Bootstrap approximation to distributions of finite population U-statistics

Mindaugas Bloznelis
Institute of Mathematics and Informatics and Vilnius University,
Naugarduko 24, Vilnius, LT-03225, Lithuania

E-Mail: mindaugas.bloznelis@mif.vu.lt

Abstract. We show that the without replacement bootstrap of Booth, Butler and Hall (1994) provides second order correct approximation to the distribution function of a Studentized U-statistic based on simple random sample drawn without replacement. In order to achieve similar approximation accuracy for the bootstrap procedure due to Bickel and Freedman (1984) and Chao and Lo (1985) we introduce randomized adjustments to the resampling fraction.

Key Words: Bootstrap, finite population, U-statistic, sampling without replacement.

Mathematics Subject Classification 2000*:* 62E20, 62D05, 62F12, 62F40, 62G20, 62G09

1 Introduction

A remarkable property of Efron's (1979) nonparametric bootstrap is the, so called, second-order correctness first observed by Singh (1981). Singh (1981) showed that the asymptotic accuracy of bootstrap approximation to the distribution function of the standardized sample mean is comparable to the one-term Edgeworth corrected normal approximation. Since then the phenomenon of the second order asymptotic accuracy of the nonparametric bootstrap has been studied by many authors (Babu and Singh (1983, 1984), Bhattacharya and Qumsiyeh (1989), Hall (1988, 1992), Helmers (1991), etc). The second order correctness of the bootstrap approximation was shown for a large class of asymptotically linear statistics admitting an Edgeworth expansion.

For samples drawn *without replacement*, the second-order correct bootstrap approximations were constructed by Booth, Butler and Hall (1994), see also Chen and Sitter (1993) and Sitter (1992). Booth, Butler and Hall (1994) considered Studentized versions of stratified sample means, and Studentized estimates based on ratio and regression estimators which are smooth functions of vector means.

The present paper focuses on $U-$statistics. There are several reasons to consider U statistics in this context. Firstly, many U statistics are not smooth functions of (multivariate) sample means and, therefore, the results (and techniques) of Booth, Butler and Hall (1994) are not applicable to them. Secondly, U statistics provide good approximations to general symmetric statistics via Hoeffding's decomposition. In this way, results obtained for U statistics can be extended to more general asymptotically linear statistics. We show that the without-replacement bootstrap introduced in Booth, Butler and Hall (1994) provides second-order correct approximation to the distribution function of Studentized $U-$statistics. Similar approximation accuracy is shown to hold also for a version of the without-replacement bootstrap by Bickel and Freedman (1984) and Chao and Lo (1985) but with a randomized resample size. In Section 2 we present the results. Proofs are given in Section 3.

Acknowledgement. I would like to thank Prof. R. Helmers for valuable discussion that motivated present study.

2 Results

Before presenting the results we introduce some notation related to finite population U statistics.

2.1. Preliminaries. Let $\mathcal{X} = \{x_1, \ldots, x_N\}$ denote a finite population. Assume that we want to estimate the population expression $u = \sum_{1 \le i < j \le N} t(x_i, x_j)$, where t is a real valued function symmetric in its arguments (i.e., $t(x, y) = t(y, x)$). The $U-$ statistic

$$\hat{u} = \hat{u}_n(\mathbb{X}) = \frac{\binom{N}{2}}{\binom{n}{2}} \sum_{1 \le i < j \le n} t(X_i, X_j) \tag{1}$$

based on the simple random sample $\mathbb{X} = \{X_1, \ldots, X_n\}$ drawn without replacement from $\mathcal{X}$ is an unbiased estimator of u. $U-$statistics generalize linear estimators. In particular, choosing $t(x, y) = \frac{1}{2\binom{N}{2}}(x + y)$ we obtain the sample mean $\hat{u} = n^{-1}(X_1 + \cdots + X_n)$. For $t = \frac{1}{\binom{N}{2}}|x - y|$ we obtain Gini's mean difference estimator.

In order to make an inference about the unknown population value u we need good approximations to the distribution function of (properly standardized) estimator $\hat{u}$. We shall construct bootstrap approximations to the distribution function of Studentized $U-$statistic $F(x) = \mathbf{P}\{\hat{u} - u \le xS\}$, where

$$S^2 = S^2(\mathbb{X}) = \frac{N - n}{N} \frac{n - 1}{n} \sum_{i=1}^{n} \bigl(\hat{u}_{n-1}(\mathbb{X}_i) - \overline{u}\bigr)^2 \tag{2}$$

is the jackknife estimator of variance of $\hat{u}$. Here $\overline{u} = \frac{1}{n} \sum_{i=1}^{n} \hat{u}_{n-1}(\mathbb{X}_i)$, $\mathbb{X}_i$ denotes the simple random sample of size $n-1$ obtained from $\mathbb{X}$ by removing the observation X_i.

It is convenient to represent $\hat{u}$ by the sum

$$\hat{u} = u + L + Q, \tag{3}$$

where $L = w \sum_{i=1}^{n} g(X_i)$ and $Q = w \sum_{1 \le i < j \le n} \psi(X_i, X_j)$ are uncorrelated. Here

$$g(X_i) = (n - 1) \frac{N - 1}{N - 2} h(X_i) \tag{4}$$

and

$$\psi(X_i, X_j) = t(X_i, X_j) - \mathbf{E}t(X_i, X_j) - \frac{N - 1}{N - 2}(h(X_i) + h(X_j)) \tag{5}$$

are uncorrelated, $h(x) = \mathbf{E}\big(t(X_1, X_2) - \mathbf{E}t(X_1, X_2)|X_1 = x\big)$ and $w = \frac{\binom{N}{2}}{\binom{n}{2}}$. By the Erdős-Rényi (1959) central limit theorem, for large n, the distribution of the linear statistic L can be approximated by the normal distribution. In the case where the linear part L dominates the statistic (this corresponds to the case of non degenerate $U-$statistics in the i.i.d. setup) the normal approximation applies to $\hat{u}$ as well. An improvement upon the normal approximation is provided by an Edgeworth expansion. The one-term Edgeworth expansion of the distribution function $F(x)$ is

$$G(x) = \Phi(x) + \frac{(1-2f)\alpha + (2-f)\alpha\, x^2 + 3\tau^2\,\kappa\,(x^2+1)}{6\tau}\Phi'(x). \qquad (6)$$

Here Φ denotes the standard normal distribution function and Φ' denotes its derivative, $f = n/N$ denotes the sample fraction, $\tau^2 = N\,f\,(1-f)$, and

$$\alpha = \sigma^{-3}\mathbf{E}g^3(X_1), \qquad \kappa = \sigma^{-3}\mathbf{E}\psi(X_1, X_2)g(X_1)g(X_2), \qquad \sigma^2 = \mathbf{E}g^2(X_1). \qquad (7)$$

Let us note, that the decomposition (3) (called Hoeffding's decomposition) was first used in Hoeffding (1948) in the case of independent observations. Using Hoeffding's decomposition Bentkus, Götze and van Zwet (1997) constructed second order aproximations to distribution functions of general asymptotically linear symmetric statistics of independent observations.

For samples drawn without replacement Hoeffding's decomposition was studied in Bloznelis and Götze (2001) and Zhao and Chen (1990). The one-term Edgeworth expansion for finite population Student statistic is shown in Babu and Singh (1985). Edgeworth expansion of finite population $U-$statistics was constructed by Kokic and Weber (1990). The expansion (6) is shown in Bloznelis (2003). For i.i.d. samples the one-term Edgeworth expansion of Studentized $U-$ statistics was shown in Helmers (1991).

2.2. Bootstrap The without replacement bootstrap (Bickel and Freedman (1984), Chao and Lo (1985), Booth, Butler and Hall (1994)) approximates the distribution of an estimator $\hat{\theta}_n(\mathbb{X})$ by the conditional distribution (given $\mathbb{X}$) of $\hat{\theta}_n(\mathbb{X}^*)$, where $\mathbb{X}^*$ is the simple random sample of size n drawn without replacement from an empirical population $\mathcal{X}^*$. Elements of empirical (bootstrap) population are replicates of the observations $X_1, \dots, X_n$. In the case where the population size is a multiple of the sample size, i.e., $N = nk$, put $\mathcal{X}^* = \cup_{j=1}^{k}\mathcal{X}_j^*$, where, for every j, $\mathcal{X}_j^* = \{X_{j1}, \dots, X_{jn}\}$ is a copy of $\mathbb{X}$. For $N = nk + l$ and $0 < l < n$, two ways of generating $\mathcal{X}^*$ are considered in the literature. Booth, Butler and Hall (1994) define the empirical population $\mathcal{X}^* = (\cup_{j=1}^{k}\mathcal{X}_j^*) \cup \{Y_1, \dots, Y_l\}$ of size N by adding elements

$Y_1, \dots, Y_l$ drawn without replacement from $\mathbb{X}$. Bickel and Freedman (1984) and Chao and Lo (1985) define $\mathcal{X}^*$ via randomization: $\mathcal{X}^* = \mathcal{X}^*_{[k]}$ with probability γ and $\mathcal{X}^* = \mathcal{X}^*_{[k+1]}$ with probability $1-\gamma$. Here, for $h = 1, 2, \dots,$ we denote $\mathcal{X}^*_{[h]} = \cup_{j=1}^{h} \mathcal{X}^*_j$. In what follows the two bootstrap procedures are called BBH bootstrap and BFCL bootstrap respectively.

We shall consider the accuracy of BBH and BFCL bootstrap approximations when both the population size and the sample size are large. It is convenient to introduce a sequence of populations $\mathcal{X}^{(\nu)}$ and a sequence of statistics $\hat{u}^{(\nu)}$ defined by (1) and based on simple random samples $\mathbb{X}^{(\nu)} = \{X_1^{(\nu)}, \dots, X_n^{(\nu)}\}$ drawn without replacement from $\mathcal{X}^{(\nu)}$. Here the kernel $t = t^{(\nu)}$ the sample size $n = n^{(\nu)}$ and the population size $N = N^{(\nu)}$ all depend on $\nu = 1, 2, 3, \dots$. In order to keep the notation simple we shall skip the superscript ν whenever this does not cause an ambiguity.

We shall assume that $n, N - n \to \infty$ as $\nu \to \infty$. Furthermore we assume that there exists a constant $C_0 > 0$ such that for every $\nu = 1, 2, \dots$ we have

$$0 < \mathbf{E} t^6(X_1, X_2)/s^6 < C_0, \qquad \text{where} \qquad s^2 = \mathbf{E} h^2(X_1) > 0. \tag{8}$$

2.2.1. Without-replacement BBH bootstrap. Write $N = nk + l$, where $0 \le l < n$. Let $\tilde{\mathbb{X}} = \{\tilde{X}_1, \dots, \tilde{X}_n\}$ be the simple random sample drawn without replacement from the population $\tilde{\mathcal{X}} := \mathcal{X}^*_{[k]} \cup \mathcal{Y}$, where $\mathcal{Y} = \{Y_1, \dots, Y_l\}$ denotes the simple random sample drawn without replacement from $\mathbb{X}$. Introduce the resampling estimator of $\hat{u}$

$$U_n(\tilde{\mathbb{X}}) = \frac{\binom{N}{2}}{\binom{n}{2}} \sum_{\{x,y\} \subset \tilde{\mathbb{X}}} v(x, y), \tag{9}$$

where

$$v(X_i, X_j) = t(X_i, X_j) \mathbb{I}_{\{i \neq j\}}, \quad \text{for} \qquad X_i, X_j \in \mathbb{X}. \tag{10}$$

Note that using the kernel v instead of t we avoid possibly undefined quantities $t(X_i, X_i)$. Write, for short, $\tilde{U} = U_n(\tilde{\mathbb{X}})$. The conditional distribution function

$$F^*(x) = P\{\tilde{U} - E(\tilde{U}|\mathbb{X}, \mathcal{Y}) \le x\, \tilde{S} \big| \mathbb{X}\},$$

is the BBH bootstrap approximation of $F(x)$. Here $\tilde{S}^2 = S^2(\tilde{\mathbb{X}})$ denotes the jackknife estimate (2) of the conditional variance of $\tilde{U}$ given $\mathbb{X}$ and $\mathcal{Y}$.

We show that F^* is the second order correct approximation of the distribution function F as $n, N - n \to \infty$, i.e.,

$$F(x) = F^*(x) + o_P(\tau^{-1}). \tag{11}$$

We outline the proof of (11). Let $\tilde{F}(x) = P\{\tilde{U} - E(\tilde{U}|\mathbb{X}, \mathcal{Y}) \le x\, \tilde{S}\big|\mathbb{X}, \mathcal{Y}\}$ denote the conditional distribution function given $\mathbb{X}$ and $\mathcal{Y}$ and let $\tilde{G}(x)$ denote the one-term Edgeworth expansion of $\tilde{F}(x)$ given by (6). Thus, $\tilde{G}(x)$ is obtained by replacing α and κ in (6) by respective conditional moments $\tilde{\alpha}$ and $\tilde{\kappa}$ of $\tilde{U} = U_n(\tilde{\mathbb{X}})$ given $\mathbb{X}$ and $\mathcal{Y}$. Let $\mathbf{E}^*$ denote the conditional expectation given $\mathbb{X}$. We have $F^*(x) = \mathbf{E}^* \tilde{F}(x)$. Write

$$\tilde{F}(x) - F(x) = r_1 + r_2 + r_3, \tag{12}$$

where $r_1 = \tilde{F}(x) - \tilde{G}(x)$, $r_2 = \tilde{G}(x) - G(x)$ and $r_3 = G(x) - F(x)$. Introduce the event $\mathcal{A} = \{|\tilde{s}^2 - s^2| < s^2/2\}$, where $\tilde{s}^2$ is the BBH bootstrap estimate of s^2 defined in (24) below. Let $\overline{\mathcal{A}}$ denotes the complement event. From (12) we obtain that

$$|F^*(x) - F(x)| \le R_1 + R_2 + R_3 + R_4, \tag{13}$$

where $R_1 = \big|E^*(\tilde{F}(x) - \tilde{G}(x))\mathbb{I}_{\mathcal{A}}\big|$, $R_2 = \big|\mathbf{E}^*(\tilde{G}(x) - G(x))\mathbb{I}_{\mathcal{A}}\big|$, $R_3 = |G(x) - F(x)|$, and $R_4 = E^* \mathbb{I}_{\overline{\mathcal{A}}}$.

Theorem 1 in Bloznelis (2003) provides an explicit estimate for $|G(x) - F(x)|$ which yields the bound $R_3 = o(\tau^{-1})$ under moment conditions on L and Q and asymptotic nonlatticeness condition on the linear part L. In order to show that $R_1 = o_P(\tau^{-1})$ we apply the law of large numbers to the estimate of $|\tilde{F}(x) - \tilde{G}(x)|$ provided by Theorem 1 (ibidem). Finally, the bounds $R_i = o_P(\tau^{-1})$, $i = 2, 4$ follow from Theorem 1 below. Collecting the bounds $R_i = o_P(\tau^{-1})$, for $i = 1, 2, 4$ and $R_3 = o(\tau^{-1})$ in (13) we obtain (11).

Theorem 1. *Assume that (8) holds. Then*

$$\mathbf{E}^*(\tilde{\alpha} - \alpha)\, \mathbb{I}_{\mathcal{A}} = O_P(n^{-1/2}), \quad \mathbf{E}^*(\tilde{\kappa} - \kappa)\, \mathbb{I}_{\mathcal{A}} = O_P(n^{-3/2}), \quad \mathbf{E}^* \mathbb{I}_{\overline{\mathcal{A}}} = O_P(n^{-1}). \tag{14}$$

2.2.2. Without-replacement BFCL bootstrap. Given integers h and m let $\mathbb{X}^\star = \{X_1^\star, \dots, X_m^\star\}$ be the simple random sample drawn without replacement from the population $\mathcal{X}^*_{[h]}$. Introduce the statistic

$$U^\star_{h,m} = U^\star_{h,m}(\mathbb{X}^\star) = \frac{\binom{N_h}{2}}{\binom{m}{2}} \sum_{\{x,y\} \subset \mathbb{X}^\star} v(x, y), \tag{15}$$

and the conditional distribution function

$$F^\star_{h,m}(x) = P\{U^\star_{h,m} - E(U^\star_{h,m}|\mathbb{X}) \le x\, S^\star_{h,m}\big|\mathbb{X}\}. \tag{16}$$

Here $N_h = nh$, the kernel v is defined in (10), and $S^{\star 2}_{h,m}$ denotes the jackknife estimate (2) of the conditional variance of $U^\star_{h,m}$ given $\mathbb{X}$.

BFCL bootstrap approximation $F^\star(x)$ of $F(x)$ is defined as follows. For $N = nk$ put $F^\star(x) = F^\star_{k,n}(x)$. For $N = nk + l$ with $0 < l < n$, put

$$F^\star(x) = \gamma F^\star_1(x) + (1-\gamma)F^\star_2(x), \tag{17}$$

where we denote, for short, $F^\star_1(x) = F^\star_{k,m_1}(x)$ and $F^\star_2(x) = F^\star_{k+1,m_2}(x)$. The integers m_1, m_2 are defined by the inequalities

$$m_1 \le fnk < m_1 + 1, \qquad \text{and} \qquad m_2 - 1 < fn(k+1) \le m_2. \tag{18}$$

The probability γ satisfies the equation

$$\gamma \frac{\sqrt{n}}{\sqrt{m_1}} + (1-\gamma)\frac{\sqrt{n}}{\sqrt{m_2}} = 1. \tag{19}$$

We show that $F^\star$ is the second order correct approximation of the distribution function F as $n, N-n \to \infty$, i.e.,

$$F(x) = F^\star(x) + o_P(\tau^{-1}). \tag{20}$$

We outline the proof of (20). For $j = 1, 2$, let $G^\star_j(x)$ denote the one-term Edgeworth expansion (6) of $F^\star_j(x)$, where α and κ are replaced by respective conditional moments $\alpha^\star_j$ and $\kappa^\star_j$ of $U^\star_j$ given $\mathbb{X}$, and f, τ are replaced by f_j, τ_j. Here we denote $U^\star_1 = U^\star_{k,m_1}$ and $U^\star_2 = U^\star_{k+1,m_2}$ and $f_1 = m_1/nk$, $f_2 = m_2/n(k+1)$, and $\tau_j^2 = m_j(1-f_j)$. We have

$$F^\star(x) - F(x) = \gamma R_{11} + (1-\gamma)R_{12} + R_2 + R_3, \tag{21}$$

where $R_{1.j} = F^\star_j(x) - G^\star_j(x)$, $R_2 = \gamma G^\star_1(x) + (1-\gamma)G^\star_2(x) - G(x)$, and $R_3 = G(x) - F(x)$. In order to show that $R_{1.j} = o_P(\tau^{-1})$ we apply Theorem 1 of Bloznelis (2003) which provides an explicit estimate for $|F^\star_j(x) - G^\star_j(x)|$. The bounds $R_{1.j} = o_P(\tau^{-1})$, $j = 1, 2$, then follow by the law of large numbers applied to this estimate. Furthermore, the bound $R_2 = o_P(\tau^{-1})$ follows from Theorem 2 below. The bound $R_3 = o(\tau^{-1})$ is shown in (13). Collecting these bounds in (21) we obtain (20).

Theorem 2. *Assume that (8) holds. Then, for $j = 1, 2$, we have as $n, N-n \to \infty$*

$$\alpha^\star_j - \alpha = O_P(n^{-1/2}), \qquad \tau_j^2 \kappa^\star_j - \tau^2 \kappa = O_P(n^{-3/2}), \tag{22}$$

$$f_j - f = O(N^{-1}), \qquad \frac{\gamma}{\tau_1} + \frac{1-\gamma}{\tau_2} - \frac{1}{\tau} = O(n^{-1}). \tag{23}$$

Remark. Theorems 1 and 2 yield the bootstrap approximation accuracy bound $O_P(\tau^{-2})$ in (11) and (20) in the case where the one-term Edgeworth expansion is valid with the error $O(\tau^{-2})$. Such error bound can be shown under a Cramér type condition on the linear part L and moment conditions on L and Q, cf. Bloznelis and Götze (2001).

3 Proofs

We give the proof of Theorem 1. The proof of Theorem 2 is much the same.

Denote $s^2 = \mathbf{E}h^2(X_1)$, $a = \mathbf{E}h^3(X_1)$ and $b = \mathbf{E}\psi(X_1, X_2)h(X_1)h(X_2)$. We have $\alpha = a/s^3$ and $\kappa = b/s^3 A$, where $A = (n-1)\frac{N-1}{N-2}$.

Let us define the bootstrap estimates of s^2, a, b and α, κ. Let $\{X_1^*, X_2^*, X_3^*\}$ be the simple random sample of size 3 drawn without replacement from the (random) set $\mathbb{X}$. The naive bootstrap estimates of s^2, a, b are the conditional moments (given $\mathbb{X}$)

$$s_0^2 = \mathbf{E}^* h_0^2(X_1^*), \qquad a_0 = \mathbf{E}^* h_0^3(X_1^*), \qquad b_0 = \mathbf{E}^* \psi_0(X_1^*, X_2^*)h_0(X_1^*)h_0(X_2^*),$$

where $h_0(x) = \mathbf{E}^*(t(X_1^*, X_2^*) - \mathbf{E}^* t(X_1^*, X_2^*)|X_1^* = x)$, and

$$\psi_0(x, y) = t(x, y) - \mathbf{E}^* t(X_1^*, X_2^*) - \frac{n-1}{n-2}(h_0(x) + h_0(y)).$$

Furthermore, $\alpha_0 = a_0/s_0^3$ and $\kappa_0 = b_0/s_0^3 A$ are naive bootstrap estimates of α and κ.

Let $\tilde{\mathbf{E}}$ denote the conditional expectation given $\mathbb{X}$ and $\mathcal{Y}$. Introduce the BBH bootstrap estimates

$$\begin{aligned} &\tilde{s}^2 = \tilde{\mathbf{E}}\tilde{h}^2(\tilde{X}_1), \quad \tilde{a} = \tilde{\mathbf{E}}\tilde{h}^3(\tilde{X}_1), \quad \tilde{b} = \tilde{\mathbf{E}}\tilde{\psi}(\tilde{X}_1, \tilde{X}_2)\tilde{h}(\tilde{X}_1)\tilde{h}(\tilde{X}_2), \\ &\tilde{\alpha} = \tilde{s}^{-3}\tilde{a}, \quad \tilde{\kappa} = A^{-1}\tilde{s}^{-3}\tilde{b}. \end{aligned} \tag{24}$$

Here

$$\begin{aligned} \tilde{h}(x) &= \tilde{\mathbf{E}}\big(v(\tilde{X}_1, \tilde{X}_2) - \tilde{E}v(\tilde{X}_1, \tilde{X}_2)\big|\tilde{X}_1 = x\big), \\ \tilde{\psi}(x, y) &= v(x, y) - \tilde{E}v(\tilde{X}_1, \tilde{X}_2) - \frac{N-1}{N-2}(\tilde{h}(x) + \tilde{h}(y)). \end{aligned}$$

In the proof we show that the bootstrap estimates $\tilde{a}, \tilde{b}, \tilde{s}^2$ are close to the corresponding naive bootstrap estimates a_0, b_0, s_0^2 and that the latter ones are consistent.

Denote $t_0 = \mathbf{E}t(X_1, X_2)$, $\tilde{t} = \mathbf{E}v(\tilde{X}_1, \tilde{X}_2)$ and for $x_i, x_j \in \mathcal{X}$ write

$$v_0(x_i, x_j) = (t(x_i, x_j) - t_0)\mathbb{I}_{\{i \neq j\}}, \qquad \tilde{v}(x_i, x_j) = v(x_i, x_j) - \tilde{t},$$
$$u_0 = \mathbf{E}^* v_0(X_1^*, X_2^*), \qquad \tilde{u} = \tilde{\mathbf{E}}\tilde{v}(\tilde{X}_1, \tilde{X}_2). \tag{25}$$

Observe, that $\mathbf{E}u_0 = \mathbf{E}\tilde{u} = 0$. Furthermore, a calculation shows that

$$\tilde{t} = (1 - \delta)t_0, \qquad \delta = k(N + l - n)/N(N - 1) < 1/n.$$

In the proof we use the fact that values of moments $a_0, b_0, s_0, \alpha_0, \kappa_0$ (respectively $\tilde{a}, \tilde{b}, \tilde{s}, \tilde{\alpha}, \tilde{\kappa}$) remain unchanged if we replace t by v_0 (respectively v by $\tilde{v}$) in the expressions defining these moments. We shall assume that given $\mathbb{X}$ the collections of random variables $\{X_1^*, X_2^*, X_3^*\}$ and $\mathcal{Y} = \{Y_1, \ldots, Y_l\}$ are conditionally independent.

Proof of Theorem 1. We can assume without loss of generality that $s^2 = 1$.

In order to prove that $\mathbf{E}^*\mathbb{I}_{\overline{\mathcal{A}}} = O_P(n^{-1})$ we show that for all $\varepsilon > 0$

$$\mathbf{P}\{|\tilde{s}^2 - 1| > \varepsilon\} = O(n^{-1}). \tag{26}$$

For this purpose we write write $\tilde{s}^2 - 1 = \bar{s}^2 - \tilde{u}^2 - 1$ where

$$\bar{s}^2 = \tilde{\mathbf{E}}\bar{h}^2(\tilde{X}_1), \qquad \bar{h}(x) = \tilde{\mathbf{E}}\big(\tilde{v}(\tilde{X}_1, \tilde{X}_2)\big|\tilde{X}_1 = x\big). \tag{27}$$

Combining Chebyshev's inequality with moment bounds (56) and (57) we obtain (26).

Let us prove that

$$\mathbf{E}^*(\alpha - \tilde{\alpha})\,\mathbb{I}_{\mathcal{A}} = O_P(n^{-1/2}). \tag{28}$$

Firstly, we replace $\tilde{\alpha}$ by $\overline{\alpha} = \overline{a}/\tilde{s}^3$, where $\overline{a} = \tilde{\mathbf{E}}\bar{h}^3(\tilde{X}_1)$. Write

$$\mathbf{E}^*(\alpha - \tilde{\alpha})\,\mathbb{I}_{\mathcal{A}} = \mathbf{E}^*(\alpha - \overline{\alpha})\,\mathbb{I}_{\mathcal{A}} + r,$$

where $r = \mathbf{E}^*(\overline{\alpha} - \tilde{\alpha})\mathbb{I}_{\mathcal{A}}$. We prove that $r = O_P(n^{-1/2})$. It suffices to show that $\mathbf{E}|r| = O(n^{-1/2})$. The latter bound follows from the inequalities

$$|\overline{\alpha} - \tilde{\alpha}|\mathbb{I}_{\mathcal{A}} \le c|\overline{a} - \tilde{a}|\mathbb{I}_{\mathcal{A}} \le c|\tilde{u}^3| + c|\tilde{u}|$$

and the moment inequalities of Lemma 3.2.

Secondly, we show that

$$\mathbf{E}^*(\alpha - \overline{\alpha})\,\mathbb{I}_{\mathcal{A}} = O_P(n^{-1/2}). \tag{29}$$

Note that $|\overline{\alpha} - \overline{a}| \le c|\overline{a}(\tilde{s}^2 - 1)|$, for $|\tilde{s}^2 - 1| < 1/10$, and $\alpha = a$ for $s^2 = 1$. We obtain

$$(\alpha - \overline{\alpha})\,\mathbb{I}_{\mathcal{A}} = (a - \overline{a})\mathbb{I}_{\mathcal{A}} + R_1 = (a - \overline{a}) + R_1 + R_2,$$

where $|R_1| \le c|\overline{a}(\tilde{s}^2 - 1)|$ and $R_2 = (\overline{a} - a)\,\mathbb{I}_{\overline{\mathcal{A}}}$. In order to prove the bounds $\mathbf{E}^* R_i = O_P(n^{-1/2})$ we show that $\mathbf{E}|R_i| = O(n^{-1/2})$, $i = 1, 2$. We apply Cauchy-Schwartz to $\mathbf{E}|R_1|$

$$\begin{aligned}
\mathbf{E}|\overline{a}(\tilde{s}^2 - 1)| \quad &= \mathbf{E}|\overline{a}(\bar{s}^2 - \tilde{u}^2 - 1)| \\
&\le c(\mathbf{E}\overline{a}^2)^{1/2}(\mathbf{E}(\bar{s}^2 - 1)^2)^{1/2} + c(\mathbf{E}\overline{a}^2)^{1/2}(\mathbf{E}\tilde{u}^4)^{1/2} \\
&= O(n^{-1/2})
\end{aligned}$$

In the last step we apply (55), (56), (57). Similarly, by (26), (55), we have

$$\mathbf{E}|R_2| \le |a|\mathbf{P}(\bar{\mathcal{A}}) + (\mathbf{E}\overline{a}^2)^{1/2}\mathbf{P}^{1/2}(\bar{\mathcal{A}}) = O(n^{-1/2}).$$

In order to prove (29) it remains to show that $\mathbf{E}^*\overline{a} - a = O_P(n^{-1/2})$. For this purpose we write $\overline{a} - a = \overline{a} - \overline{a}_0 + \overline{a}_0 - a$, where

$$\overline{a}_0 = \mathbf{E}^* h_*^3(X_1^*), \qquad h_*(x) = \mathbf{E}^*(v_0(X_1^*, X_2^*)|X_1^* = x), \tag{30}$$

and combine the bounds

$$\mathbf{E}^*\overline{a} - \overline{a}_0 = O_P(n^{-1/2}), \qquad \overline{a}_0 - a = O_P(n^{-1/2}). \tag{31}$$

The second bound follows from (44). Let us prove the first bound of (31). For this purpose it is convenient to write the function $\bar{h}$ in the form

$$\bar{h}(x) = \frac{1}{N-1}\big(kS_x + \bar{S}_x + t_0\delta'\big) = \delta_n h_*(x) + \frac{1}{N-1}\bar{\Delta}_x, \tag{32}$$

for $x \in \mathbb{X}$, where we denote $\delta_n = \frac{N}{N-1}\frac{n-1}{n}$,

$$\begin{aligned}
S_x &= \sum_{y\in\mathbb{X}} v_0(x, y), \qquad \bar{S}_x = \sum_{y\in\mathcal{Y}} v_0(x, y), \\
\bar{\Delta}_x &= \Delta_x + t_0\delta', \qquad \Delta_x = \bar{S}_x - \frac{l}{n}S_x
\end{aligned}$$

and where $\delta' = \mathbb{I}_{\{x\notin\mathcal{Y}\}} - k(n-l)/N$ satisfies $|\delta'| \le 1$. We have

$$\overline{a} = \frac{1}{N}\Big(k\sum_{x\in\mathbb{X}} \bar{h}^3(x) + \sum_{x\in\mathcal{Y}} \bar{h}^3(x)\Big) = \frac{1}{N}\big(I_1 + I_2\big),$$

where

$$I_1 = \bigl(k + \frac{l}{n}\bigr) \sum_{x\in\mathbb{X}} \bar{h}^3(x), \qquad I_2 = \sum_{x\in\mathcal{Y}} \bar{h}^3(x) - \frac{l}{n} \sum_{x\in\mathbb{X}} \bar{h}^3(x).$$

In order to prove (31) we shall show that

$$\mathbf{E}^* N^{-1} I_1 - \overline{a}_0 = O_P(n^{-1/2}), \qquad \mathbf{E}^* N^{-1} I_2 = O_P(n^{-1/2}). \tag{33}$$

Let us prove the first bound. It follows from (32) that

$$N^{-1} I_1 = \delta_n^3 \overline{a}_0 + R, \qquad |R| \le c \sum_{i=1}^{3} N^{-i} r_i, \tag{34}$$

where $r_i = \mathbf{E}^* |h_*^{3-i}(X_1^*) \bar{\Delta}_{X_1^*}^i|$. Combining of (58), (55) and Hoelder's inequality gives $\mathbf{E} r_i = O(l^{i/2})$. Therefore, we have $R = O_P(N^{-1/2})$ and (33) follows from (34).

Let us show the second bound of (33). It follows from (32) that $I_2 = \delta_n^3 R_1 + R_2$, where

$$R_1 = \sum_{x\in\mathcal{Y}} \bigl(h_*^3(x) - \mathbf{E}^* h_*^3(Y_1)\bigr), \qquad |R_2| \le c \sum_{i=1}^{3} N^{-i}(r_i' + r_i''),$$

$$r_i' = \sum_{x\in\mathbb{X}} |h_*^{3-i}(x) \bar{\Delta}_x^i|, \qquad r_i'' = \frac{l}{n} \sum_{x\in\mathbb{X}} |h_*^{3-i}(x) \bar{\Delta}_x^i|.$$

It follows, by symmetry and Hoelder's inequality that

$$\mathbf{E} r_i' \le n \mathbf{E} |h_*^{3-i}(X_1^*) \bar{\Delta}_{X_1^*}^i| \le n \bigl(\mathbf{E} h_*^{6-2i}(X_1^*)\bigr)^{1/2} \bigl(\mathbf{E} \bar{\Delta}_{X_1^*}^{2i}\bigr)^{1/2} \le cn\, l^{i/2}.$$

In the last step we applied (55) and (58). The same bound holds for $\mathbf{E} r_i''$ as well. We obtain the bound $N^{-1} \mathbf{E}^* R_2 = O_P(N^{-1/2})$. Furthermore, we have, see (61), $\mathbf{E}^* R_1^2 \le l \mathbf{E}^* \bigl(h_*^3(Y_1) - \mathbf{E}^* h_*^3(Y_1)\bigr)^2$. Therefore, by (55), $\mathbf{E} R_1^2 \le cl$. We obtain $N^{-1} \mathbf{E}^* R_1 = O_P(N^{-1/2})$ thus showing (33). The proof of (28) is complete.

Let us prove the bound $\mathbf{E}^*(\kappa - \tilde{\kappa})\, \mathbb{I}_{\mathcal{A}} = O_P(n^{-3/2})$. In view of the assumption $s^2 = 1$ it suffices to show that

$$\mathbf{E}^*(b - \tilde{b}/\tilde{s}^3)\, \mathbb{I}_{\mathcal{A}} = O_P(n^{-1/2}). \tag{35}$$

Before the proof of (35) we introduce some notation. For $\bar{h}$ and h_* defined (27) and (30) write

$$b' = \tilde{E} v'(\tilde{X}_1, \tilde{X}_2), \qquad v'(x,y) = \tilde{v}(x,y) h_*(x) h_*(y),$$
$$b'' = \mathbf{E}^* v_0(X_1^*, X_2^*) h_*(X_1^*) h_*(X_2^*), \qquad \bar{b} = \tilde{E} \tilde{v}(\tilde{X}_1, \tilde{X}_2) \bar{h}(\tilde{X}_1) \bar{h}(\tilde{X}_2).$$

The proof of (35) is similar to that of (28). Combining the bound, see (56), (57), $\mathbf{E}(\tilde{s}^2-1)^2 = O(n^{-1})$ and the bound, see (26), $\mathbf{P}(\bar{\mathcal{A}}) = O(n^{-1})$ we show that

$$\mathbf{E}^*(b - \tilde{b}/\tilde{s}^3)\mathbb{I}_{\mathcal{A}} = \mathbf{E}^*(b - \tilde{b}) + O_P(n^{-1/2}). \qquad (36)$$

Furthermore, invoking (56) and the identity $\tilde{\mathbf{E}}(\tilde{h}(\tilde{X}_2)|\tilde{X}_1) = \frac{-1}{N-1}\tilde{h}(\tilde{X}_1)$ we obtain

$$\mathbf{E}^*\tilde{b} = \mathbf{E}^*\bar{b} + O_P(n^{-1/2}). \qquad (37)$$

Next we replace $\bar{b}$ by b', that is, we show that

$$\mathbf{E}^*\bar{b} = \mathbf{E}^*b' + O_P(n^{-1/2}). \qquad (38)$$

Note that (38) follows from the identity, see (32),

$$\bar{b} = \delta_n^2 b' + \frac{2\delta_n}{N-1}R_1 + \frac{1}{(N-1)^2}R_2,$$

and the bounds $\mathbf{E}^*R_i = O_P(l^{1/2})$, $i = 1, 2$. Here

$$R_1 = \tilde{\mathbf{E}}\tilde{v}(\tilde{X}_1, \tilde{X}_2)h_*(\tilde{X}_2)\bar{\Delta}_{\tilde{X}_1}, \qquad R_2 = \tilde{\mathbf{E}}\tilde{v}(\tilde{X}_1, \tilde{X}_2)\bar{\Delta}_{\tilde{X}_1}\bar{\Delta}_{\tilde{X}_2}.$$

We shall show the bound $\mathbf{E}^*R_1 = O_P(l^{1/2})$ only. To this aim we prove that $\mathbf{E}\mathbf{E}^*|R_1| \le cl^{1/2}$. Combining Cauchy-Schwartz

$$|R_1| \le (AB)^{1/2}, \qquad A = \tilde{\mathbf{E}}\tilde{v}^2(\tilde{X}_1, \tilde{X}_2)(h_*(\tilde{X}_2))^2, \qquad B = \tilde{\mathbf{E}}\bar{\Delta}^2_{\tilde{X}_1}$$

with the simple bound $\mathbf{E}\mathbf{E}^*A = O(1)$ and inequalities

$$B = \frac{1}{N}\sum_{x\in\tilde{\mathcal{X}}} \bar{\Delta}_x^2 \le N^{-1}(k+1)n\mathbf{E}^*\bar{\Delta}^2_{X_1^*}$$

and, see (58), $\mathbf{E}\mathbf{E}^*\bar{\Delta}^2_{X_1^*} \le cl$ we obtain the inequality $\mathbf{E}\mathbf{E}^*R_1 \le cl^{1/2}$.

Finally, (35) follows from (36), (37), (38) and the following bounds

$$b'' - \mathbf{E}^*b' = O_P(n^{-1/2}), \qquad b - b'' = O_P(n^{-1/2}). \qquad (39)$$

The second bound of (39) follows from (44). Let us prove the first bound.

We have

$$b' = \frac{1}{\binom{N}{2}}\Big(k^2 \sum_{\{x,y\}\subset\mathbb{X}} v'(x,y) + \binom{k}{2}\sum_{x\in\mathbb{X}} v'(x,x) + +k\sum_{x\in\mathbb{X}}\sum_{y\in\mathcal{Y}} v'(x,y) + \sum_{\{x,y\}\subset\mathcal{Y}} v'(x,y)\Big)$$

$$= \frac{1}{\binom{N}{2}}\Big(N' I_1 + kI_2 + (\binom{k}{2} + \frac{kl}{n})I_3 + I_4\Big),$$

$$I_1 = \sum_{\{x,y\}\subset\mathbb{X}} v'(x,y), \qquad I_2 = \sum_{x\in\mathbb{X}}\sum_{y\in\mathcal{Y}} v'(x,y) - \frac{l}{n}\sum_{x\in\mathbb{X}}\sum_{y\in\mathbb{X}} v'(x,y),$$

$$I_3 = \sum_{x\in\mathbb{X}} v'(x,x), \qquad I_4 = \sum_{\{x,y\}\subset\mathcal{Y}} v'(x,y) - \frac{\binom{l}{2}}{\binom{n}{2}}\sum_{\{x,y\}\subset\mathbb{X}} v'(x,y).$$

Here $N' = k^2 + 2k\frac{l}{n} + \frac{\binom{l}{2}}{\binom{n}{2}}$ satisfies $1 \geq \frac{\binom{n}{2}}{\binom{N}{2}}N' \geq 1 - \frac{2}{n}$. Therefore, the first bound of (39) follows from the bounds

$$\frac{1}{\binom{n}{2}}\mathbf{E}^* I_1 = b'' + O_P(n^{-1/2}), \qquad \frac{k^2}{N^2}\mathbf{E}^* I_3 = O_P(n^{-1/2}), \tag{40}$$

$$\frac{k}{N^2}\mathbf{E}^* I_2 = O_P(n^{-1/2}), \qquad \frac{1}{N^2}\mathbf{E}^* I_4 = O_P(n^{-1/2}). \tag{41}$$

The first bound of (40) follows from the identity

$$\frac{1}{\binom{n}{2}}\mathbf{E}^* I_1 = \mathbf{E}^* v'(X_1^*, X_2^*) = b'' + (t_0 - \tilde{t})\mathbf{E}^* h_*(X_1^*)h_*(X_2^*)$$

and the inequality $|\tilde{t} - t_0| \leq cn^{-1}$. The second bound of (40) follows from the identity $\mathbf{E}\mathbf{E}^* I_3 = -n\tilde{t}\mathbf{E}h_*^2(X_1^*)$.

In order to prove (41) we show that

$$\mathbf{E}\mathbf{E}^* I_2 = O(nl^{1/2}), \qquad \mathbf{E}\mathbf{E}^* I_4 = O(l^{3/2}). \tag{42}$$

Given $\mathbb{X}$ let $Z_1, \dots, Z_l$ be the simple random sample drawn without replacement from the population $\{z_i = \sum_{x\in\mathbb{X}} v'(x, X_i),\ X_i \in \mathbb{X}\}$. We have $I_2 = \sum_{i=1}^{l}(Z_i - \mathbf{E}^* Z_i)$ and, see (62), $\mathbf{E}^* I_2^2 \leq cl\mathbf{E}^*(Z_1 - \mathbf{E}^* Z_1)^2$. The first bound of (42) follows from the inequalities

$$(\mathbf{E}\mathbf{E}^* I_2)^2 \leq \mathbf{E}\mathbf{E}^* I_2^2 \leq l\mathbf{E}\mathbf{E}^*(Z_1 - \mathbf{E}^* Z_1)^2 = O(n^2 l).$$

Let us prove the second bound of (42). Given $\mathbb{X}$ the random variable I_4 is a centered $U-$statistic based on the simple random sample $Y_1, \dots, Y_l$. An application of (62) gives $\mathbf{E}^* I_4^2 \leq cl^3\mathbf{E}^*(v'(Y_1, Y_2) - \mathbf{E}^* v'(Y_1, Y_2))^2$. Finally, the inequalities $(\mathbf{E}\mathbf{E}^* I_4)^2 \le \mathbf{E}\mathbf{E}^* I_4^2 = O(l^3)$ show (42). The proof of Theorem 1 is complete.

Lemma 3.1. *Assume that (8) holds. Then as $n, (N-n) \to \infty$ we have*

$$
\begin{aligned}
&s^{-2}s_0^2 - 1 = O_P(n^{-1/2}), \qquad u_0 = O_P(n^{-1/2}),\\
&\alpha_0 - \alpha = O_P(n^{-1/2}), \qquad \kappa_0 - \kappa = O_P(n^{-3/2}).
\end{aligned}
\tag{43}
$$

Proof of Lemma 3.1. Stochastic bounds (43) follow by the law of large numbers. We can assume without loss of generality that $s^2 = 1$. In order to prove (43) it suffices to show that

$$
\begin{aligned}
&s_0^2 - 1 = O_P(n^{-1/2}), \qquad u_0 = O_P(n^{-1/2}),\\
&a_0 - a = O_P(n^{-1/2}), \qquad b_0 - b = O_P(n^{-1/2}).
\end{aligned}
\tag{44}
$$

We give the proof of (44) for s_0^2 and u_0 only. The proof for a_0 and b_0 is much the same. The first two bounds of (44) follow by Chebyshev's inequality from the moment inequalities

$$\mathbf{E}u_0 = 0, \qquad \mathbf{Var}\, u_0 \le c n^{-1}\mathbf{E}t^2(X_1, X_2), \tag{45}$$

$$\bigl|\mathbf{E}s_0^2 - 1\bigr| \le c n^{-1}\mathbf{E}t^2(X_1, X_2), \qquad \mathbf{Var}\, s_0^2 \le c n^{-1}\mathbf{E}t^4(X_1, X_2). \tag{46}$$

We prove (45), (46). The identity $\mathbf{E}u_0 = 0$ is obvious. The second bound of (45) follows from Lemma 3.3 (i). Let us prove (46). For $x \in \mathbb{X}$ write

$$h_0(x) = \frac{1}{n-1}S_x - u_0, \qquad S_x = \sum_{y \in \mathbb{X}\setminus\{x\}} v_0(x, y).$$

We have

$$
\begin{aligned}
s_0^2 \quad &= \mathbf{E}^*\bigl(\frac{1}{n-1}S_{X_1^*} - u_0\bigr)^2 = \mathbf{E}^*\bigl(\frac{1}{n-1}S_{X_1^*}\bigr)^2 - u_0^2\\
&= \frac{1}{n(n-1)^2}\sum_{x\in\mathbb{X}} S_x^2 - u_0^2.
\end{aligned}
\tag{47}
$$

A calculation shows that

$$
\begin{aligned}
\sum_{x\in\mathbb{X}} S_x^2 &= 2\sum_{\{x,y\}\subset\mathbb{X}} v_0^2(x,y) + 2\sum_{x\in\mathbb{X}}\ \sum_{\{y,z\}\subset\mathbb{X}\setminus\{x\}} v_0(x,y)v_0(x,z)\\
&= 2\binom{n}{2}\mathbf{E}^* v_0^2(X_1^*, X_2^*) + 2n\binom{n-1}{2}\mathbf{E}^* v_0(X_1^*, X_2^*)v_0(X_1^*, X_3^*)
\end{aligned}
\tag{48}
$$

From (47) and (48) we obtain

$$s_0^2 = \frac{1}{n-1}\mathbf{E}^* v_0^2(X_1^*, X_2^*) + \frac{n-2}{n-1}\mathbf{E}^* v_0(X_1^*, X_2^*) v_0(X_1^*, X_3^*) - u_0^2. \qquad (49)$$

Therefore, we have, by symmetry, that

$$\mathbf{E}s_0^2 = \frac{1}{n-1}\mathbf{E}v_0^2(X_1, X_2) + \frac{n-2}{n-1}\mathbf{E}v_0(X_1, X_2) v_0(X_1, X_3) - \mathbf{E}u_0^2. \qquad (50)$$

The proof of the following identity is similar to that of (49),

$$s^2 = \frac{1}{N-1}\mathbf{E}v_0^2(X_1, X_2) + \frac{N-2}{N-1}\mathbf{E}v_0(X_1, X_2) v_0(X_1, X_3). \qquad (51)$$

It follows from (50) and (51) that

$$s^2 - \mathbf{E}s_0^2 = \mathbf{E}u_0^2 + R, \qquad \text{where} \qquad |R| \le c n^{-1} \mathbf{E}t^2(X_1, X_2).$$

Finally, invoking (45) we obtain the first bound of (46).

Let us prove the second bound of (46). Denote $\varphi_2(x, y) = v_0^2(x, y) - \mathbf{E}v_0^2(X_1, X_2)$ and

$$\begin{aligned} \varphi_3(x, y, z) \quad &= v_0(x, y) v_0(x, z) + v_0(y, x) v_0(y, z) + v_0(z, x) v_0(z, y) \\ &-3\mathbf{E}v_0(X_1, X_2) v_0(X_1, X_3), \\ \varphi_4(x, y, z, w) \quad &= v_0(x, y) v_0(z, w) + v_0(x, z) v_0(y, w) + v_0(y, z) v_0(x, w) \\ &-3\mathbf{E}v_0(X_1, X_2) v_0(X_3, X_4). \end{aligned}$$

It follows from (49) and (50) that

$$s_0^2 - \mathbf{E}s_0^2 = \frac{1}{n-1} T_2 + \frac{1}{9}\frac{n-2}{n-1} T_3 - T', \qquad (52)$$

where

$$T_2 = \frac{1}{\binom{n}{2}} \sum_{\{x,y\} \subset \mathbb{X}} \varphi_2(x, y), \qquad T_3 = \frac{1}{\binom{n}{3}} \sum_{\{x,y,z\} \subset \mathbb{X}} \varphi_3(x, y, z).$$

and $T' = u_0^2 - \mathbf{E}u_0^2$. A calculation shows that

$$\begin{aligned} T' \quad &= \frac{1}{\binom{n}{2}} T_2 + \frac{4}{3}\frac{n-2}{n(n-1)} T_3 + \frac{1}{3}\frac{(n-2)(n-3)}{n(n-1)} T_4, \qquad (53) \\ T_4 \quad &= \frac{1}{\binom{n}{4}} \sum_{\{x,y,z,w\} \subset \mathbb{X}} \varphi_4(x, y, z, w). \end{aligned}$$

It follows from (52) and (53) that

$$\mathbf{Var}\, s_0^2 \le cn^{-2}\mathbf{E}T_2^2 + c\mathbf{E}T_3^2 + c\mathbf{E}T_4^2. \tag{54}$$

Lemma 3.3 (i) shows that $\mathbf{E}T_i^2 \le cn^{-1}\mathbf{E}\varphi_i^2(X_1,\dots,X_i)$. Since $\mathbf{E}\varphi_i^2(X_1,\dots,X_i)$ $\le c\mathbf{E}t^4(X_1,X_2)$ we obtain bound (46) from (54). Proof of Lemma 3.1 is complete.

Lemma 3.2. *Assume that (8) holds with the constant $C_0 > 0$. Suppose that $s^2 = 1$. Then there exists a number $c > 0$ depending only on C_0 such that*

$$\mathbf{E}\bar{h}^6(\tilde{X}_1) < c, \qquad \mathbf{E}h_*^6(X_1^*) < c, \tag{55}$$

$$\mathbf{E}\tilde{u}^2 \le c\,n^{-1}, \qquad \mathbf{E}\tilde{u}^4 \le c\,n^{-1}. \tag{56}$$

$$|\mathbf{E}\bar{s}^2 - 1| \le c\,n^{-1}, \qquad \mathbf{Var}\,\bar{s}^2 \le c\,n^{-1}. \tag{57}$$

For $i \ge 3$ there exists a constant c_i such that for any $x \in \mathbb{X}$ we have

$$\mathbf{E}^*|\Delta_x|^i \le c_i\, l^{i/2}\mathbf{E}^*\big(|v_0(X_1^*,X_2^*)|^i \big| X_1^* = x\big). \tag{58}$$

Recall that $\tilde{u}, \bar{s}^2, \bar{h}, h_*$ and Δ_x are defined in (25), (27), (30) and (32).

Proof. Bounds (55) follow from (8).

Let us prove (56). We shall show the first bound only. The proof of the second bound is much the same. A calculation shows that

$$\tilde{u} = \frac{1}{\binom{N}{2}} \sum_{\{x,y\}\subset\tilde{\mathcal{X}}} \tilde{v}(x,y) = \frac{k^2}{\binom{N}{2}}T_1 + \frac{2k+1}{\binom{N}{2}}T_2 + \frac{k}{\binom{N}{2}}T_3,$$

$$T_1 = \sum_{\{x,y\}\subset\mathbb{X}} v_0(x,y), \qquad T_2 = \sum_{\{x,y\}\subset\mathcal{Y}} v_0(x,y), \qquad T_3 = \sum_{x\in\mathbb{X}\setminus\mathcal{Y}}\sum_{y\in\mathcal{Y}} v_0(x,y).$$

Moment inequalities of Lemma 3.3 imply the bounds $\mathbf{E}T_i^2 \le cn^3\mathbf{E}v_0^2(X_1,X_2)$, $i = 1,2,3$. We obtain

$$\mathbf{E}\tilde{u}^2 \le c(k^4 + (2k+1)^2 + k^2)\frac{n^3}{N^4}\mathbf{E}v_0^2(X_1,X_2) \le c\frac{1}{n}\mathbf{E}v_0^2(X_1,X_2)$$

thus showing the first bound of (56).

Let us prove (57). Using the notation (32) we write

$$\bar{s}^2 = (N-1)^{-2}\Big((t_0\delta')^2 + 2t_0\delta'\tilde{\mathbf{E}}S + \tilde{\mathbf{E}}S^2\Big),$$

where $S = kS_{\tilde{X}_1} + \bar{S}_{\tilde{X}_1}$. We have $\mathbf{E}S = 0$. A calculation shows that

$$\mathbf{E}S^2 = N^2\big(\mathbf{E}v_0(X_1, X_2)v_0(X_1, X_3) + r\big), \qquad \text{where} \qquad |r| \le \frac{c}{n}\mathbf{E}v_0^2(X_1, X_2).$$

Combining this bound with (51) we obtain $\mathbf{E}\bar{s}^2 = 1 + O(n^{-1})$.

In order to prove that $\mathbf{Var}\,\bar{s}^2 = O(n^{-1})$ we show that $\mathbf{Var}\,(N^{-2}\tilde{\mathbf{E}}S^2) = O(n^{-1})$ and $\mathbf{Var}\,(N^{-2}\tilde{\mathbf{E}}S) = O(n^{-1})$. We sketch the proof of the first bound. The proof of the second bound is much simpler. Write

$$\begin{aligned} N(\tilde{\mathbf{E}}S^2) \quad &= k\sum_{x\in\mathbb{X}}\big(kS_x + \bar{S}_x\big)^2 + \sum_{x\in\mathcal{Y}}\big(kS_x + \bar{S}_x\big)^2 \\ &= k^3\sum_{x\in\mathbb{X}} S_x^2 + 2k^2\sum_{x\in\mathbb{X}} S_x\bar{S}_x + k\sum_{x\in\mathbb{X}}\bar{S}_x^2 \\ &+k^2\sum_{x\in\mathcal{Y}} S_x^2 + 2k\sum_{x\in\mathcal{Y}} S_x\bar{S}_x + \sum_{x\in\mathcal{Y}}\bar{S}_x^2. \end{aligned} \tag{59}$$

Given nonintersecting sets $A, B \subset [n]$ denote

$$U_1(A,B) = \sum_{x\in A}\sum_{y\in B} t^2(x,y), \qquad U_i(A,B) = \sum_{\{x,y\}\subset A}\sum_{z\in B}\varphi_i(x,y;z),$$
$$\varphi_2(x,y;z) = t(x,y)\big(t(x,z) + t(y,z)\big), \qquad \varphi_3(x,y;z) = t(x,z)t(y,z).$$

Write $\tilde{S}_x = \sum_{y\in\mathcal{Z}} v(x,y)$, where $\mathcal{Z} = \mathbb{X}\setminus\mathcal{Y}$. We split $S_x = \tilde{S}_x + \bar{S}_x$ in every (but not the first and last) summand of (59). It follows from the identities

$$\sum_{x\in\mathcal{Z}}\tilde{S}_x\bar{S}_x = U_2(\mathcal{Z},\mathcal{Y}), \qquad \sum_{x\in\mathcal{Y}}\tilde{S}_x\bar{S}_x = U_2(\mathcal{Y},\mathcal{Z}),$$
$$\sum_{x\in\mathcal{Y}}\tilde{S}_x^2 = U_1(\mathcal{Y},\mathcal{Z}) + 2U_3(\mathcal{Z},\mathcal{Y}), \qquad \sum_{x\in\mathcal{Z}}\bar{S}_x^2 = U_1(\mathcal{Z},\mathcal{Y}) + 2U_3(\mathcal{Y},\mathcal{Z})$$

that

$$\begin{aligned} N(\tilde{\mathbf{E}}S^2) \quad &= k^3\sum_{x\in\mathbb{X}} S_x^2 + (3k^2 + 3k + 1)\sum_{x\in\mathcal{Y}}\bar{S}_x^2 \\ &+(2k^2 + k)U_1(\mathcal{Z},\mathcal{Y}) + k^2U_1(\mathcal{Y},\mathcal{Z}) \\ &+2k^2U_2(\mathcal{Z},\mathcal{Y}) + (4k^2 + 2k)U_2(\mathcal{Y},\mathcal{Z}) \\ &+(4k^2 + 2k)U_3(\mathcal{Y},\mathcal{Z}) + 2k^2U_3(\mathcal{Z},\mathcal{Y}). \end{aligned} \tag{60}$$

Finally, we construct upper bounds for the variances of various summands of (60) separately. To the first two summands we apply (54). To the next two summands we apply Lemma 3.3(ii). To the remaining summands we apply Lemma 3.3 (iii).

Let us show (58). We shall apply the finite-population version of Rosenthal's inequality. Let $Z_1, \dots, Z_k$ (respectively $Z_1', \dots, Z_k'$) denote a sample drawn without replacement (respectively with replacement) from the finite population $\{z_1, \dots, z_K\}$. A result of Hoeffding [19] implies that $\mathbf{E}(Z_1 + \dots + Z_k)^i \le \mathbf{E}(Z_1' + \dots + Z_k')^i$ for $i \ge 2$. Assume that $\mathbf{E}Z_1' = 0$. Then Rosenthal's inequality implies $\mathbf{E}(Z_1' + \dots + Z_k')^i \le c_i k^{i/2} \mathbf{E}|Z_1'|^i$. Since Z_1 and Z_1' have the same (uniform) distribution, we obtain the inequality

$$\mathbf{E}(Z_1 + \dots + Z_k)^i \le c_i k^{i/2} \mathbf{E}|Z_1|^i. \tag{61}$$

We apply this inequality to random variables $Z_i = v_0(x, Y_i) - \mathbf{E}^* v_0(x, Y_i)$, $Y_i \in \mathcal{Y}$ and obtain

$$\mathbf{E}^*|\Delta_x|^i \le c_i l^{i/2} \mathbf{E}^*|Z_1|^i \le \bar{c}_i l^{i/2} \mathbf{E}^*\bigl(|v_0(X_1^*, X_2^*)|^i \big| X_1^* = x\bigr).$$

We have shown (58) thus completing the proof of Lemma 3.2.

Auxiliary moment inequalities are collected in Lemma 3.3.

Lemma 3.3. *Let $N > n \ge k \ge 1$. Assume that φ_k is a real symmetric function defined on $\mathcal{X}^k$. Suppose that $\mathbf{E}\varphi_k(X_1, \dots, X_k) = 0$.*

(i) There exists a constant $c_k > 0$ depending on k only such that

$$\mathbf{E}\Bigl(\sum_{\{y_1,\dots,y_k\}\subset \mathbb{X}} \varphi_k(y_1, \dots, y_k)\Bigr)^2 \le c_k n^{2k-1}\bigl(1 - \frac{n}{N}\bigr)\mathbf{E}\varphi_k^2(X_1, \dots, X_k). \tag{62}$$

(ii) Let $A, B \subset \{1, \dots, n\}$ are non-empty non-intersecting subsets. For positive integers $r + t = k$ let

$$T = \sum_{\{i_1,\dots,i_r\}\subset A} \sum_{\{j_1,\dots,j_t\}\subset B} \varphi_k(X_{i_1}, \dots, X_{i_r}, X_{j_1} \dots, X_{j_t}).$$

There exists a constant $c_{r,t} > 0$ depending on r and t such that

$$\mathbf{E}T^2 \le c_{r,t}|A|^{2r}|B|^{2t}\Bigl(\frac{1}{|A|} + \frac{1}{|B|}\Bigr)\mathbf{E}\varphi_k^2(X_1, \dots, X_k). \tag{63}$$

(iii) Let φ be a real function defined on $\mathcal{X}^3$ such that $\varphi(x, y; z) = \varphi(y, x; z)$. Assume that $\mathbf{E}\varphi(X_1, X_2; X_3) = 0$. For non-intersecting subsets $A, B \subset [n]$ the random variable $T = \sum_{i \in B} \sum_{\{j,k\}\subset A} \varphi(X_j, X_k; X_i)$ satisfies

$$\mathbf{E}T^2 \le c|A|^3|B|\bigl(|A| + |B|\bigr)\mathbf{E}\varphi^2(X_1, X_2; X_3). \tag{64}$$

Proofs of these moment inequalities are technical and routine. They can be found in the extended version of the paper [8].

4 Proof of Lemma 3.3

Proof. We prove (62) for $k = 3$. The proof an arbitrary k is similar. Denote $M = \mathbf{E}\varphi^2(X_1, X_2, X_3)$. Let us write Hoeffding's decomposition of $T = \sum_{\{x,y,z\}\subset\mathbb{X}} \varphi_3(x, y, z)$, see, e.g., Bloznelis and Götze (2001),

$$T = \binom{n-1}{2} \sum_{x\in\mathbb{X}} g_1(x) + \binom{n-2}{1} \sum_{\{x,y\}\in\mathbb{X}} g_2(x, y) + \sum_{\{x,y,z\}\in\mathbb{X}} g_3(x, y, z), \quad (65)$$

where

$$g_1(x) = \frac{N-1}{N-3} h_\varphi(x), \qquad h_\varphi(x) = \mathbf{E}(\varphi_3(X_1, X_2, X_3)|X_1 = x),$$
$$g_2(x, y) = \frac{N-2}{N-4}\Big(\mathbf{E}\big(\varphi_3(X_1, X_2, X_3)|X_1 = x, X_2 = y\big) - \frac{N-1}{N-2}\big(h_\varphi(x) + h_\varphi(y)\big)\Big),$$
$$g_3(x_1, x_2, x_3) = \varphi_3(x_1, x_2, x_3) - \sum_{i=1}^{3} g_1(x_i) - \sum_{1\le i<j\le 3} g_2(x_i, x_j).$$

It follows from formula (2.6) of Bloznelis and Götze (2001) that

$$\mathbf{E}T^2 = \binom{n-1}{2}^2 \frac{\binom{n}{1}\binom{N-n}{1}}{\binom{N-1}{1}} v_1^2 + \binom{n-2}{1}^2 \frac{\binom{n}{2}\binom{N-n}{2}}{\binom{N-2}{2}} v_2^2 + \frac{\binom{n}{3}\binom{N-n}{3}}{\binom{N-3}{3}} v_3^2,$$

where we denote $v_i^2 = \mathbf{E}g_i^2(X_1, \ldots, X_i)$. The simple inequalities $v_i^2 \le c_i\mathbf{E}\varphi_3^2(X_1, X_2, X_3)$, for $i = 1, 2, 3$, complete the proof of (62).

Proof of (63). Let us prove (63) for $r = t = 1$. Denote $M = \mathbf{E}\varphi^2(X_1, X_2)$. Write Hoeffding's decomposition $\varphi(X_1, X_2) = g_1(X_1) + g_1(X_2) + g_2(X_1, X_2)$. Here

$$g_1(x) = \frac{N-1}{N-2}\mathbf{E}(\varphi_2(X_1, X_2)|X_1 = x), \quad g_2(x, y) = \varphi_2(x, y) - g_1(x) - g_2(y).$$

We have $T = T_1 + T_2$, where the random variables

$$T_1 = |B| \sum_{i\in A} g_1(X_i) + |A| \sum_{j\in B} g_1(X_j), \qquad T_2 = \sum_{i\in A}\sum_{j\in B} g_2(X_i, X_j)$$

are uncorrelated. Therefore, $\mathbf{E}T^2 = \mathbf{E}T_1^2 + \mathbf{E}T_2^2$. A simple calculation shows that

$$\begin{aligned}\mathbf{E}T_1^2 \quad &= \Big(|A|^2 \frac{|B|(N-|B|)}{N-1} + |B|^2 \frac{|A|(N-|A|)}{N-1}\Big)\mathbf{E}g_1^2(X_1) \\ &+ |A|^2|B|^2 \mathbf{E}g_1(X_1)g_1(X_2) \\ &\le c|A|^2|B|^2\Big(\frac{1}{|A|} + \frac{1}{|B|}\Big)\mathbf{E}\varphi_2^2(X_1, X_2).\end{aligned}$$

In order to evaluate $\mathbf{E}T_2^2$ write,

$$T_2^2 \quad = \sum_{i\in A}\sum_{j\in B} g_2(X_i, X_j)\Big(g_2(X_i, X_j) + S_i + S_j + S_{ij}\Big), \tag{66}$$

$$S_i \quad = \sum_{r\in A\setminus\{i\}} g_2(X_r, X_j), \qquad S_j = \sum_{r\in B\setminus\{j\}} g_2(X_i, X_r),$$

$$S_{ij} \quad = \sum_{r\in A\setminus\{i\}} \sum_{t\in B\setminus\{j\}} g_2(X_r, X_t). \tag{67}$$

We have, by symmetry,

$$\begin{aligned}\mathbf{E}T_2^2 \quad &= |A||B|\Big(\sigma_{2.0} + (|A| + |B| - 2)\sigma_{2.1} + (|A|-1)(|B|-1)\sigma_{2.2}\Big), \\ \sigma_{2.0} \quad &= \mathbf{E}g_2^2(X_1, X_2), \qquad \sigma_{2.1} = \mathbf{E}g_2(X_1, X_2)g_2(X_1, X_3), \\ \sigma_{2.2} \quad &= \mathbf{E}g_2(X_1, X_2)g_2(X_3, X_4).\end{aligned}$$

Invoking the identities

$$\sigma_{2.1} = -\frac{1}{N-2}\sigma_{2.0}, \qquad \sigma_{2.2} = \frac{1}{N-2}\frac{1}{N-3}\sigma_{2.0},$$

and using the inequality $\sigma_{2.0} \le c\mathbf{E}\varphi_2^2(X_1, X_2)$ we obtain

$$\mathbf{E}T_2^2 \le c|A||B|\mathbf{E}\varphi_2^2(X_1, X_2).$$

This bound in combination with the bound for $\mathbf{E}T_1^2$ gives (63).

Let us prove (63) for arbitrary r and t. A function $g : \mathcal{X}^s \to R$ is called symmetric if it is invariant under permutation of its s arguments. For a symmetric function g and a subset $D = \{i_1, \ldots, i_s\} \subset [n]$ we write $g(D) = g(X_{i_1}, \ldots, X_{i_s})$. Let A_r (respectively B_t) denote r-subset of A (respectively t-subset of B). Using these notation we can write

$$T = \sum_{A_r\subset A} \sum_{B_t\subset B} \varphi_k(A_r \cup B_t).$$

Given subset $D \subset A$ of size $r+t$ we write Hoeffding's decomposition of $\varphi_k(D)$, see, e.g., Bloznelis and Götze (2001),

$$\varphi_k(D) = \sum_{s=1}^{k} \sum_{D_s \subset D} g_s(D_s).$$

Here D_s denotes a subset of D of size s. The function $g_s(D_s)$ is a linear combination of the conditional expectations $\mathbf{E}(\varphi_k(D)|X_{i_1} \ldots, X_{i_u})$, for $\{i_1, \ldots, i_u\} \subset D_s$ and $u = 1, 2, \ldots s$, such that

$$\mathbf{E} g_s(D_s) g_u(D_u) = 0, \qquad \text{for} \qquad u \neq s. \tag{68}$$

We have

$$\mathbf{E} g_s^2(D_s) \le c(k,s)\sigma_\varphi^2, \qquad \text{where} \qquad \sigma_\varphi^2 := \mathbf{E}\varphi_k^2(X_1, \ldots, X_k). \tag{69}$$

We replace $\varphi_k(A_r \cup B_t)$ by its Hoeffding's decomposition and obtain

$$\begin{aligned} T &= \sum_{A_r \subset A} \sum_{B_t \subset B} \sum_{s=1}^{k} \sum_{u+v=s} \sum_{A_u \subset A_r} \sum_{B_v \subset B_t} g_s(A_u \cup B_t) = \sum_{s=1}^{k} T_s, \\ T_s &= \sum_{u+v=s} \binom{|A|-u}{r-u} \binom{|B|-v}{t-v} \sum_{A_u \subset A} \sum_{B_v \subset B} g_s(A_u \cup B_v). \end{aligned}$$

Here the sum $\sum_{u+v=s}$ is taken over all ordered pairs (u, v) of non-negative integers satisfying $u + v = s$. Furthermore, for $v = 0$ (respectively $u = 0$) we put $\sum_{A_u} \sum_{B_v} = \sum_{A_u}$ (respectively $\sum_{A_u} \sum_{B_v} = \sum_{B_v}$). Using the key property (68) of Hoeffding's decomposition we obtain $\mathbf{E}T^2 = \sum_s \mathbf{E}T_s^2$. In what follows we construct upper bounds for $\mathbf{E}T_s^2$

$$\mathbf{E}T_s^2 \le c_{r,t}|A|^{2r}|B|^{2t}\left(\frac{1}{|A|} + \frac{1}{|B|}\right)\sigma_\varphi^2. \tag{70}$$

These bounds imply (63).

For $s = 1$ we have

$$T_1 = \binom{|B|}{t}\binom{|A|-1}{r-1} \sum_{i \in A} g_1(X_i) + \binom{|A|}{r}\binom{|B|-1}{t-1} \sum_{j \in B} g_1(X_j).$$

Invoking the simple inequality $(b_1 + \cdots + b_m)^2 \le m(b_1^2 + \cdots + b_m^2)$ (take $m = 2$) and using the inequalities $\mathbf{E}(\sum_{i \in A} g_1(X_i))^2 \le |A|\mathbf{E}g_1^2(A_1) \le |A|\sigma_\varphi^2$ we obtain (70).

For $s > 1$ write

$$\begin{aligned} \mathbf{E}T_s^2 &\leq 2(s+1) \sum_{u+v=s} \binom{|A|-u}{r-u}^2 \binom{|B|-v}{t-v}^2 \mathbf{E}T_{uv}^2, \qquad (71) \\ T_{uv} &= \sum_{A_u \subset A} \sum_{B_v \subset B} g_s(A_u \cup B_v). \end{aligned}$$

Furthermore, by symmetry, we have

$$\mathbf{E}T_{uv}^2 = \binom{|A|}{u}\binom{|B|}{v} \mathbf{E}g_s(A_u^0 \cup B_v^0) T_{uv} \qquad (72)$$

where A_u^0 and B_v^0 are fixed subsets of A and B of sizes u and v respectively. A simple calculation shows that

$$\begin{aligned} \mathbf{E}g_s(A_u^0 \cup B_v^0)T_{uv} &= \sum_{i=0}^{u}\sum_{j=0}^{v} \sum_{\substack{A_u \subset A \\ |A_u \cap A_u^0|=i}} \sum_{\substack{B_v \subset B \\ |B_v \cap B_v^0|=j}} \mathbf{E}g_s(A_u^0 \cup B_v^0) g_s(A_u \cup B_v) \\ &= \sum_{i=0}^{u}\sum_{j=0}^{v} \binom{|A|-u}{u-i}\binom{|B|-v}{v-j}\binom{u}{i}\binom{v}{j} a_{s,i+j}. \end{aligned}$$

Here for $|\{i_1,\ldots,i_s\} \cap \{j_1,\ldots,j_s\}| = u$ we denote

$$a_{s,u} = \mathbf{E}g_s(X_{i_1},\ldots,X_{i_s})g_s(X_{j_1},\ldots,X_{j_s}).$$

Invoking the identity, see (5.9) in Bloznelis and Götze (2001),

$$a_{s,u} = \frac{(-1)^{s-u}}{\binom{N-s}{s-u}} \mathbf{E}g_s^2(A_s),$$

we obtain

$$\mathbf{E}g_s(A_u^0 \cup B_v^0)T_{uv} \leq c_{r,t} \sum_{i=0}^{u}\sum_{j=0}^{v} \frac{|A|^{u-i}|B|^{v-j}}{N^{s-(i+j)}} \sigma_\varphi^2 \leq c_{r,t}\sigma_\varphi^2.$$

In the last step we used the fact that $|A|, |B| < N$ and $s + v = s$. It follows from (72) that

$$\mathbf{E}T_{uv}^2 \leq c_{r,t}|A|^u|B|^v\sigma_\varphi^2.$$

Finally, the substitution of this inequality in (71) gives the bound (70).

Proof of (64). Introduce the conditional Hoeffding's decomposition of $T_i = \sum_{\{j,k\}\subset A} \varphi(X_j, X_k; X_i)$ given X_i, for $i \notin A$. Let $E_i(\cdots) = \mathbf{E}(\cdots|X_i)$ denote the conditional expectation given X_i. Denote

$$m(X_i) = \mathbf{E}_i\varphi(X_j, X_k; X_i), \qquad \overline{\varphi}(x, y; z) = \varphi(x, y; z) - m(z)$$
$$g_1(x; z) = \frac{N-2}{N-3}\mathbf{E}\big(\overline{\varphi}(X_j, X_k, X_i)\big|X_i = z, X_j = x\big)$$

and write

$$\overline{\varphi}(x, y; z) = g_1(x; z) + g_1(y; z) + g_2(x, y; z).$$

A simple calculation shows that, for distinct i, j, k and arbitrary $r \neq i$,

$$\mathbf{E}\big(g_2(X_j, X_k; X_i)\big|X_r, X_i\big) = 0. \tag{73}$$

It follows from (73) that

$$\mathbf{E}_i g_2(X_j, X_k; X_i) g_1(X_r; X_i) = 0. \tag{74}$$

We have the conditional decomposition

$$\begin{aligned} T_i &= \binom{|A|}{2} m(X_i) + \overline{T}_i, \qquad \overline{T}_i = (|A| - 1)U_1 + U_2, \\ U_1 &= \sum_{j\in A} g_1(X_j; X_i), \qquad U_2 = \sum_{\{j,k\}\subset A} g_2(X_j, X_k; X_i). \end{aligned} \tag{75}$$

It follows from (73), (74) and the variance formula (2.6) of Bloznelis and Götze (2001) that

$$\mathbf{E}_i\overline{T}_i^2 = (|A| - 1)^2\mathbf{E}_i U_1^2 + \mathbf{E}_i U_2^2 \le c|A|^3\mathbf{E}_i\overline{\varphi}^2(X_1, X_2; X_3). \tag{76}$$

From (75) we obtain

$$T = \binom{|A|}{2} \sum_{i\in B} m(X_i) + \sum_{i\in B} \overline{T}_i =: A_1 + A_2.$$

Using the simple inequality $(b_1 + \cdots + b_s)^2 \le s(b_1^2 + \cdots + b_s^2)$ we obtain, by symmetry,

$$\mathbf{E}T^2 \le 2\mathbf{E}A_1^2 + 2\mathbf{E}A_2^2, \qquad \mathbf{E}A_2^2 \le |B|^2\mathbf{E}\overline{T}_1^2 = |B|^2\mathbf{E}(\mathbf{E}_1\, \overline{T}_1^2).$$

Note that $\mathbf{E}A_1^2 \le |A|^4|B|\mathbf{E}^2 m(X_1)$. This inequality combined with bound (76) and the simple inequalities

$$\mathbf{E}\mathbf{E}_i\overline{\varphi}^2(X_1, X_2; X_3) \le \mathbf{E}\varphi^2(X_1, X_2; X_3), \qquad \mathbf{E}^2 m(X_1) \le \mathbf{E}\varphi^2(X_1, X_2; X_3)$$

implies (64).

References

[1] Babu, G.J., and Singh, K. (1983). Inference on means using bootstrap. *Annals of Statistics,* **11**, 999-1003.
[2] Babu, G.J., and Singh, K. (1984). On one-term Edgeworth correction by Efron's bootstrap. *Sankhya Ser. A.*, **46**, 219-232.
[3] Babu, G.J., and Singh, K. (1985). Edgeworth Expansions for Sampling Without Replacement from Finite Populations. *Journal of Multivariate Analysis*, **17**, 261–278.
[4] Bentkus, V., Götze, F., and van Zwet, W. R. (1997). An Edgeworth Expansion for Symmetric Statistics. *Annals of Statistics,* **25**, 851–896.
[5] Bhattacharya, R. N., and Qumsiyeh, M. (1989). Second Order and L^p-Comparisons Between the Bootstrap and Empirical Edgeworth Expansion Methodologies. *Annals of Statistics,* **17**, 160–169.
[6] Bickel , P.J., and Freedman, D.A. (1984). Asymptotic Normality and the Bootstrap in Stratified Sampling. *Annals of Statistics,* **12**, 470–482.
[7] Bloznelis, M. (2003). An Edgeworth Expansion for Studentized Finite Population Statistics. *Acta Applicandae Mathematicae,* **78**, 51-60.
[8] Bloznelis, M., and Götze, F. (2001). Orthogonal Decomposition of Finite Population Statistic and Its Applications to Distributional Asymptotics. *Annals of Statistics,* **29**, 899–917.
[9] Booth, J.G., Butler, R.W. and Hall P. (1994). Bootstrap methods for finite populations. *J. Amer. Statist. Assoc.* **89** 1282–1289
[10] Chao, M. T. and Lo, S. H. (1985). A bootstrap method for finite population. *Sankhya Ser. A* **47** 399–405.
[11] Chen, J. and Sitter, R. R. (1993). Edgeworth expansion and the bootstrap for stratified sampling without replacement from a finite population. *Canad. J. Statist.*, **21** 347–357.
[12] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics,* **7**, 1-26.

[13] Erdős, P. and Rényi, A. (1959). On the central limit theorem for samples from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci.* **4** 49–61.
[14] Hall, P. (1992). *The bootstrap and Edgeworth expansion.* Springer, New York.
[15] Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *Ann. Statist.* **16** 927–985.
[16] Helmers, R. (1991). On the Edgeworth expansion and the bootstrap approximation for a Studentized U-statistic. *Ann. Statist.* **19** 470–484.

[17] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293-325.
[18] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Stat. Assoc.* **58** 13–30.
[19] Kokic, P.N. and Weber, N.C. (1990). An Edgeworth expansion for U-statistics based on samples from finite populations. *Ann. Probab.* **18** 390–404.
[20] Singh, K., (1981). On the Asymptotic Accuracy of Efron's Bootstrap. *Annals of Statistics,* **9**, 1187-1195.
[21] Sitter, R.R. (1992). A resampling procedure for complex survey data. *J. Amer. Statist. Assoc.* **87** 755–765.
[22] Zhao, L. C. and Chen, X. R. (1990). Normal approximation for finite-population U-statistics. *Acta Math. Appl. Sinica* **6** 263–272.