_{4,3}m^{-a}
\end{aligned}$$

by condition (C4). In addition, there exists a constant $C_{4,4} > 0$ such that

$$\begin{aligned}
& |E(\Delta_{N,2}(m) \exp[it\{W_N - \Delta_{N,2}(m)\}] | \mathcal{F}_N)| \\
&= \left| E \left(\sum_{i=1}^m \sum_{j=i+1}^N U_{N,i,j} \exp \left[it \left\{ \sum_{\ell=1}^N X_{N,\ell} + \Delta_{N,2} - \Delta_{N,2}(m) \right\} \right] | \mathcal{F}_N \right) \right| \\
&\leq \sum_{i=1}^m \sum_{j=i+1}^N \left| \prod_{\substack{\ell=1, \dots, m \\ \ell \neq i, j}} E\{\exp(itX_{N,\ell}) | \mathcal{F}_N\} \right| \\
&\quad \times \left| E(U_{N,i,j} \exp[it\{X_{N,i} + X_{N,j} + \Delta_{N,2} - \Delta_{N,2}(m)\}] | \mathcal{F}_N) \right| \\
&\leq C_{4,4}(m-2)^{-a}n_0^{1/2}N^{-1}m.
\end{aligned}$$

Finally, for a positive constant $C_{4,5}$,

$$\begin{aligned}
& \frac{1}{\pi} \int_{n_0^{1/4}(\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N}(t)| dt \\
&\leq \frac{C_{4,5}}{\pi} \int_{n_0^{1/4}(\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} (m^{-a} + |t|m^{1-a}n_0^{1/2}N^{-1} + t^2n_0^{-1}N^{-1}m) dt \\
&= o(n_0^{-1/2}),
\end{aligned}$$

as $a > 2$. Therefore, we finish the proof of this lemma. \square

Proof of Theorem 3.1. According to Lemma 3.1, $\mu_{N,Poi}^{(3)} = O(n_0^{-2}N^3)$. This cooperates with

(C2) indicates that

$$\frac{\mu_{N,Poi}^{(3)}}{V_{N,Poi}^{3/2}} = O(n_0^{-1/2}). \quad (\text{A.17})$$

In addition, by (1) and (2) of Lemma 3.1, we can prove that

$$\hat{V}_{N,Poi}^{-3/2} \hat{\mu}_{N,Poi}^{(3)} = O_p(n_0^{-1/2}).$$

Similarly, we can show that $\hat{V}_{N,Poi}^{-3/2} \hat{\tau}_{N,Poi}^{(3)} = O_p(n_0^{-1/2})$ according to Lemma 3.1. Finally, by (C2) and Lemma 3.1,

$$\hat{V}_{N,Poi}^{-3/2} \hat{\mu}_{N,Poi}^{(3)} - V_{N,Poi}^{-3/2} \mu_{N,Poi}^{(3)} = o_p(n_0^{-1/2})$$

and

$$\hat{V}_{N,Poi}^{-3/2} \hat{\tau}_{N,Poi}^{(3)} - V_{N,Poi}^{-3/2} \tau_{N,Poi}^{(3)} = o_p(n_0^{-1/2}).$$

Combine these results with Lemma 9.5, we have proved Theorem 3.1. □

The proof of Theorem 3.2 uses the following lemma.

Lemma 9.6. *Let (N_1^*, \dots, N_n^*) be a multinomial random vector with distribution $\text{MN}(N; \rho)$, where $\rho = (\rho_1, \dots, \rho_n)$ and $\rho_i = \pi_{N,i}^{-1} / \left(\sum_{j=1}^n \pi_{N,j}^{-1} \right)$ for $i = 1, \dots, n$. Denote $\mathcal{F}_N^* = \{y_{N,1}^*, \dots, y_{N,n}^*\}$ to be the bootstrap finite population generated from the realized sample $\{y_{N,1}, \dots, y_{N,n}\}$ and the random vector (N_1^*, \dots, N_n^*) with each N_i^* indicates the number of replicates of $y_{N,i}$ in \mathcal{F}_N^* . Define $n_1 = \sum_{i=1}^n \mathbb{I}(N_i^* \geq 1)$ as the number of distinct $y_{N,i}$, $i = 1, \dots, n$ in \mathcal{F}_N^* . Then, as $N \rightarrow \infty$,*

$$\mathbb{P}_*(n_1 \geq \min\{n_0, m\}) \rightarrow 1, \quad (\text{A.18})$$

where $m = \lfloor n_0^{-1/2} N / (\log n_0) \rfloor$ is the integer part of $n_0^{-1/2} N / (\log n_0)$ and \mathbb{P}_* is the probability measure for the first step of the proposed bootstrap method conditional on the realized sample $\{y_{N,1}, \dots, y_{N,n}\}$.

Proof of Lemma 9.6. We prove this lemma under two different case scenarios: $m \geq n_0$ and $m < n_0$.

First, consider the case of $m \geq n_0$, so we have that $n_0 = o(N^{2/3})$. By the strong law of large numbers, $nn_0^{-1} = \sum_{i=1}^N I_{N,i} (\sum_{i=1}^N \pi_{N,i})^{-1} \rightarrow 1$ with probability 1. It suffices to show that

$$\mathbb{P}_* (n_1 \neq n) \rightarrow 0. \quad (\text{A.19})$$

Mentioned that

$$\begin{aligned} \mathbb{P}_* (n_1 \neq n) &= \mathbb{P}_* \left(\sum_{i=1}^n \mathbb{I}(N_i^* \geq 1) < n \right) \\ &= \mathbb{P}_* \left(\bigcup_{i=1}^n \{N_i^* = 0\} \right) \\ &\leq \sum_{i=1}^n \mathbb{P}_* (N_i^* = 0) \\ &\leq \sum_{i=1}^n (1 - \rho_i)^N \\ &\leq n(1 - C_1 C_2^{-1} n^{-1})^N \\ &\rightarrow 0, \end{aligned}$$

where the last inequality is due to that $C_1 C_2^{-1} n^{-1} \leq \rho_i \leq C_1^{-1} C_2 n^{-1}$ for $i = 1, \dots, n$.

Next, consider the case of $m < n_0$ and it is sufficient to prove that

$$\mathbb{P}_* (n_1 \geq m) \rightarrow 1. \quad (\text{A.20})$$

When $\alpha < 1$, $n_0 = o(N)$ and $m = o(N^{2/3})$. Thus,

$$\begin{aligned}
\mathbb{P}_*(n_1 < m) &= \mathbb{P}_* \left(\sum_{i=1}^n \mathbb{I}(N_i^* \geq 1) < m \right) \\
&\leq \mathbb{P}_* \left(\sum_{i=1}^m \mathbb{I}(N_i^* \geq 1) < m \right) \\
&= \mathbb{P}_* \left(\bigcup_{i=1}^m \{N_i^* = 0\} \right) \\
&\leq \sum_{i=1}^m \mathbb{P}_*(N_i^* = 0) \\
&\leq \sum_{i=1}^m (1 - \rho_i)^N \\
&\leq m(1 - C_1 C_2^{-1} n^{-1})^N \\
&\rightarrow 0,
\end{aligned}$$

which implies (A.20) immediately.

When $\alpha = 1$, $n_0 = O(N)$ and $m = o(N^{1/2})$. Using Stirling's formula, there exists a positive constant $C_{5,1}$ such that

$$\begin{aligned}
&\mathbb{P}_*(n_1 < m) \\
&\leq \sum_{i=n-m}^n \mathbb{P}_*(n_1 = n - i) \\
&\leq \sum_{i=n-m}^n \frac{n!}{(n-i)!i!} \{(n-i)C_1^{-1}C_2 n^{-1}\}^N \\
&\leq C_{5,1} \sum_{i=n-m}^n \left\{ \frac{n}{2\pi i(n-i)} \right\}^{1/2} \left(\frac{n-i}{i} \right)^i \left(\frac{n}{n-i} \right)^n \left\{ \frac{C_2(n-i)}{C_1 n} \right\}^n,
\end{aligned}$$

where the second inequality uses the fact that $C_1 C_2^{-1} n^{-1} \leq \rho_i \leq C_1^{-1} C_2 n^{-1}$. Mentioned that

$m = o(N^{1/2})$ and $n = O(N)$, we have

$$\begin{aligned} \left\{ \frac{n}{2\pi i(n-i)} \right\}^{1/2} &= O(1), \\ \left(\frac{n-i}{i} \right)^i &\leq \left(\frac{m}{n-m} \right)^i = o(N^{-i/2}) \\ \left(\frac{n}{n-i} \right)^n \left\{ \frac{C_2(n-i)}{C_1 n} \right\}^n &= \left(\frac{C_2}{C_1} \right)^n. \end{aligned}$$

for $i = n - m, \dots, n$. Thus,

$$\begin{aligned} &\left\{ \frac{n}{2\pi i(n-i)} \right\}^{1/2} \left(\frac{n-i}{i} \right)^i \left(\frac{n}{n-i} \right)^n \left\{ \frac{C_2(n-i)}{C_1 n} \right\}^n \\ &= o(N^{-i/2} (C_1^{-1} C_2)^n). \end{aligned}$$

Finally,

$$\begin{aligned} &\mathbb{P}_*(n_1 < m) \\ &= o\left(\sum_{i=n-m}^n N^{-i/2} (C_1^{-1} C_2)^n \right) \\ &= o(m N^{-(n-m)/2} (C_1^{-1} C_2)^n) \\ &= o(\exp\{- (n-m) \log N/2 + n \log(C_1^{-1} C_2) + \log m\}) \\ &\rightarrow 0, \end{aligned}$$

which finalize the proof of this lemma. □

Proof of Theorem 3.2. We first show

$$N^{-1} \sum_{i=1}^N (\pi_{N,i}^{-1} I_{N,i} - 1) \rightarrow 0 \tag{A.21}$$

almost surely (\mathbb{P}_{Poi}). Denote $D_N^{(2)}$ to be the event $\{N^{-1} |\sum_{i=1}^N (\pi_{N,i}^{-1} I_{N,i} - 1)| > \epsilon\}$, where ϵ is

a fixed positive number. Similar to the argument used in proving Lemma 3.1, we have

$$\begin{aligned}
& \mathbb{P}_{Poi} \left(D_N^{(2)} \right) \\
& \leq \epsilon^{-4} N^{-4} E \left\{ \left| \sum_{i=1}^N (\pi_{N,i}^{-1} I_{N,i} - 1) \right|^4 \mid \mathcal{F}_N \right\} \\
& = \epsilon^{-4} N^{-4} \left[\sum_{i=1}^N (1 - \pi_{N,i}) \{ (1 - \pi_{N,i})^3 \pi_{N,i}^{-3} + 1 \} \right. \\
& \quad \left. + \sum_{(i,j) \in \Gamma_N} (1 - \pi_{N,i})(1 - \pi_{N,j}) \pi_{N,i}^{-1} \pi_{N,j}^{-1} \right] \\
& \leq C_{6,1} \epsilon^{-4} (n_0^{-3} + n_0^{-2}),
\end{aligned}$$

where $C_{6,1}$ is a positive constant with respect to N and $\Gamma_N = \{(i, j) : i, j = 1, \dots, N \text{ and } i \neq j\}$. This immediately implies

$$\sum_{N=1}^{\infty} \mathbb{P}_{Poi}(D_N^{(2)}) < \infty,$$

for any arbitrary positive ϵ and we have proved (A.21) by the Borel-Cantelli Lemma.

Next, for any $0 \leq \delta \leq 8$,

$$\begin{aligned}
E \left(n_0^{-1} \sum_{i=1}^N |y_{N,i}|^\delta I_{N,i} \mid \mathcal{F}_N \right) & = n_0^{-1} \sum_{i=1}^N |y_{N,i}|^\delta \pi_{N,i} \\
& \leq C_2 N^{-1} \sum_{i=1}^N |y_{N,i}|^\delta \\
& < \infty,
\end{aligned} \tag{A.22}$$

where the first inequality holds by (C1) and the last inequality holds by (C3). Thus, by (A.22) and the Markov's inequality, we have

$$n_0^{-1} \sum_{i=1}^N y_{N,i}^\delta I_{N,i} = O_p(1) \tag{A.23}$$

for $0 \leq \delta \leq 8$. In addition, as

$$\begin{aligned} E \left(n_0^{-1} \sum_{i=1}^N I_{N,i} \mid \mathcal{F}_N \right) &= 1, \\ \text{var} \left(n_0^{-1} \sum_{i=1}^N I_{N,i} \mid \mathcal{F}_N \right) &\leq n_0^{-2} \sum_{i=1}^N \pi_{N,i} \\ &\leq C_2 n_0^{-1}, \end{aligned}$$

we have

$$n_0^{-1} \sum_{i=1}^N I_{N,i} = 1 + o_p(1). \quad (\text{A.24})$$

In the first step of our proposed bootstrap method, $y_{N,1}^*, \dots, y_{N,N}^*$ are independently and identically distributed (i.i.d.) with $\mathbb{P}_*(y_{N,i}^* = y_{N,j}) = \rho_{N,j} = \pi_{N,j}^{-1} \left(\sum_{\ell=1}^n \pi_{N,\ell}^{-1} \right)^{-1}$. Mentioned that, for a positive constant $C_{6,2}$,

$$\begin{aligned} &E_* \left\{ N^{-1} \sum_{i=1}^N (y_{N,i}^*)^8 \right\} \\ &= \sum_{i=1}^n \pi_{N,i}^{-1} \left(\sum_{j=1}^n \pi_{N,j}^{-1} \right)^{-1} y_{N,i}^8 \\ &= \left(N^{-1} \sum_{i=1}^N \pi_{N,i}^{-1} I_{N,i} \right)^{-1} \sum_{i=1}^N (N^{-1} \pi_{N,i}^{-1}) y_{N,i}^8 I_{N,i} \\ &\leq C_{6,2} n_0^{-1} \sum_{i=1}^N y_{N,i}^8 I_{N,i} \\ &= O_p(1), \end{aligned} \quad (\text{A.25})$$

where the first equality holds by the property of the proposed bootstrap method, the inequality holds by (A.21) and (C1), and the last equality holds by (A.23). Thus, by (A.25) and Markov's inequality, we have

$$N^{-1} \sum_{i=1}^N (y_{N,i}^*)^8 = O_p(1). \quad (\text{A.26})$$

Similarly, we can prove that for any subset of the bootstrap finite population, say, $\{y_{N,\ell_1}^*, \dots, y_{N,\ell_{m_0}}^*\} \subset \{y_{N,1}^*, \dots, y_{N,N}^*\}$ and all $0 \leq \delta \leq 8$,

$$m_0^{-1} \sum_{i=1}^{m_0} (y_{N,\ell_i}^*)^\delta = O_p(1), \quad (\text{A.27})$$

here m_0 can be any positive integer less than N .

Denote $V_{N,Poi}^* = \sum_{i=1}^N E_{**}(\hat{Y}_{N,Poi}^* - Y_N^*)^2 = \sum_{i=1}^N (y_{N,i}^*)^2 (1 - \pi_{N,i}^*)(\pi_{N,i}^*)^{-1}$, then by Lemma 3.1, (A.21) and Condition (C2),

$$\begin{aligned} & E_* (n_0 N^{-2} V_{N,Poi}^*) \\ &= E_* \left\{ n_0 N^{-2} \sum_{i=1}^n N_i^* y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-1} \right\} \\ &= n_0 N^{-2} \sum_{i=1}^n N \pi_{N,i}^{-1} \left(\sum_{j=1}^n \pi_{N,j}^{-1} \right)^{-1} y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-1} \\ &= \left(N^{-1} \sum_{i=1}^N \pi_{N,i}^{-1} I_{N,i} \right)^{-1} \left(n_0 N^{-2} \sum_{i=1}^n y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} \right) \\ &= \left(N^{-1} \sum_{i=1}^N \pi_{N,i}^{-1} I_{N,i} \right)^{-1} \left(n_0 N^{-2} \hat{V}_{N,Poi} \right) \\ &\rightarrow \sigma_1^2 \end{aligned}$$

in probability. In addition, for a positive constant $C_{6,3}$,

$$\begin{aligned}
& \text{var}_* (n_0 N^{-2} V_{N,Poi}^*) \\
&= n_0^2 N^{-4} \sum_{i=1}^n \text{var}_*(N_i^*) y_{N,i}^4 (1 - \pi_{N,i})^2 \pi_{N,i}^{-2} \\
&\quad + n_0^2 N^{-4} \sum_{(i,j) \in \Gamma_n} \text{cov}_*(N_i^*, N_j^*) y_{N,i}^2 y_{N,j}^2 (1 - \pi_{N,i})(1 - \pi_{N,j}) \pi_{N,i}^{-1} \pi_{N,j}^{-1} \\
&= n_0^2 N^{-4} \sum_{i=1}^n N \rho_{N,i} (1 - \rho_{N,i}) y_{N,i}^4 (1 - \pi_{N,i})^2 \pi_{N,i}^{-2} \\
&\quad - n_0^2 N^{-4} \sum_{(i,j) \in \Gamma_n} N \rho_{N,i} \rho_{N,j} y_{N,i}^2 y_{N,j}^2 (1 - \pi_{N,i})(1 - \pi_{N,j}) \pi_{N,i}^{-1} \pi_{N,j}^{-1} \\
&\leq C_{6,3} \left\{ N^{-1} n_0^{-1} \sum_{i=1}^N y_{N,i}^4 I_{N,i} + N^{-1} \left(n_0^{-1} \sum_{i=1}^N y_{N,i}^2 I_{N,i} \right) \left(n_0^{-1} \sum_{i=1}^N y_{N,i}^2 I_{N,i} \right) \right\} \\
&= o_p(N^{-1}),
\end{aligned}$$

where $\Gamma_n = \{(i, j) : i, j = 1, \dots, n \text{ and } i \neq j\}$. Thus, we have

$$n_0 N^{-2} V_{N,Poi}^* = \sigma_1^2 + o_p(1). \quad (\text{A.28})$$

Denote $T_{N,Poi}^* = (\hat{V}_{N,Poi}^*)^{-1/2} (\hat{Y}_{N,Poi}^* - Y_N^*)$, where $Y_N^* = \sum_{i=1}^N y_{N,i}^*$,

$$\begin{aligned}
\hat{Y}_{N,Poi}^* &= \sum_{i=1}^n m_i^* y_{N,i} \pi_{N,i}^{-1} \\
&= \sum_{i=1}^N y_{N,i}^* (\pi_{N,i}^*)^{-1} I_{N,i}^*
\end{aligned}$$

and

$$\begin{aligned}
\hat{V}_{N,Poi}^* &= \sum_{i=1}^n m_i^* y_{N,i}^2 (1 - \pi_{N,i}) \pi_{N,i}^{-2} \\
&= \sum_{i=1}^N (y_{N,i}^*)^2 (1 - \pi_{N,i}^*) (\pi_{N,i}^*)^{-2} I_{N,i}^*,
\end{aligned}$$

here $I_{N,i}^*$ is the bootstrap counterpart of $I_{N,i}$ and $I_{N,i}^* \sim \text{Ber}(\pi_{N,i}^*)$ conditional on the bootstrap

finite population \mathcal{F}_N^* .

In view of (A.26) and (A.28), by similar arguments used in the proof of Lemma 9.3, we can show that

$$T_{N,Poi}^* = \Delta_{N,1}^* + \Delta_{N,2}^* + \Delta_{N,3}^* + \Delta_{N,4}^*,$$

where

$$\begin{aligned} \Delta_{N,1}^* &= (V_{N,Poi}^*)^{-1/2} \sum_{i=1}^N y_{N,i}^* (\pi_{N,i}^*)^{-1} (I_{N,i}^* - \pi_{N,i}^*), \\ \Delta_{N,2}^* &= -2^{-1} (V_{N,Poi}^*)^{-3/2} \sum_{(i,j) \in \Gamma_N} y_{N,i}^* (y_{N,j}^*)^2 (1 - \pi_{N,j}^*) (\pi_{N,i}^*)^{-1} (\pi_{N,j}^*)^{-2} \\ &\quad \times (I_{N,i}^* - \pi_{N,i}^*) (I_{N,j}^* - \pi_{N,j}^*), \\ \Delta_{N,3}^* &= -2^{-1} (V_{N,Poi}^*)^{-3/2} \sum_{i=1}^N (y_{N,i}^*)^3 (1 - \pi_{N,i}^*)^2 (\pi_{N,i}^*)^{-2}, \end{aligned}$$

and $\Delta_{N,4}^*$ satisfies

$$\mathbb{P}_{Poi}^* \left(|\Delta_{N,4}^*| \geq n_0^{-1/2} (\log n_0)^{-1} \right) = o_p(n_0^{-1/2}),$$

where \mathbb{P}_{Poi}^* is the counterpart of \mathbb{P}_{Poi} conditional on the bootstrap finite population $\{y_{N,1}^*, \dots, y_{N,N}^*\}$.

Let $\hat{F}_{N,Poi}^*(z) = \mathbb{P}_{Poi}^*(T_{N,Poi}^* \leq z)$ be the cumulative distribution function of $T_{N,Poi}^*$ conditional on the bootstrap finite population \mathcal{F}_N^* , we proceed to prove

$$\begin{aligned} &\hat{F}_{N,Poi}^*(z) \tag{A.29} \\ &= \Phi(z) + \left\{ \frac{\mu_{N,Poi}^{(3)*}}{6(V_{N,Poi}^*)^{3/2}} (1 - z^2) + \frac{\tau_{N,Poi}^{(3)*}}{2(V_{N,Poi}^*)^{3/2}} z^2 \right\} \phi(z) + o_p(n_0^{-1/2}), \end{aligned}$$

uniformly in $z \in \mathbb{R}$, where

$$\mu_{N,Poi}^{(3)*} = \sum_{i=1}^N E_{**}(\hat{Y}_{N,Poi}^* - Y_N^*)^3 = \sum_{i=1}^N (y_{N,i}^*)^3 (1 - \pi_{N,i}^*) \{ (1 - \pi_{N,i}^*)^2 (\pi_{N,i}^*)^{-2} - 1 \}$$

and

$$\tau_{N, Poi}^{(3)*} = \sum_{i=1}^N (y_{N,i}^*)^3 (1 - \pi_{N,i}^*) (\pi_{N,i}^*)^{-2}.$$

Denote $W_N^* = \Delta_{N,1}^* + \Delta_{N,2}^*$ and let $\phi_{W_N^*}(t)$ be the characteristic function (c.f.) of W_N^* conditional on the bootstrap finite population \mathcal{F}_N^* , that is,

$$\phi_{W_N^*}(t) = E_{**} \{ \exp(itW_N^*) \} = E_{**} [\exp\{it(\Delta_{N,1}^* + \Delta_{N,2}^*)\}].$$

Regarding the proof of Lemma 9.5, it is enough to show that

$$\frac{1}{\pi} \int_{n_0^{1/4} (\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N^*}(t)| dt = o_p(n_0^{-1/2}), \quad (\text{A.30})$$

for any arbitrary $\varepsilon \in (0, 1)$ in order to finalize (A.29). Here a_ε is chosen to satisfy $|dF_{E,N1}(z)/dx| \leq a_\varepsilon \varepsilon$ with $F_{E,N1}(z)$ defined as

$$F_{E,N1}(z) = \Phi(z) + 6^{-1} \left(V_{N, Poi}^{-3/2} \mu_{N, Poi}^{(3)} - 3V_{N, Poi}^{-3/2} \tau_{N, Poi}^{(3)} \right) (1 - z^2) \phi(z).$$

Denote $X_{N,i}^* = (V_{N, Poi}^*)^{-1/2} y_{N,i}^* (\pi_{N,i}^*)^{-1} (I_{N,i}^* - \pi_{N,i}^*)$ and

$$\begin{aligned} U_{N,i,j}^* &= -2^{-1} V_{N, Poi}^{-3/2} \{ y_{N,i}^* (1 - \pi_{N,i}^*) (\pi_{N,i}^*)^{-1} + y_{N,j}^* (1 - \pi_{N,j}^*) (\pi_{N,j}^*)^{-1} \} \\ &\quad \times y_{N,i}^* y_{N,j}^* (\pi_{N,i}^*)^{-1} (\pi_{N,j}^*)^{-1} (I_{N,i}^* - \pi_{N,i}^*) (I_{N,j}^* - \pi_{N,j}^*), \end{aligned}$$

then $\Delta_{N,1}^* = \sum_{i=1}^N X_{N,i}^*$, $\Delta_{N,2}^* = \sum_{1 \leq i < j \leq N} U_{N,i,j}^*$ and

$$W_N^* = \sum_{i=1}^N X_{N,i}^* + \sum_{1 \leq i < j \leq N} U_{N,i,j}^*.$$

Take $m = \lfloor n_0^{-1/2} N / (\log n_0) \rfloor$, we prove (A.30) under two different case scenarios: $m \geq n_0$ and $m < n_0$.

First, consider the case of $m \geq n_0$, we have that $n_0 = o(N^{2/3})$. According to Lemma

9.6, we assume $y_{N,1}^* = y_{N,1}, \dots, y_{N,n}^* = y_{N,n}$, without loss of generality. Then, by (A.28), condition (C5) and the fact that

$$\begin{aligned} & E \{ \exp (t X_{N,i}) \mid \mathcal{F}_N \} \\ &= \pi_{N,i} \exp \left\{ t V_{N,Poi}^{-1/2} y_{N,i} (1 - \pi_{N,i}) \pi_{N,i}^{-1} \right\} + (1 - \pi_{N,i}) \exp \left(- t V_{N,Poi}^{-1/2} y_{N,i} \right), \end{aligned}$$

we arrive at

$$\begin{aligned} & \left| \prod_{i=1}^m E_{**} \{ \exp (t X_{N,i}^*) \} \right| \\ & \leq \left| \prod_{i=1}^n E_{**} \{ \exp (t X_{N,i}^*) \} \right| \\ &= \left| \prod_{i=1}^n \left[\pi_{N,i}^* \exp \left\{ t (V_{N,Poi}^*)^{-1/2} y_{N,i}^* (1 - \pi_{N,i}^*) (\pi_{N,i}^*)^{-1} \right\} \right. \right. \\ & \quad \left. \left. + (1 - \pi_{N,i}^*) \exp \left\{ - t (V_{N,Poi}^*)^{-1/2} y_{N,i}^* \right\} \right] \right| \\ &= \left| \prod_{i=1}^n \left[\pi_{N,i} \exp \left\{ t (V_{N,Poi}^*)^{-1/2} y_{N,i} (1 - \pi_{N,i}) \pi_{N,i}^{-1} \right\} \right. \right. \\ & \quad \left. \left. + (1 - \pi_{N,i}) \exp \left\{ - t (V_{N,Poi}^*)^{-1/2} y_{N,i} \right\} \right] \right| \\ &= O_p(n_0^{-a}). \end{aligned} \tag{A.31}$$

By (A.27), we have that $m^{-1} \sum_{i=1}^m (y_{N,i}^*)^4 = O_p(1)$. Similar to the proof of Lemma 9.3, define

$$\Delta_{N,2}^*(m) = \sum_{i=1}^m \sum_{j=i+1}^N U_{N,i,j}^*,$$

then $E_{**}\{\Delta_{N,2}^*(m)\}^2 = O_p(n_0^{-1}N^{-1}m)$. Furthermore, for positive constants $C_{6,4}$ and $C_{6,5}$,

$$\begin{aligned}
& |\phi_{W_N^*}(t)| \\
& \leq |E_{**}(\exp[lt\{W_N^* - \Delta_{N,2}^*(m)\}])| \\
& \quad + |t| |E_{**}(\Delta_{N,2}^*(m) \exp[lt\{W_N^* - \Delta_{N,2}^*(m)\}])| \\
& \quad + C_{6,4}t^2n_0^{-1}N^{-1}m \\
& \leq C_{6,5} \left(n_0^{-a} + |t|n_0^{3/2-a}N^{-1} + t^2n_0^{-1}N^{-1}m \right)
\end{aligned}$$

in probability and this immediately implies

$$\begin{aligned}
& \frac{1}{\pi} \int_{n_0^{1/4}(\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} |\phi_{W_N^*}^*(t)| dt \\
& \leq \frac{C_{6,5}}{\pi} \int_{n_0^{1/4}(\log n_0)^{-1} \leq |t| \leq a_\varepsilon n_0^{1/2}} \frac{1}{|t|} (n_0^{-a} + |t|n_0^{3/2-a}N^{-1} + t^2n_0^{-1}N^{-1}m) dt \\
& = o_p(n_0^{-1/2}).
\end{aligned}$$

Next, consider the case of $m < n_0$. According to Lemma 9.6, we assume $y_{N,1}^* = y_{N,1}, \dots, y_{N,m}^* = y_{N,m}$, without loss of generality. Then, use the same technique as (A.31), we can show that

$$\left| \prod_{i=1}^m E_{**}\{\exp(tX_{N,i}^*)\} \right| = O_p(m^{-a}).$$

The following procedure is similar to the case of $m \geq n_0$. Thus, we arrive at

$$\begin{aligned}
& \hat{F}_{N,Poi}^*(z) \\
& = \Phi(z) + \left\{ \frac{\mu_{N,Poi}^{(3)*}}{6(V_{N,Poi}^*)^{3/2}}(1 - z^2) + \frac{\tau_{N,Poi}^{(3)*}}{2(V_{N,Poi}^*)^{3/2}}z^2 \right\} \phi(z) + o_p(n_0^{-1/2}).
\end{aligned}$$

It remains to show that

$$\hat{V}_{N,Poi}^{-3/2} \hat{\mu}_{N,Poi}^{(3)} - (V_{N,Poi}^*)^{-3/2} \mu_{N,Poi}^{(3)*} = o_p(n_0^{-1/2})$$

and

$$\hat{V}_{N,Poi}^{-3/2} \hat{\tau}_{N,Poi}^{(3)} - (V_{N,Poi}^*)^{-3/2} \tau_{N,Poi}^{(3)*} = o_p(n_0^{-1/2}).$$

Mentioned that

$$\begin{aligned} & \hat{V}_{N,Poi}^{-3/2} \hat{\mu}_{N,Poi}^{(3)} - (V_{N,Poi}^*)^{-3/2} \mu_{N,Poi}^{(3)*} \\ = & n_0^{-1/2} \left\{ \left(n_0 N^{-2} \hat{V}_{N,Poi} \right)^{-3/2} \left(n_0^2 N^{-3} \hat{\mu}_{N,Poi}^{(3)} \right) \right. \\ & \left. - \left(n_0 N^{-2} V_{N,Poi}^* \right)^{-3/2} \left(n_0^2 N^{-3} \mu_{N,Poi}^{(3)*} \right) \right\} \end{aligned}$$

and

$$\begin{aligned} & \hat{V}_{N,Poi}^{-3/2} \hat{\tau}_{N,Poi}^{(3)} - (V_{N,Poi}^*)^{-3/2} \tau_{N,Poi}^{(3)*} \\ = & n_0^{-1/2} \left\{ \left(n_0 N^{-2} \hat{V}_{N,Poi} \right)^{-3/2} \left(n_0^2 N^{-3} \hat{\tau}_{N,Poi}^{(3)} \right) \right. \\ & \left. - \left(n_0 N^{-2} V_{N,Poi}^* \right)^{-3/2} \left(n_0^2 N^{-3} \tau_{N,Poi}^{(3)*} \right) \right\}, \end{aligned}$$

it suffices to prove that

$$n_0 N^{-2} (V_{N,Poi}^* - \hat{V}_{N,Poi}) \rightarrow 0, \quad (\text{A.32})$$

$$n_0^2 N^{-3} (\mu_{N,Poi}^{(3)*} - \hat{\mu}_{N,Poi}^{(3)}) \rightarrow 0, \quad (\text{A.33})$$

$$n_0^2 N^{-3} (\tau_{N,Poi}^{(3)*} - \hat{\tau}_{N,Poi}^{(3)}) \rightarrow 0, \quad (\text{A.34})$$

in probability conditional on the series of realized samples. For (A.32), it is a consequence

result of $n_0 N^{-2} V_{N, Poi}^* = \sigma_1^2 + o_p(1)$ and Lemma 3.1. For (A.33), consider

$$\begin{aligned}
& E_*(n_0^2 N^{-3} \mu_{N, Poi}^{(3)*}) \\
&= n_0^2 N^{-3} \sum_{i=1}^n \frac{N \pi_{N,i}^{-1}}{\sum_{j=1}^n \pi_{N,j}^{-1}} y_{N,i}^3 (1 - \pi_{N,i}) \{(1 - \pi_{N,i})^2 \pi_{N,i}^{-2} - 1\} \\
&= n_0^2 N^{-3} \hat{\mu}_{N, Poi}^{(3)} \{1 + o(1)\}
\end{aligned} \tag{A.35}$$

almost surely, where the second equality holds by (A.21). Next, consider

$$\begin{aligned}
& \text{var}_*(n_0^2 N^{-3} \mu_{N, Poi}^{(3)*}) \\
&\leq 4n_0^4 N^{-6} \sum_{i=1}^n \frac{N \pi_{N,i}^{-1}}{\sum_{j=1}^n \pi_{N,j}^{-1}} y_{N,i}^6 \pi_{N,i}^{-4} \\
&\leq 4C_1^{-5} n_0^{-1} N^{-1} \sum_{i=1}^N y_{N,i}^6 \{1 + o(1)\} = o(1)
\end{aligned} \tag{A.36}$$

almost surely, where the second inequality holds by (A.21) and the last equality holds by (C3). By (2) in Lemma 3.1, (A.35) and (A.36), we have proved (A.33), and the proof of (A.34) is similar. This concludes the proof of Theorem 3.2. \square

Lemma 9.7. *Let i, j, k, l be pairwise distinct positive integers, which are no larger than N . Suppose that (C6) holds. Under SRS, we have*

$$E[(I_{N,i} \pi_{N,i}^{-1} - 1)^4 \mid \mathcal{F}_N] = O(n^{-3} N^3), \tag{A.37}$$

$$E[(I_{N,i} \pi_{N,i}^{-1} - 1)^3 (I_{N,j} \pi_{N,j}^{-1} - 1) \mid \mathcal{F}_N] = O(N^2 n^{-2}), \tag{A.38}$$

$$E[(I_{N,i} \pi_{N,i}^{-1} - 1)^2 (I_{N,j} \pi_{N,j}^{-1} - 1)^2 \mid \mathcal{F}_N] = O(N^2 n^{-2}), \tag{A.39}$$

$$\begin{aligned}
& E[(I_{N,i} \pi_{N,i}^{-1} - 1)(I_{N,j} \pi_{N,j}^{-1} - 1)(I_{N,k} \pi_{N,k}^{-1} - 1)^2 \mid \mathcal{F}_N] \\
&= O(N n^{-2}),
\end{aligned} \tag{A.40}$$

$$\begin{aligned}
& E[(I_{N,i} \pi_{N,i}^{-1} - 1)(I_{N,j} \pi_{N,j}^{-1} - 1)(I_{N,k} \pi_{N,k}^{-1} - 1)(I_{N,l} \pi_{N,l}^{-1} - 1) \mid \mathcal{F}_N] \\
&= O(n^{-2}).
\end{aligned} \tag{A.41}$$

Proof of Lemma 9.7. Consider

$$\begin{aligned}
E[(I_{N,i}\pi_{N,i}^{-1} - 1)^4 \mid \mathcal{F}_N] &= \pi_{N,i}(\pi_{N,i}^{-1} - 1)^4 + (1 - \pi_{N,i}) \\
&= (1 - \pi_{N,i})[(1 - \pi_{N,i})^3\pi_{N,i}^{-3} + 1] \\
&\leq N^3n^{-3},
\end{aligned}$$

where the last inequality holds by the fact that $(1 - x)^3 + x^3 \leq 1$ for $x \in [0, 1]$. Thus, we have proved (A.37).

Denote $\#A$ to be the number of elements that equal to 1 in set A . Under SRS, we have

$$\begin{aligned}
P_N(\#\{I_{N,i}, I_{N,j}\} = 2) &= \frac{n(n-1)}{N(N-1)}, \\
P_N(\{I_{N,i} = 1, I_{N,j} = 0\}) &= \frac{n(N-n)}{N(N-1)}, \\
P_N(\#\{I_{N,i}, I_{N,j}\} = 0) &= \frac{(N-n)(N-n-1)}{N(N-1)},
\end{aligned}$$

Under SRS, we have $\pi_{N,i}^{-1} - 1 = (N - n)n^{-1}$ for $i = 1, \dots, N$. Consider

$$\begin{aligned}
&E[(I_{N,i}\pi_{N,i}^{-1} - 1)^3(I_{N,j}\pi_{N,j}^{-1} - 1) \mid \mathcal{F}_N] \\
&= \frac{(N-n)^4}{n^4} \frac{n(n-1)}{N(N-1)} - \frac{(N-n)^3}{n^3} \frac{n(N-n)}{N(N-1)} \\
&\quad - \frac{N-n}{n} \frac{n(N-n)}{N(N-1)} + \frac{(N-n)(N-n-1)}{N(N-1)} \\
&= -\frac{(N-n)^4}{n^3N(N-1)} + O(1) \\
&= O(N^2n^{-2}), \tag{A.42}
\end{aligned}$$

where the last equality holds by the facts that $(N - n)^4[n^3N(N - 1)]^{-1} = O(N^2n^{-3})$ and $N^2n^{-2} = O(1)$ if $n \asymp N$. Thus, we have proved (A.38) by (A.42).

Consider

$$\begin{aligned}
& E[(I_{N,i}\pi_{N,i}^{-1} - 1)^2(I_{N,j}\pi_{N,j}^{-1} - 1)^2 \mid \mathcal{F}_N] \\
&= \frac{(N-n)^4}{n^4} \frac{n(n-1)}{N(N-1)} + 2 \frac{(N-n)^2}{n^2} \frac{n(N-n)}{N(N-1)} + \frac{(N-n)(N-n-1)}{N(N-1)} \\
&= O(N^2n^{-2}) + O(Nn^{-1}) + O(1) \\
&= O(N^2n^{-2}),
\end{aligned}$$

which proves (A.39).

Similar with the case for two terms, we have the following results under SRS. That is,

$$\begin{aligned}
P_N(\#\{I_{N,i}, I_{N,j}, I_{N,k}\} = 3) &= \frac{n(n-1)(n-2)}{N(N-1)(N-2)}, \\
P_N(\{I_{N,i} = 1, I_{N,j} = 1, I_{N,k} = 0\}) &= \frac{n(n-1)(N-n)}{N(N-1)(N-2)}, \\
P_N(\{I_{N,i} = 1, I_{N,j} = 0, I_{N,k} = 0\}) &= \frac{n(N-n)(N-n-1)}{N(N-1)(N-2)}, \\
P_N(\#\{I_{N,i}, I_{N,j}, I_{N,k}\} = 0) &= \frac{(N-n)(N-n-1)(N-n-2)}{N(N-1)(N-2)}.
\end{aligned}$$

Consider

$$\begin{aligned}
& E[(I_{N,i}\pi_{N,i}^{-1} - 1)(I_{N,j}\pi_{N,j}^{-1} - 1)(I_{N,k}\pi_{N,k}^{-1} - 1)^2 \mid \mathcal{F}_N] \\
&= \frac{(N-n)^4}{n^4} \frac{n(n-1)(n-2)}{N(N-1)(N-2)} + \frac{(N-n)^2}{n^2} \frac{n(n-1)(N-n)}{N(N-1)(N-2)} \\
&\quad - 2 \frac{(N-n)^3}{n^3} \frac{n(n-1)(N-n)}{N(N-1)(N-2)} - 2 \frac{N-n}{n} \frac{n(N-n)(N-n-1)}{N(N-1)(N-2)} \\
&\quad + \frac{(N-n)^2}{n^2} \frac{n(N-n)(N-n-1)}{N(N-1)(N-2)} + \frac{(N-n)(N-n-1)(N-n-2)}{N(N-1)(N-2)} \\
&= \frac{(n-1)(n-2)(N-n)^4}{n^3N(N-1)(N-2)} + \frac{(N-n)^3}{nN(N-1)} - 2 \frac{(n-1)(N-n)^4}{n^2N(N-1)(N-2)} \\
&\quad + \frac{(N-n)(N-n-1)(N-n-2) - 2(N-n)^2(N-n-1)}{N(N-1)(N-2)}. \tag{A.43}
\end{aligned}$$

After some algebra, the first three terms of (A.43) are

$$\begin{aligned}
& \frac{(n-1)(n-2)(N-n)^4}{n^3N(N-1)(N-2)} + \frac{(N-n)^3}{nN(N-1)} \\
& - 2 \frac{(n-1)(N-n)^4}{n^2N(N-1)(N-2)} \\
& = \frac{(N-n)^3}{N(N-1)(N-2)} + O(Nn^{-2}).
\end{aligned} \tag{A.44}$$

Together with (A.43) and (A.44), we have

$$\begin{aligned}
& E[(I_{N,i}\pi_{N,i}^{-1} - 1)(I_{N,j}\pi_{N,j}^{-1} - 1)(I_{N,k}\pi_{N,k}^{-1} - 1)^2 \mid \mathcal{F}_N] \\
& = O(Nn^{-2}) + \frac{(N-n)^3 - (N-n)(N-n-1)(N-n+2)}{N(N-1)(N-2)} \\
& = O(Nn^{-2}),
\end{aligned} \tag{A.45}$$

where the last equality holds by (C6) and the fact that the second term in the first equality converges to 0. Thus, we have shown (A.40) by (A.45).

For four terms under SRS, we have

$$\begin{aligned}
P_N(\#\{I_{N,i}, I_{N,j}, I_{N,k}, I_{N,l}\} = 4) &= \frac{n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)}, \\
P_N(\{I_{N,i} = 1, I_{N,j} = 1, I_{N,k} = 1, I_{N,l} = 0\}) &= \frac{n(n-1)(n-2)(N-n)}{N(N-1)(N-2)(N-3)}, \\
P_N(\{I_{N,i} = 1, I_{N,j} = 1, I_{N,k} = 0, I_{N,l} = 0\}) &= \frac{n(n-1)(N-n)(N-n-1)}{N(N-1)(N-2)(N-3)}, \\
P_N(\{I_{N,i} = 1, I_{N,j} = 0, I_{N,k} = 0, I_{N,l} = 0\}) &= \frac{n(N-n)(N-n-1)(N-n-2)}{N(N-1)(N-2)(N-3)}, \\
P_N(\#\{I_{N,i}, I_{N,j}, I_{N,k}, I_{N,l}\} = 0) &= \frac{(N-n)(N-n-1)(N-n-2)(N-n-3)}{N(N-1)(N-2)(N-3)}.
\end{aligned}$$

Now, consider

$$\begin{aligned}
& E[(I_{N,i}\pi_{N,i}^{-1} - 1)(I_{N,j}\pi_{N,j}^{-1} - 1)(I_{N,k}\pi_{N,k}^{-1} - 1)(I_{N,l}\pi_{N,l}^{-1} - 1) \mid \mathcal{F}_N] \\
= & \frac{(N-n)^4}{n^4} \frac{n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)} \\
& - 4 \frac{(N-n)^3}{n^3} \frac{n(n-1)(n-2)(N-n)}{N(N-1)(N-2)(N-3)} \\
& + 6 \frac{(N-n)^2}{n^2} \frac{n(n-1)(N-n)(N-n-1)}{N(N-1)(N-2)(N-3)} \\
& - 4 \frac{(N-n)}{n} \frac{n(N-n)(N-n-1)(N-n-2)}{N(N-1)(N-2)(N-3)} \\
& + \frac{(N-n)(N-n-1)(N-n-2)(N-n-3)}{N(N-1)(N-2)(N-3)}. \tag{A.46}
\end{aligned}$$

Consider

$$\begin{aligned}
& \frac{(n-1)(n-2)(n-3)}{n^3} (N-n)^4 \\
& - 4 \frac{(n-1)(n-2)}{n^2} (N-n)^4 \\
& + 6 \frac{(n-1)}{n} (N-n)^3 (N-n-1) \\
& - 4(N-n)^2 (N-n-1)(N-n-2) \\
& + (N-n)(N-n-1)(N-n-2)(N-n-3) \\
= & \frac{3(N-n)^4}{n^2} - \frac{6(N-n)^4}{n^3} + \frac{6(N-n)^3}{n} \\
& + 3(N-n)^2 - 6(N-n) \\
= & O(N^4 n^{-2}), \tag{A.47}
\end{aligned}$$

where the last equality is valid by (C6). Together with (A.46) and (A.47), we have

$$\begin{aligned}
& E[(I_{N,i}\pi_{N,i}^{-1} - 1)(I_{N,j}\pi_{N,j}^{-1} - 1)(I_{N,k}\pi_{N,k}^{-1} - 1)(I_{N,l}\pi_{N,l}^{-1} - 1) \mid \mathcal{F}_N] \\
= & O(N^4 n^{-2}) \{N(N-1)(N-2)(N-3)\}^{-1} \\
= & O(n^{-2}).
\end{aligned}$$

Thus, we have proved (A.41). □

Proof of Lemma 4.1. Based on basic algebra and (C3), we can show (9), so the proof is omitted here.

Note that

$$\begin{aligned} s_{N,SRS}^2 &= n^{-1} \sum_{i=1}^n y_i^2 - \left(n^{-1} \sum_{i=1}^n y_i \right)^2, \\ \sigma_{N,SRS}^2 &= N^{-1} \sum_{i=1}^N y_i^2 - \left(N^{-1} \sum_{i=1}^N y_i \right)^2. \end{aligned}$$

To show (10), it is enough to show that

$$n^{-1} \sum_{i=1}^n y_i^2 - N^{-1} \sum_{i=1}^N y_i^2 \rightarrow 0, \tag{A.48}$$

$$n^{-1} \sum_{i=1}^n y_i - N^{-1} \sum_{i=1}^N y_i \rightarrow 0 \tag{A.49}$$

almost surely.

First, we show (A.48), and we have

$$n^{-1} \sum_{i=1}^n y_i^2 - N^{-1} \sum_{i=1}^N y_i^2 = N^{-1} \sum_{i=1}^N (I_{N,i} \pi_{N,i}^{-1} - 1) y_{N,i}^2.$$

Based on (C6) and the proof of Lemma 3.1, it is enough to show that

$$N^{-4} E \left[\left\{ \sum_{i=1}^N (I_{N,i} \pi_{N,i}^{-1} - 1) y_{N,i}^2 \right\}^4 \mid \mathcal{F}_N \right] = O(n^{-2}).$$

By some basic algebra and (C3), we have

$$\sum_{i \neq j} y_{N,i}^6 y_{N,j}^2 = O(N^2), \quad (\text{A.50})$$

$$\sum_{i \neq j \neq k} y_{N,i}^2 y_{N,j}^2 y_{N,k}^4 = O(N^3), \quad (\text{A.51})$$

$$\sum_{i \neq j \neq k \neq l} y_{N,i}^2 y_{N,j}^2 y_{N,k}^2 y_{N,l}^2 = O(N^4), \quad (\text{A.52})$$

where $i \neq j \neq k$ and $i \neq j \neq k \neq l$ are defined similarly as $i \neq j$ for (A.4).

Consider

$$\begin{aligned} & N^{-4} E \left[\left\{ \sum_{i=1}^N (I_{N,i} \pi_{N,i}^{-1} - 1) y_{N,i}^2 \right\}^4 \mid \mathcal{F}_N \right] \\ = & N^{-4} \sum_{i=1}^N E \{ (I_{N,i} \pi_{N,i}^{-1} - 1)^4 \mid \mathcal{F}_N \} y_{N,i}^8 \\ & + N^{-4} \sum_{i \neq j} E \{ (I_{N,i} \pi_{N,i}^{-1} - 1)^2 (I_{N,j} \pi_{N,j}^{-1} - 1)^2 \mid \mathcal{F}_N \} y_{N,i}^4 y_{N,j}^4 \\ & + N^{-4} \sum_{i \neq j} E \{ (I_{N,i} \pi_{N,i}^{-1} - 1)^3 (I_{N,j} \pi_{N,j}^{-1} - 1) \mid \mathcal{F}_N \} y_{N,i}^6 y_{N,j}^2 \\ & + N^{-4} \sum_{i \neq j \neq k} E \{ (I_{N,i} \pi_{N,i}^{-1} - 1) (I_{N,j} \pi_{N,j}^{-1} - 1) (I_{N,k} \pi_{N,k}^{-1} - 1)^2 \mid \mathcal{F}_N \} y_{N,i}^2 y_{N,j}^2 y_{N,k}^4 \\ & + N^{-4} \sum_{i \neq j \neq k \neq l} E \{ (I_{N,i} \pi_{N,i}^{-1} - 1) (I_{N,j} \pi_{N,j}^{-1} - 1) (I_{N,k} \pi_{N,k}^{-1} - 1) (I_{N,l} \pi_{N,l}^{-1} - 1) \mid \mathcal{F}_N \} y_{N,i}^2 y_{N,j}^2 y_{N,k}^2 y_{N,l}^2 \\ = & O(n^{-2}), \end{aligned}$$

where the last equality holds by Lemma 9.7, (A.4) and (A.50) to (A.52). Thus, we have proved (A.48). Similarly, we can prove (A.49).

Note that

$$\mu_{N,SRS}^{(3)} = N^{-1} \sum_{i=1}^N y_{N,i}^3 - 3\bar{Y}_N N^{-1} \sum_{i=1}^N y_{N,i}^2 + 2\bar{Y}_N^3.$$

To show (11), by (A.48) and (A.49), it remains to show

$$n^{-1} \sum_{i=1}^n y_i^3 - N^{-1} \sum_{i=1}^N y_{N,i}^3 \rightarrow 0, \quad (\text{A.53})$$

in probability.

Note that

$$n^{-1} \sum_{i=1}^n y_i^3 - N^{-1} \sum_{i=1}^N y_{N,i}^3 = N^{-1} \sum_{i=1}^N (I_{N,i} \pi_{N,i}^{-1} - 1) y_{N,i}^3.$$

Consider

$$E \left(n^{-1} \sum_{i=1}^n y_i^3 - N^{-1} \sum_{i=1}^N y_{N,i}^3 \mid \mathcal{F}_N \right) = 0 \quad (\text{A.54})$$

$$\begin{aligned} \text{var} \left(n^{-1} \sum_{i=1}^n y_i^3 - N^{-1} \sum_{i=1}^N y_{N,i}^3 \mid \mathcal{F}_N \right) &= n^{-1} (1 - nN^{-1}) \sigma_{N,3}^2 \\ &= O(n^{-1}), \end{aligned} \quad (\text{A.55})$$

where (A.54) holds by the sampling design, $\sigma_{N,3}^2$ is the finite population variance of $\{y_{N,1}^3, \dots, y_{N,N}^3\}$, and the second equality of (A.55) holds by (C3). Together with (A.54) and (A.55), we have proved (A.53). \square

Proof of Theorem 4.2. First, we show that

$$N^{-1} \sum_{i=1}^N (y_{N,i}^*)^8 = O_p(1). \quad (\text{A.56})$$

Consider

$$N^{-1} E_* \left\{ \sum_{i=1}^N (y_{N,i}^*)^8 \right\} = n^{-1} \sum_{i=1}^n y_{N,i}^8 \quad (\text{A.57})$$

$$\begin{aligned} E \left(n^{-1} \sum_{i=1}^n y_{N,i}^8 \mid \mathcal{F}_N \right) &= N^{-1} \sum_{i=1}^N y_{N,i}^8 \\ &= O(1). \end{aligned} \quad (\text{A.58})$$

Together with (A.57) and (A.58), we have proved (A.56) using the Markov inequality.

By (C8), there exists a strongly non-latticed distribution G_{SRS} such that

$$N^{-1} \sum_{i=1}^N \exp(\iota y_i) \rightarrow \int \exp(\iota x) dG_{SRS}(x) \quad (\text{A.59})$$

as $N \rightarrow \infty$ for $t \in \mathbb{R}$. Next, we show that

$$N^{-1} \sum_{i=1}^N \exp(\iota y_i^*) \rightarrow \int \exp(\iota x) dG_{SRS}(x) \quad (\text{A.60})$$

in probability as $N \rightarrow \infty$ for $t \in \mathbb{R}$, where $\{y_i^* : i = 1, \dots, N\}$ is the bootstrap finite population.

By Euler's formula, we have

$$\begin{aligned} N^{-1} \sum_{i=1}^N \exp(\iota y_i) &= N^{-1} \sum_{i=1}^N \cos(\iota y_i) + \iota N^{-1} \sum_{i=1}^N \sin(\iota y_i), \\ N^{-1} \sum_{i=1}^N \exp(\iota y_i^*) &= N^{-1} \sum_{i=1}^n N_i^* \cos(\iota y_i) + \iota N^{-1} \sum_{i=1}^n N_i^* \sin(\iota y_i). \end{aligned}$$

It is enough to show that

$$N^{-1} \sum_{i=1}^n N_i^* \cos(\iota y_i) = N^{-1} \sum_{i=1}^N \cos(\iota y_i) + o_p(1), \quad (\text{A.61})$$

$$N^{-1} \sum_{i=1}^n N_i^* \sin(\iota y_i) = N^{-1} \sum_{i=1}^N \sin(\iota y_i) + o_p(1). \quad (\text{A.62})$$

We only show the result (A.61), and (A.62) can be obtained in a similar manner.

By the first step of the proposed bootstrap method, we have

$$N^{-1} E_* \left\{ \sum_{i=1}^n N_i^* \cos(\iota y_i) \right\} = n^{-1} \sum_{i=1}^n \cos(\iota y_i) \quad (\text{A.63})$$

$$\begin{aligned} N^{-2} \text{var}_* \left\{ \sum_{i=1}^n N_i^* \cos(\iota y_i) \right\} &\leq N^{-2} \sum_{i=1}^n N_i n^{-1} \\ &= o(1), \end{aligned} \quad (\text{A.64})$$

where the inequality of (A.64) holds by the negative correlation among N_i^* and the fact that $|\cos(ty_i)| \leq 1$. Using similar argument in (A.54) and (A.55), we can have shown (A.61) by results in (A.63) and (A.64). By (A.61) and (A.62), we have proved (A.60).

Thus, by (C6), (A.56) and (A.60), we have

$$\hat{F}_{N,SRS}^*(z) = \Phi(z) - \frac{(1 - n/N)^{1/2} \mu_{N,SRS}^{(3)*}}{6n^{1/2}(\sigma_{N,SRS}^*)^3} \left\{ \frac{1 - 2n/N}{1 - n/N} (z^2 - 1) - 3z^2 \right\} \phi(z) + o_p(n^{-1/2})$$

almost surely conditional on the generated bootstrap finite population, where $\mu_{SRS}^{(3)*}$ and $(\sigma_{SRS}^*)^2$ are the bootstrap central third moment and variance.

Based on Lemma 4.1, it remains to show that

$$\mu_{N,SRS}^{(3)*} - \hat{\mu}_{N,SRS}^{(3)} \rightarrow 0 \quad (\text{A.65})$$

$$(\sigma_{N,SRS}^*)^2 - s_{N,SRS}^2 \rightarrow 0 \quad (\text{A.66})$$

in probability. By some algebra, it is equivalent to show

$$N^{-1} \sum_{i=1}^N (y_{N,i}^*)^3 - n^{-1} \sum_{i=1}^n y_{N,i}^3 \rightarrow 0, \quad (\text{A.67})$$

$$N^{-1} \sum_{i=1}^N (y_{N,i}^*)^2 - n^{-1} \sum_{i=1}^n y_{N,i}^2 \rightarrow 0, \quad (\text{A.68})$$

$$N^{-1} \sum_{i=1}^N y_{N,i}^* - n^{-1} \sum_{i=1}^n y_{N,i} \rightarrow 0, \quad (\text{A.69})$$

in probability.

Consider

$$N^{-1} E_* \left\{ \sum_{i=1}^N (y_{N,i}^*)^3 \right\} = n^{-1} \sum_{i=1}^n y_{N,i}^3 \quad (\text{A.70})$$

$$\begin{aligned} \text{var}_* \left\{ N^{-1} \sum_{i=1}^N (y_{N,i}^*)^3 \right\} &\leq N^{-2} \sum_{i=1}^n N n^{-1} y_{N,i}^6 \\ &= o_p(1), \end{aligned} \quad (\text{A.71})$$

where the last equality of (A.71) is derived by the Markov inequality and a similar procedure for (A.58). Thus, we have proved (A.67) by (A.70) and (A.71). Similarly, we can prove (A.68) and (A.69). Therefore, we have shown (A.65) and (A.66), which concludes the proof of Theorem 4.2. \square

Proof of Lemma 5.1. Mentioned that $Z_{N,1}, \dots, Z_{N,n} \stackrel{i.i.d.}{\sim} G_{N,PPS}$ and $\mathbb{P}(Z_{N,i} = p_{N,k}^{-1} y_{N,k}) = p_{N,k}$ for $i = 1, \dots, n; k = 1, \dots, N$, we have

$$\begin{aligned} E(N^{-\delta} |Z_{N,i}|^\delta | \mathcal{F}_N) &= N^{-\delta} \sum_{i=1}^N p_{N,i}^{-(\delta-1)} |y_{N,i}|^\delta \\ &= O\left(N^{-1} \sum_{i=1}^N |y_{N,i}|^\delta\right) \\ &= O(1), \end{aligned} \tag{A.72}$$

for all positive $\delta \leq 8$, where the second equality holds by (C9), and the last equality holds by (C3).

By the strong law of large numbers,

$$n^{-1} \sum_{i=1}^n N^{-2} Z_{N,i}^2 - E(N^{-2} Z_{N,i}^2 | \mathcal{F}_N) \rightarrow 0 \tag{A.73}$$

almost surely and

$$n^{-1} \sum_{i=1}^n N^{-1} Z_{N,i} - E(N^{-1} Z_{N,i} | \mathcal{F}_N) \rightarrow 0 \tag{A.74}$$

almost surely.

Note that

$$\begin{aligned} N^{-2} s_{N,PPS}^2 &= n^{-1} \sum_{i=1}^n N^{-2} Z_{N,i}^2 - \text{big}\left(n^{-1} \sum_{i=1}^n N^{-1} Z_{N,i}\right)^2, \\ N^{-2} \sigma_{N,PPS}^2 &= E(N^{-2} Z_{N,i}^2 | \mathcal{F}_N) - \{E(N^{-1} Z_{N,i} | \mathcal{F}_N)\}^2. \end{aligned}$$

By (A.73) and (A.74), we have proved (14).

Notice that

$$\begin{aligned}
N^{-3}\mu_{PPS}^{(3)} &= N^{-3}\sum_{i=1}^N p_{N,i}(p_{N,i}^{-3}y_{N,i}^3 - 3p_{N,i}^{-2}y_{N,i}^2Y_N + 3p_{N,i}^{-1}y_{N,i}Y_N^2 - Y_N^3) \\
&= N^{-3}\sum_{i=1}^N p_{N,i}^{-2}y_{N,i}^3 - 3N^{-3}Y_N\sum_{i=1}^N p_{N,i}^{-1}y_{N,i}^2 + 2N^{-3}Y_N^3 \\
&= O(1),
\end{aligned} \tag{A.75}$$

where $Y_N = E(Z_{N,i} | \mathcal{F}_N) = \sum_{i=1}^N y_{N,i}$, and the last equality of (A.75) holds by (C9) and (A.72). In addition, for $\zeta = 1, 2, 3$, we have

$$E\left[\left\{n^{-1}\sum_{i=1}^n N^{-\zeta}Z_{N,i}^{\zeta} - E(N^{-\zeta}Z_{N,i}^{\zeta} | \mathcal{F}_N)\right\} | \mathcal{F}_N\right] = 0$$

and

$$\text{var}\left(n^{-1}\sum_{i=1}^n N^{-\zeta}Z_{N,i}^{\zeta} | \mathcal{F}_N\right) \leq n^{-1}N^{-2\zeta}E(Z_{N,i}^{2\zeta} | \mathcal{F}_N) = O(n^{-1}).$$

By Markov's inequality,

$$n^{-1}\sum_{i=1}^n N^{-\zeta}Z_{N,i}^{\zeta} - E(N^{-\zeta}Z_{N,i}^{\zeta} | \mathcal{F}_N) = O_p(n^{-1/2})$$

for $\zeta = 1, 2, 3$, from which we can prove that $N^{-3}(\hat{\mu}_{N,PPS}^{(3)} - \mu_{N,PPS}^{(3)}) = O_p(n^{-1/2})$. Thus, we complete the proof of Lemma 5.1. \square

Proof of Theorem 5.1. Rewrite

$$\begin{aligned}
T_{N,PPS} &= \hat{V}_{N,PPS}^{-1/2}(\hat{Y}_{N,PPS} - Y_N) \\
&= n^{1/2}\left(n^{-1}\sum_{i=1}^n N^{-1}Z_{N,i} - N^{-1}Y_N\right) \\
&\quad \times \left\{n^{-1}\sum_{i=1}^n (N^{-1}Z_{N,i})^2 - \left(n^{-1}\sum_{i=1}^n N^{-1}Z_{N,i}\right)^2\right\}^{-1/2},
\end{aligned}$$

where $N^{-1}Y_N = E(N^{-1}Z_{N,i} \mid \mathcal{F}_N)$. By (A.72), we have $E(|N^{-1}Z_{N,i}|^3 \mid \mathcal{F}_N) < \infty$. Using the results of Hall (1987), as the distribution of $Z_{N,i}$ is non-lattice, we have

$$\hat{F}_{N,PPS}(z) = \Phi(z) + \frac{\mu_{N,PPS}^{(3)}}{6\sqrt{n}\sigma_{N,PPS}^3}(2z^2 + 1)\phi(z) + o(n^{-1/2}),$$

where $\sigma_{N,PPS}^2 = E\{(Z_{N,i} - Y_N)^2 \mid \mathcal{F}_N\} = \sum_{i=1}^N p_{N,i}(p_{N,i}^{-1}y_{N,i} - Y_N)^2$ and $\mu_{N,PPS}^{(3)} = E\{(Z_{N,i} - Y_N)^3 \mid \mathcal{F}_N\} = \sum_{i=1}^N p_{N,i}(p_{N,i}^{-1}y_{N,i} - Y_N)^3$. Finally, according to Lemma 5.1, we have shown Theorem 5.1. \square

Proof of Theorem 5.2. Mentioned that $\mathbb{P}_{PPS}(p_{N,a,i} = p_{N,k}) = p_{N,k}$ for $i = 1, \dots, n$; $k = 1, \dots, N$ and $p_{N,a,1}, \dots, p_{N,a,n}$ are independent. For any a such that $0 \leq \delta \leq 8$,

$$E \left\{ \left| n^{-1} \sum_{i=1}^n N^{-1} p_{N,a,i}^{-1} y_{N,a,i}^\delta \right| \mid \mathcal{F}_N \right\} \leq N^{-1} \sum_{i=1}^N |y_{N,i}|^\delta < \infty.$$

Thus, by SLLN,

$$n^{-1} \sum_{i=1}^n N^{-1} p_{N,a,i}^{-1} y_{N,a,i}^\delta \rightarrow N^{-1} \sum_{i=1}^N y_{N,i}^\delta \quad (\text{A.76})$$

with probability 1 for all $0 \leq \delta \leq 8$.

In the first step of our proposed bootstrap method, $y_{N,1}^*, \dots, y_{N,n}^*$ are independently and identically distributed (i.i.d.) with $\mathbb{P}_*(y_{N,i}^* = y_{N,a,j}) = \rho_{N,j} = p_{N,a,j}^{-1} \left(\sum_{\ell=1}^n p_{N,a,\ell}^{-1} \right)^{-1}$ for $i = 1, \dots, n$; $j = 1, \dots, n$. Let \mathbb{P}_* be the probability measure for the first step of the proposed bootstrap method conditional on the realized sample $\{y_{N,a,1}, \dots, y_{N,a,n}\}$. Then, we

have

$$\begin{aligned}
& E_* \left\{ N^{-1} \sum_{i=1}^N (y_{N,i}^*)^8 \right\} \\
&= \sum_{i=1}^n p_{N,a,i}^{-1} \left(\sum_{j=1}^n p_{N,a,j}^{-1} \right)^{-1} y_{N,a,i}^8 \\
&= \left(n^{-1} \sum_{i=1}^n N^{-1} p_{N,a,i}^{-1} \right)^{-1} n^{-1} \sum_{i=1}^n N^{-1} p_{N,a,i}^{-1} y_{N,a,i}^8 \\
&= O_p(1), \tag{A.77}
\end{aligned}$$

from which, we get that conditional on the series of realized samples,

$$N^{-1} \sum_{i=1}^N (y_{N,i}^*)^8 = O_p(1). \tag{A.78}$$

Recall that $C_N^* = \sum_{i=1}^n N_{a,i}^* p_{N,a,i}$, where $N_{a,i}^*$ is the number of repetitions of the i -th realized sample in the proposed bootstrap method. Next, we show

$$C_N^* = 1 + o_p(1). \tag{A.79}$$

Consider

$$\begin{aligned}
E_*(C_N^*) &= \left(n^{-1} \sum_{i=1}^n N^{-1} p_{N,a,i}^{-1} \right)^{-1}, \tag{A.80} \\
\text{var}_*(C_N^*) &\leq N \left(\sum_{i=1}^n p_{N,a,i}^{-1} \right)^{-1} \sum_{i=1}^n p_{N,a,i} \\
&= O \left(N^{-1} \left(n^{-1} \sum_{i=1}^n N^{-1} p_{N,a,i}^{-1} \right)^{-1} \right), \tag{A.81}
\end{aligned}$$

where the equality of (A.81) holds by (C9). By (A.80) and (A.81), we have shown (A.79) according to (A.76) with $\delta = 0$.

Similarly, we can show

$$nN_{a,i}^*p_{N,a,i} = 1 + o_p(1) \quad (\text{A.82})$$

for $i = 1, \dots, n$.

Let $((C_N^*)^{-1}p_{N,b,i}^*, y_{N,b,i}^*)$ be the quantities for the i -th selected element from the bootstrap finite population \mathcal{F}_N^* . Denote $Z_{N,i}^* = C_N^*(p_{N,b,i}^*)^{-1}y_{N,b,i}^*$ for $i = 1, \dots, n$. Then $\mathbb{P}_{PPS}^*(Z_{N,i}^* = C_N^*(p_{N,k}^*)^{-1}y_{N,k}^*) = (C_N^*)^{-1}p_{N,k}^*$ for $i = 1, \dots, n; k = 1, \dots, N$, where \mathbb{P}_{PPS}^* is the counterpart of \mathbb{P}_{PPS} conditional on the bootstrap finite population \mathcal{F}_N^* .

Conditional on the bootstrap finite population \mathcal{F}_N^* , denote $F_{N,PPS}^*(z) = \mathbb{P}_{PPS}^*(T_{N,PPS}^* \leq z)$ as the distribution of $T_{N,PPS}^* = (\hat{V}_{N,PPS}^*)^{-1/2}(\hat{Y}_{N,PPS}^* - Y_N^*)$, where $Y_N^* = \sum_{i=1}^N y_{N,i}^*$, $\hat{Y}_{N,PPS}^* = n^{-1} \sum_{i=1}^n Z_{N,i}^*$ and $V_{N,PPS}^* = n^{-2} \sum_{i=1}^n (Z_{N,i}^* - \bar{Z}_N^*)^2$ with $\bar{Z}_N^* = n^{-1} \sum_{i=1}^n Z_{N,i}^* = \hat{Y}_{N,PPS}^*$.

Consider

$$\begin{aligned} E \left(n^{-1} \sum_{i=1}^n N^{-3} |Z_{N,i}|^3 \mid \mathcal{F}_N \right) &= E(N^{-3} |Z_{N,i}|^3 \mid \mathcal{F}_N) \\ &= O(1), \end{aligned} \quad (\text{A.83})$$

$$\begin{aligned} \text{var} \left(n^{-1} \sum_{i=1}^n N^{-3} |Z_{N,i}|^3 \mid \mathcal{F}_N \right) &= n^{-1} \text{var}(N^{-3} |Z_{N,i}|^3 \mid \mathcal{F}_N) \\ &\leq n^{-1} E(N^{-6} Z_{N,i}^6 \mid \mathcal{F}_N) \\ &= O(n^{-1}), \end{aligned} \quad (\text{A.84})$$

where the results of (A.83) and (A.84) are based on (A.72).

Recall that $E_{**}(\cdot)$ is the expectation with respect to the sampling design conditional on the bootstrap finite population and $\{p_{N,k}^* : k = 1, \dots, N\}$ consists of $N_{a,i}^*$ copies of $p_{N,a,i}$ for

$i = 1, \dots, n$. Consider

$$\begin{aligned}
& E_{**} (|N^{-1}Z_{N,i}^*|^3) \\
&= N^{-3} \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* |C_N^* (p_{N,i}^*)^{-1} y_{N,i}^*|^3 \\
&= N^{-3} (C_N^*)^2 \sum_{i=1}^N (p_{N,i}^*)^{-2} |y_{N,i}^*|^3 \\
&= N^{-3} (C_N^*)^2 \sum_{i=1}^n N_{a,i}^* p_{N,a,i}^{-2} |y_{N,a,i}|^3 \\
&= n^{-1} \sum_{i=1}^n |N^{-1} p_{N,a,i}^{-1} y_{N,a,i}|^3 \{1 + o_p(1)\} \\
&= O_p(1), \tag{A.85}
\end{aligned}$$

where the fourth equality holds by (A.82), and last equality holds by Lemma 5.1, (C9), (A.83) and (A.84).

Consider the characteristic function of $N^{-1}Z_{N,i}$ and $N^{-1}Z_{N,i}^*$. Specifically, the characteristic function of $N^{-1}Z_{N,i}$ is

$$\psi_{Z,N}(t) = \sum_{i=1}^N p_{N,i} \exp(itN^{-1}y_{N,i}/p_{N,i})$$

and the characteristic function of $N^{-1}Z_{N,i}^*$, conditional on the bootstrap finite population \mathcal{F}_N^* , is

$$\psi_{Z,N}^*(t) = \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* \exp(itN^{-1}C_N^* y_{N,i}^*/p_{N,i}^*)$$

To show that the distribution of $Z_{N,i}^*$ is non-lattice in probability conditional on the bootstrap finite population \mathcal{F}_N^* , it is enough to show that, for any fixed $t_0 > 0$,

$$\sup_{|t| \leq t_0} |\psi_{Z,N}^*(t) - \psi_{Z,N}(t)| \rightarrow 0 \tag{A.86}$$

in probability as $n \rightarrow \infty$. By remarking that

$$\begin{aligned}
& \sup_{|t| \leq t_0} |\psi_{Z,N}^*(t) - \psi_{Z,N}(t)| \\
& \leq \sup_{|t| \leq t_0} \left| \psi_{Z,N}^*(t) - \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* \exp(itN^{-1} y_{N,i}^*/p_{N,i}^*) \right| \\
& \quad + \sup_{|t| \leq t_0} \left| \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* \exp(itN^{-1} y_{N,i}^*/p_{N,i}^*) - \psi_{Z,N}(t) \right|. \tag{A.87}
\end{aligned}$$

First,

$$\begin{aligned}
& \sup_{|t| \leq t_0} \left| \psi_{Z,N}^*(t) - \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* \exp(itN^{-1} y_{N,i}^*/p_{N,i}^*) \right| \\
& \leq \sup_{|t| \leq t_0} \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* |\exp(itN^{-1} C_N^* y_{N,i}^*/p_{N,i}^*) - \exp(itN^{-1} y_{N,i}^*/p_{N,i}^*)| \\
& \leq t_0 N^{-1} \sum_{i=1}^N (C_N^*)^{-1} |y_{N,i}^*| |C_N^* - 1| \rightarrow 0
\end{aligned}$$

in probability as $n \rightarrow \infty$. Second,

$$\begin{aligned}
& \sup_{|t| \leq t_0} \left| \sum_{i=1}^N (C_N^*)^{-1} p_{N,i}^* \exp(itN^{-1} y_{N,i}^*/p_{N,i}^*) - \psi_{Z,N}(t) \right| \\
& = \sup_{|t| \leq t_0} \left| (C_N^*)^{-1} \sum_{i=1}^n N_i^* p_{N,a,i} \exp(itN^{-1} y_{N,a,i}/p_{N,a,i}) \right. \\
& \quad \left. - \sum_{i=1}^N p_{N,i} \cos(tN^{-1} y_{N,i}/p_{N,i}) - \iota \sum_{i=1}^N p_{N,i} \sin(tN^{-1} y_{N,i}/p_{N,i}) \right| \\
& \leq \left| (C_N^*)^{-1} \sum_{i=1}^n N_i^* p_{N,a,i} \cos(tN^{-1} y_{N,a,i}/p_{N,a,i}) \right. \\
& \quad \left. - \sum_{i=1}^N p_{N,i} \cos(tN^{-1} y_{N,i}/p_{N,i}) \right| \\
& \quad + \left| (C_N^*)^{-1} \sum_{i=1}^n N_i^* p_{N,a,i} \sin(tN^{-1} y_{N,a,i}/p_{N,a,i}) \right. \\
& \quad \left. - \sum_{i=1}^N p_{N,i} \sin(tN^{-1} y_{N,i}/p_{N,i}) \right|.
\end{aligned}$$

It suffices to show that

$$\begin{aligned} & (C_N^*)^{-1} \sum_{i=1}^n N_i^* p_{N,a,i} \cos(tN^{-1}y_{N,a,i}/p_{N,a,i}) \\ &= \sum_{i=1}^N p_{N,i} \cos(tN^{-1}y_{N,i}/p_{N,i}) + o_p(1), \end{aligned} \quad (\text{A.88})$$

$$\begin{aligned} & (C_N^*)^{-1} \sum_{i=1}^n N_i^* p_{N,a,i} \sin(tN^{-1}y_{N,a,i}/p_{N,a,i}) \\ &= \sum_{i=1}^N p_{N,i} \sin(tN^{-1}y_{N,i}/p_{N,i}) + o_p(1). \end{aligned} \quad (\text{A.89})$$

We can show (A.88) since (A.79) and (A.82) hold, $|\cos(ty_{N,a,i}/p_{N,a,i})| \leq 1$ and $Z_{N,1}, \dots, Z_{N,n}$ are independent and identically distributed random variables. Similarly, we can show (A.89). By (A.77), (C11), and the fact that the distribution of $Z_{N,i}^*$ is non-lattice in probability, we have the following result by Hall (1987):

$$\hat{F}_{N,PPS}^*(z) = \Phi(z) + \frac{\mu_{N,PPS}^{(3)*}}{6\sqrt{n}(\sigma_{N,PPS}^*)^3}(2z^2 + 1)\phi(z) + o_p(n^{-1/2}) \quad (\text{A.90})$$

uniformly in $z \in \mathbb{R}$, where $\mu_{N,PPS}^{(3)*} = E_{**}\{(Z_{N,i}^* - Y_N^*)^3\}$.

Based on a similar argument made for (A.85), we have

$$N^{-3}(\mu_{N,PPS}^{(3)*} - \hat{\mu}_{N,PPS}^{(3)}) = o_p(1), \quad (\text{A.91})$$

$$N^{-2}\{(\sigma_{N,PPS}^*)^2 - s_{N,PPS}^2\} = o_p(1). \quad (\text{A.92})$$

Together with Lemma 5.1, (A.90) to (A.92) and (C10), we have proved Theorem 5.2. \square

9.2 Design-unbiased estimates for the two-stage sampling designs

For the two-stage sampling designs in the second simulation study, Poisson sampling and PPS sampling are used in the first stage, and an SRS design is independently conducted within each selected cluster in the second stage. For the sampling designs in the first stage,

denote $\pi_i = n_1 N_i N^{-1}$ to be the first-order inclusion probability for Poisson sampling, and $p_i = N_i N^{-1}$ to be the selection probability for PPS sampling. In this section, we comply to the notation convention in Section 1.2.8 of Fuller (2009) to discuss the variance estimation under the two-stage sampling designs.

For the two-stage sampling design, where Poisson sampling is applied in the first stage, the design-based estimator of \bar{Y} is

$$\tilde{Y} = N^{-1} \sum_{i \in A} \pi_i^{-1} \hat{Y}_{i,\cdot} = n_1^{-1} \sum_{i \in A} N_i^{-1} \hat{Y}_{i,\cdot},$$

where A is the index of the selected clusters in the first stage, $\hat{Y}_{i,\cdot} = N_i n_2^{-1} \sum_{j \in B_i} y_{i,j}$ is a design-unbiased estimate of the cluster total $Y_{i,\cdot} = \sum_{j=1}^{N_i} y_{i,j}$ under the SRS design, and B_i is the index set of the sample within the selected cluster indexed by i . It can be shown that the same form holds when PPS sampling is used in the first stage.

First, we discuss the variance estimator of \tilde{Y} for the two-stage sampling design where Poisson sampling is used in the first stage. As shown in Section 1.2.8 by Fuller (2009), the variance of \tilde{Y} can be decomposed into two parts. That is,

$$\text{var}(\tilde{Y} \mid U_N) = V_1 + V_2, \tag{A.93}$$

where $V_1 = E[\text{var}\{\tilde{Y} \mid (A, U_N)\} \mid U_N]$ and $V_2 = \text{var}[E\{\tilde{Y} \mid (A, U_N)\} \mid U_N]$.

Consider

$$\text{var}\{\tilde{Y} \mid (A, U_N)\} = N^{-2} \sum_{i \in A} \pi_i^{-2} \text{var}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}, \tag{A.94}$$

where the equality holds since the SRS design is independently conducted within each selected cluster, $\text{var}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\} = N_i(N_i - n_2)n_2^{-1}S_i^2$, $S_i^2 = (N_i - 1)^{-1} \sum_{j=1}^{N_i} (y_{i,j} - \bar{Y}_{i,\cdot})^2$ is the finite population variance within the i -th cluster, and $\bar{Y}_{i,\cdot} = N_i^{-1}Y_{i,\cdot}$ is the finite population mean of the i -th cluster. Since the sample variance $s_i^2 = (n_2 - 1)^{-1} \sum_{j \in B_i} (y_{i,j} - \tilde{Y}_{i,\cdot})^2$ is an

unbiased estimator of S_i^2 , where $\tilde{Y}_{i,\cdot} = N_i^{-1}\hat{Y}_{i,\cdot}$ is the estimated cluster mean, the first term of (A.93) can be estimated by

$$\hat{V}_1 = N^{-2} \sum_{i \in A} \pi_i^{-2} \hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}, \quad (\text{A.95})$$

where $\hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\} = N_i(N_i - n_2)n_2^{-1}s_i^2$.

For the second term of (A.93), consider

$$E\{\tilde{Y} \mid (A, U_N)\} = N^{-1} \sum_{i \in A} \pi_i^{-1} Y_{i,\cdot}.$$

Since Poisson sampling is used in the first stage, we have

$$\text{var}[E\{\tilde{Y} \mid (A, U_N)\} \mid U_N] = N^{-2} \sum_{i=1}^H \pi_i^{-1} (1 - \pi_i) Y_{i,\cdot}^2, \quad (\text{A.96})$$

which can be estimated by $N^{-2} \sum_{i \in A} \pi_i^{-2} (1 - \pi_i) Y_{i,\cdot}^2$. Notice that

$$\begin{aligned} E\{\hat{Y}_{i,\cdot}^2 \mid (A, U_N)\} &= [E\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}]^2 + \text{var}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\} \\ &= Y_{i,\cdot}^2 + \text{var}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}. \end{aligned} \quad (\text{A.97})$$

By (A.96) and (A.97) and the fact that s_i^2 is an unbiased estimator of S_i^2 , the second term of (A.93) can be estimated by

$$\hat{V}_2 = N^{-2} \sum_{i \in A} \pi_i^{-2} (1 - \pi_i) [\hat{Y}_{i,\cdot}^2 - \hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}]. \quad (\text{A.98})$$

By (A.95) and (A.98), the variance of \tilde{Y} can be estimated by

$$\tilde{V} = N^{-2} \left[\sum_{i \in A} \pi_i^{-2} (1 - \pi_i) \hat{Y}_{i,\cdot}^2 + \sum_{i \in A} \pi_i^{-1} \hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\} \right],$$

when Poisson sampling is used in the first stage.

Next, we use variance decomposition (A.93) to derive the variance estimator of \tilde{Y} under the two-stage sampling design where PPS sampling is applied in the first stage. The result shown in (A.94) holds, and we can still use (A.95) to approximate V_1 .

Consider

$$\begin{aligned} & \text{var}[E\{\tilde{Y} \mid (A, U_N)\} \mid U_N] \\ &= N^{-2}n_1^{-1}(n_1 - 1)^{-1} \left(\sum_{i \in A} Z_{i,\cdot}^2 - n_1 \bar{Z}^2 \right), \end{aligned} \quad (\text{A.99})$$

where the equality holds by the property of PPS sampling, $Z_{i,\cdot} = Y_{i,\cdot} p_i^{-1}$ and $\bar{Z} = n_1^{-1} \sum_{i \in A} Z_{i,\cdot}$. Based on (A.97), we can estimate $Z_{i,\cdot}^2$ by

$$p_i^{-2}[\hat{Y}_{i,\cdot}^2 - \hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}].$$

Consider

$$\begin{aligned} E\{\tilde{Z}^2 \mid (A, U_N)\} &= \bar{Z}^2 + \text{var}\{\tilde{Z} \mid (A, U_N)\} \\ &= \bar{Z}^2 + n_1^{-2} \sum_{i \in A} p_i^{-2} \text{var}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}, \end{aligned}$$

where $\tilde{Z} = n_1^{-1} \sum_{i \in A} \hat{Y}_{i,\cdot} p_i^{-1}$. Thus, we can estimate \bar{Z}^2 by

$$\tilde{Z}^2 - n_1^{-2} \sum_{i \in A} p_i^{-2} \hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}.$$

By (A.95), (A.99) and the two approximations above, we can obtain the variance estimate of \tilde{Y} by

$$\tilde{V} = N^{-2}n_1^{-1}(n_1 - 1)^{-1} \sum_{i \in A} p_i^{-2} [\hat{Y}_{i,\cdot}^2 + (n_1 - 2)n_1^{-1} \hat{V}\{\hat{Y}_{i,\cdot} \mid (A, U_N)\}] - N^{-2} \tilde{Z}^2$$

for the two-stage sampling design with PPS sampling is used in the first stage.

References

- Antal, E. and Tillé, Y. (2011). A direct bootstrap method for complex sampling designs from a finite population, *J. Amer. Statist. Assoc.* **106**(494): 534–543.
- Athreya, K. B. and Lahiri, S. N. (2006). *Measure Theory and Probability Theory*, Springer Science & Business Media, New York.
- Babu, G. J. and Singh, K. (1984). On one term edgeworth correction by efron's bootstrap, *Sankhya A* **46**(2): 219–232.
- Babu, G. J. and Singh, K. (1985). Edgeworth expansions for sampling without replacement from finite populations, *J. Multivariate Anal.* **17**(3): 261–278.
- Beaumont, J. F. and Patak, Z. (2012). On the generalized bootstrap for sample surveys with special attention to Poisson sampling, *Int. Stat. Rev.* **80**(1): 127–148.
- Bickel, P. J. and Freedman, D. A. (1984). Asymptotic normality and the bootstrap in stratified sampling, *Ann. Statist.* **12**(2): 470–482.
- Booth, J. G., Butler, R. W. and Hall, P. (1994). Bootstrap methods for finite populations, *J. Amer. Statist. Assoc.* **89**(428): 1282–1289.
- Chao, M.-T. and Lo, S.-H. (1985). A bootstrap method for finite population, *Sankhya A* **47**(3): 399–405.
- DiCiccio, T. J. and Romano, J. P. (1995). On bootstrap procedures for second-order accurate confidence limits in parametric models, *Statist. Sinica* **5**(1): 141–160.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife, *Ann. Statist.* **7**(1): 1–26.
- Feller, W. (2008). *An Introduction to Probability Theory and Its Applications*, Vol. 2, John Wiley, New York.

- Fuller, W. A. (2009). *Sampling Statistics*, John Wiley, Hoboken.
- Gross, S. (1980). Median estimation in sample surveys, *Proc. Sect. Surv. Res.*, pp. 181–184.
- Hall, P. (1987). Edgeworth expansion for Student’s t statistic under minimal moment conditions, *Ann. Probab.* (3): 920–931.
- Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer Science & Business Media, New York.
- Hansen, M. H. and Hurwitz, W. N. (1943). On the theory of sampling from finite populations, *Ann. Math. Statist.* **14**(4): 333–362.
- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe, *J. Amer. Statist. Assoc.* **47**(260): 663–685.
- Klenke, A. (2014). *Probability Theory: A Comprehensive Course*, Universitext, second edn, Springer-Verlag London Ltd., London.
- Lai, T. L. and Wang, J. Q. (1993). Edgeworth expansions for symmetric statistics with applications to bootstrap methods, *Statist. Sinica* **3**(3): 517–542.
- Mashreghi, Z., Haziza, D. and Léger, C. (2016). A survey of bootstrap methods in finite population sampling, *Stat. Surv.* **10**: 1–52.
- Petrov, V. V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*, Clarendon Press, Oxford.
- Rao, J. N. K. and Wu, C. F. J. (1988). Resampling inference with complex survey data, *J. Amer. Statist. Assoc.* **83**(401): 231–241.
- Rao, J. N. K., Wu, C. F. J. and Yue, K. (1992). Some recent work on resampling methods for complex surveys, *Surv. Methodol.* **18**(2): 209–217.

- Shao, J. and Sitter, R. R. (1996). Bootstrap for imputed survey data, *J. Amer. Statist. Assoc.* **91**(435): 1278–1288.
- Sitter, R. R. (1992a). A resampling procedure for complex survey data, *J. Amer. Statist. Assoc.* **87**(419): 755–765.
- Sitter, R. R. (1992b). Comparing three bootstrap methods for survey data, *Canad. J. Statist.* **20**(2): 135–154.
- Sverchkov, M. and Pfeffermann, D. (2004). Prediction of finite population totals based on the sample distribution, *Surv. Methodol.* **30**(1): 79–92.