EMPIRICAL EDGEWORTH EXPANSION
FOR FINITE POPULATION STATISTICS. I

M. Bloznelis

April 2000

ABSTRACT. For symmetric asymptotically linear statistics based on simple random samples, we construct the one-term empirical Edgeworth expansion, where the moments defining the true Edgeworth expansion are replaced by their jackknife estimators. In order to establish the validity of the empirical Edgeworth expansion (in probability), we prove the consistency of the jackknife estimators.

1. INTRODUCTION

Given a set $\mathcal{X} = \{x_1, \ldots, x_N\}$, let $\mathbb{X} = \{X_1, \ldots, X_n\}$ denote a simple random sample of size $n < N$ drawn without replacement from $\mathcal{X}$. That is, for every n-subset $\mathcal{B} \subset \mathcal{X}$, we have $\mathbf{P}\{\mathbb{X} = \mathcal{B}\} = \binom{N}{n}^{-1}$. Let φ be a real function defined on n-subsets of $\mathcal{X}$. We write $\varphi(X_1, \ldots, X_n)$ to denote the value of φ at a (random) subset $\mathbb{X}$. The random variable $T = \varphi(X_1, \ldots, X_n)$ is called a symmetric finite population statistic. "Symmetric" refers to the fact that T is invariant under permutations of the sample.

Bloznelis and Götze [10], [11] constructed the one-term Edgeworth expansion

$$G(x) = \Phi(x) - \frac{(q-p)\alpha + 3\kappa}{6\tau} \Phi'''(x)$$

for the distribution function

$$F(x) = \mathbf{P}\{T - \mathbf{E}\,T \le x\sigma_T\}, \qquad \sigma_T^2 = \mathbf{Var}\,T.$$

Here

$$\tau^2 = Npq, \qquad p = n/N, \qquad q = 1 - p.$$

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

1

The moments

$$\alpha = \sigma_1^{-3} \mathbf{E}\, g_1^3(X_1) \qquad \text{and} \qquad \kappa = \sigma_1^{-3} \tau^2 \mathbf{E}\, g_2(X_1, X_2) g_1(X_1) g_1(X_2) \qquad (1.1)$$

refer to the linear part $\sum_i g_1(X_i)$ and the quadratic part $\sum_{i<j} g_2(X_i, X_j)$ of Hoeffding's decomposition

$$T = \mathbf{E}\, T + \sum_{1 \le i \le n} g_1(X_i) + \sum_{1 \le i < j \le n} g_2(X_i, X_j) + \dots. \qquad (1.2)$$

Here we assume that the linear part does not vanish, that is,

$$\sigma_1^2 > 0, \qquad \text{where} \qquad \sigma_1^2 = \mathbf{E}\, g_1^2(X_1). \qquad (1.3)$$

For basic facts about Hoeffding's decomposition of finite population statistics, we refer to [10]. Functions g_1 and g_2 can be expressed by linear combinations of the conditional expectations $h_1(x) = \mathbf{E}\left(T' \,\middle|\, X_1 = x\right)$ and $h_2(x, y) = \mathbf{E}\left(T' \,\middle|\, X_1 = x, X_2 = y\right)$, where $T' = T - \mathbf{E}\, T$,

$$g_1(x_1) = \frac{N-1}{N-n}\, h_1(x_1), \qquad (1.4)$$

$$g_2(x_1, x_2) = \frac{N-2}{N-n}\, \frac{N-3}{N-n-1}\, \left(h_2(x_1, x_2) - \frac{N-1}{N-2}\, (h_1(x_1) + h_1(x_2))\right).$$

Expressions for higher order symmetric kernels g_k, $k = 3, 4, \dots, $n, are determined in [10].

Note that, in order to write a one-term Edgeworth expansion, one does not need to evaluate all terms of Hoeffding's decomposition, but the moments of the linear and quadratic parts only. Even this problem can be difficult to solve. In such cases, one can use the empirical Edgeworth expansion

$$\hat{G}(x) = \Phi(x) - \frac{(q-p)\hat{\alpha} + 3\hat{\kappa}}{6\tau}\, \Phi'''(x),$$

where the true moments α and κ are replaced by their estimators. In the present paper, we consider the jackknife estimators $\hat{\alpha}$ and $\hat{\kappa}$; see (1.5) below.

In order to define the jackknife estimators we need few additional observations. In what follows, $\{X_1, \dots, X_m\}$, $m = n, n+1, n+2$, denote simple random samples drawn without replacement from $\mathcal{X}$. It is convenient to represent the sample $X_1, \dots, X_m$ by the set of first m variables of the random permutation $(X_1, \dots, X_N)$ of the ordered set $(x_1, \dots, x_N)$.

For $1 \le k \le n+1$, $1 \le r \le n+2$, and $1 \le i \ne j \le n+2$, denote

$$V_k = \overline{T} - T_{(k)}, \qquad \tilde{V}_r = \tilde{T} - \overline{T}_{(r)}, \qquad W_{ij} = \tilde{T} - \overline{T}_{(i)} - \overline{T}_{(j)} + T_{(i,j)},$$

where

$$\overline{T} = \frac{1}{n+1} \sum_{k=1}^{n+1} T_{(k)}, \qquad \overline{T}_{(r)} = \frac{1}{n+1} \sum_{1 \le j \le n+2, \, j \ne r} T_{(r,j)},$$

$$\tilde{T} = \frac{1}{\binom{n+2}{2}} \sum_{1 \le i < j \le n+2} T_{(i,j)},$$

and where $T_{(k)} = \varphi(X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n, X_{n+1})$ denotes the value of φ at the n-set of observations $\{X_1, \ldots, X_{n+1}\} \setminus \{X_k\}$ and $T_{(i,j)}$ denotes the value of φ at the n-set of observations $\{X_1, \ldots, X_{n+1}, X_{n+2}\} \setminus \{X_i, X_j\}$.

Write

$$\hat{\sigma}_J^2 = \sum_{k=1}^{n+1} V_k^2, \qquad \hat{\alpha}_J = \frac{\sqrt{n}}{\hat{\sigma}_J^3} \sum_{k=1}^{n+1} V_k^3,$$

$$\hat{\kappa}_J = \frac{2\sqrt{n}}{\hat{\sigma}_J^3} \sum_{1 \le i < j \le n+2} W_{ij} \tilde{V}_i \tilde{V}_j.$$

We show in Lemmas 2.1, 2.2, and 2.3 below that

$$\hat{\alpha} := \hat{\alpha}_J, \qquad \hat{\sigma}^2 := q\hat{\sigma}_J^2, \qquad \text{and} \qquad \hat{\kappa} := q\hat{\kappa}_J \tag{1.5}$$

are consistent estimators of α, σ_T^2, and κ.

There is a rich literature devoted to jackknife variance estimators of statistics based on *independent observations*; see, e.g., [26] and references therein. It is well known that the classical Quenouille–Tukey jackknife estimator of variance $\hat{\sigma}_J^2$ is asymptotically consistent provided that the underlying statistic is sufficiently "smooth"; see, e.g., [19], [20], and [27].

In order to obtain a consistent variance estimator in the case of *samples drawn without replacement*, one needs a finite population correction; see (1.5).

Beran [5] and Putter and van Zwet [23] used jackknife estimators of moments (corresponding to α and κ) to define empirical Edgeworth expansions for statistics based on *independent and identically distributed* observations. Putter and van Zwet [23] proved the consistency of the estimator $\hat{\alpha}_J$. However, their estimator of κ

$$\hat{\lambda}_2 = 2\sqrt{n}\sigma_J^{-3} \sum_{1 \le i < j \le n+2} W_{ij} V_i V_j$$

involves an unspecified quantity V_{n+2}. Replacement of V_i and V_j by $\tilde{V}_i$ and $\tilde{V}_j$ in $\hat{\lambda}_2$ results in the estimator $\hat{\kappa}_J$, which is a symmetric statistic of the observations $X_1, \ldots, X_{n+2}$. The consistency of $\hat{\kappa}_J$ in the i.i.d. case can be proved by using the argument of Lemma 5.2 of [23].

In order to obtain a consistent estimator of κ in the case of *samples drawn without replacement*, we introduce a finite population correction; see (1.5).

The moments α and κ are related to the linear and quadratic parts only. In order to establish the consistency of $\hat{\alpha}$ and $\hat{\kappa}$ we impose appropriate smoothness conditions which control higher order nonlinear terms of the decomposition (1.2). The smoothness conditions are formulated in terms of second moments of finite differences. We write $n_* = \min\{n, N - n\}$ and define

$$D^j T = \varphi(X_1, \ldots, X_n) - \varphi(X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n, X'_j), \quad X'_j = X_{n+j}.$$

Higher order difference operations are defined recursively:

$$D^{j_1, j_2} T = D^{j_2}\big(D^{j_1} T\big), \qquad D^{j_1, j_2, j_3} T = D^{j_3}\big(D^{j_2}(D^{j_1} T)\big), \ldots.$$

They are symmetric, i.e., $D^{j_1, j_2} T = D^{j_2, j_1} T$, etc. Given $k < n_*$, we write

$$\delta_j = \delta_j(T) = \mathbf{E}\left(n_*^{(j-1)} \mathbb{D}_j T\right)^2, \qquad \mathbb{D}_j T = D^{1,2,\ldots,j} T, \qquad 1 \le j \le k. \qquad (1.6)$$

The paper is organized as follows. In Section 2, we formulate our results. Proofs are given in Section 3. Some more technical calculations are postponed to Section 4. In Section 5, we collect moment inequalities for different parts of the decomposition (1.2). In Appendix 6, we estimate the convergence rate in probablity of finite population sample means.

Sections 4 - 6 will appear in the second part of this paper "Empirical Edgeworth expansion for finite population statistics. II," which will be published in another issue of this journal.

2. Results

To make the presentation mathematically rigorous we introduce the following model. Let $\mathcal{X}_\nu = \{x_{1,\nu}, \ldots, x_{N_\nu,\nu}\}$, $\nu = 1, 2, \ldots$, be a sequence of populations, and let $T_\nu = \varphi_\nu(X_{1,\nu}, \ldots, X_{n_\nu,\nu})$ be a sequence of symmetric statistics. Here $X_{1,\nu}, \ldots, X_{n_\nu,\nu}$ denote a sample of size $n_\nu \le N_\nu$ drawn without replacement from $\mathcal{X}_\nu$. Write

$$\tau_\nu^2 = N_\nu p_\nu q_\nu, \qquad p_\nu = n_\nu / N_\nu, \qquad q_\nu = 1 - p_\nu, \qquad \sigma_{T,\nu}^2 = \mathbf{Var}\, T_\nu,$$

and let α_ν, κ_ν, and $\sigma_{1,\nu}^2$ denote the moments defined by (1.1) and (1.3) with respect to the Hoeffding decomposition

$$T_\nu = \mathbf{E}\, T_\nu + \sum_{1 \le i \le n_\nu} g_{1,\nu}(X_{i,\nu}) + \sum_{1 \le i < j \le n_\nu} g_{2,\nu}(X_{i,\nu}, X_{j,\nu}) + \ldots.$$

Given ν, by $\hat{\sigma}_\nu^2$, $\hat{\alpha}_\nu$, and $\hat{\kappa}_\nu$ we denote jackknife estimators of $\sigma_{T,\nu}^2$, α_ν, and κ_ν defined by (1.5). Furthermore, write $\xi_\nu = \sigma_{1,\nu}^{-1} g_{1,\nu}(X_{1,\nu})$ and denote

$$\Psi(t) = |\mathbf{E}\,\exp\{it\xi_\nu\}|, \qquad \beta_{s,\nu} = \mathbf{E}\,|\xi_\nu|^s, \qquad \beta_{s,\nu}^*(\varepsilon) = \mathbf{E}\,|\xi_\nu|^s \mathbb{I}_{|\xi_\nu|^s > \varepsilon\tau_\nu^2},$$

$$\gamma_{s,\nu} = \sigma_{1,\nu}^{-s}\tau_\nu^{2s}\,\mathbf{E}\,|g_{2,\nu}(X_{1,\nu}, X_{2,\nu})|^s, \qquad \delta_{k,\nu} = \delta_k(T_\nu), \qquad \zeta_\nu = \sigma_{1,\nu}\tau_\nu.$$

We assume that there exist absolute constants $0 < C_1 < C_2 < \infty$ such that

$$C_1 \le \sigma_{T,\nu}^2 \le C_2, \qquad \nu = 1, 2, \ldots. \tag{2.1}$$

Furthermore, we assume that $\min\{n_\nu; N_\nu - n_\nu\} \to \infty$ as $\nu \to \infty$. In particular, we have $\tau_\nu^2 \to \infty$ and $N_\nu \to \infty$.

2.1 Edgeworth expansions.

Theorem 2.1 below provides the one-term Edgeworth expansion

$$G_\nu(x) = \Phi(x) - \frac{(q_\nu - p_\nu)\alpha_\nu + 3\kappa_\nu}{6\tau_\nu}\,\Phi'''(x)$$

for the distribution function $F_\nu(x) = \mathbf{P}\{T_\nu - \mathbf{E}\,T_\nu \le x\sigma_{T,\nu}\}$.

For particular classes of statistics that are smooth functions of finite population sample means, Edgeworth expansions were constructed by Babu and Singh [1] and Babu and Bai [2]. Theorem 2.1 below provides a uniform analogue of the corresponding result proved in [10].

Theorem 2.1. *Assume that $\tau_\nu \to \infty$. Suppose that there exist positive numbers C_1, C_2, C_3, and $s > 3$, $t > 2$, sequences $\varepsilon_\nu \downarrow 0$ and $\eta_\nu \uparrow \infty$, and a positive decreasing function ϕ on $(0, +\infty)$ such that (2.1) is satisfied,*

$$\delta_{3,\nu} \le \varepsilon_\nu \tau_\nu^{-1}, \tag{2.2}$$

$$\beta_{s,\nu} \le C_3, \qquad \gamma_{t,\nu} \le C_3, \tag{2.3}$$

and

$$\Psi(t) \le 1 - \phi(t), \qquad C_3^{-1} \le |t| \le \eta_\nu \tag{2.4}$$

for $\nu = 1, 2, \ldots$.

Then there exists a sequence $\psi_\nu \downarrow 0$ depending on C_1, C_2, C_3, s, t, $\{\varepsilon_\nu\}$, $\{\eta_\nu\}$, and ϕ only such that, for $\nu = 1, 2, \ldots$,

$$\sup_x |F_\nu(x) - G_\nu(x)| < \psi_\nu \tau_\nu^{-1}. \tag{2.5}$$

Our main result Theorem 2.2 establishes the validity (in probability) of the empirical Edgeworth expansion

$$\hat{G}_\nu(x) = \Phi(x) - \frac{(q_\nu - p_\nu)\hat{\alpha}_\nu + 3\hat{\kappa}_\nu}{6\tau_\nu}\,\Phi'''(x).$$

Theorem 2.2. *Assume that $\tau_\nu \to \infty$. Suppose that there exist positive numbers C_1, C_2, C_3, and $s > 3$, $t > 2$, sequences $\varepsilon_\nu \downarrow 0$ and $\eta_\nu \uparrow \infty$, and a positive decreasing function ϕ on $(0, +\infty)$ such that (2.1), (2.3), and (2.4) are satisfied and*

$$\delta_{2,\nu} \le \varepsilon_\nu n_\nu^{-1/3}, \qquad \delta_{3,\nu} \le \varepsilon_\nu \tau_\nu^{-1}, \tag{2.6}$$

for $\nu = 1, 2, \ldots$.

Then there exists a sequence $\psi_\nu \downarrow 0$, depending only on C_1, C_2, C_3, s, t, $\{\varepsilon_\nu\}$, $\{\eta_\nu\}$, and ϕ such that, for $\nu = 1, 2, \ldots$,

$$\mathbf{P}\big\{\sup_x |F_\nu(x) - \hat{G}_\nu(x)| > \psi_\nu \tau_\nu^{-1}\big\} < \psi_\nu. \tag{2.7}$$

Note that conditions (2.1), (2.3), (2.4), and (2.2) (respectively (2.6)) define the (uniformity) classes of sequences $\{(\mathcal{X}_\nu, \varphi_\nu), \nu = 1, 2, \ldots\}$, for which the bound (2.5) (respectively (2.7)) holds uniformly. This way of formulating results is convenient for applications like sub-sampling (see [21], [3], and [8]), jackknife histogram (see [28], [25], [13]), and finite population bootstrap (see [7], [14], [1], and [12]). All of these resampling schemes deal with classes of sequences $\{(\mathcal{X}_\nu, \varphi_\nu), \nu = 1, 2, \ldots\}$.

Empirical Edgeworth expansions for distribution functions of symmetric statistic based on i.i.d. observations were studied by a number of authors: Beran [5], Bhattacharya and Qumsiyeh [6], Helmers [16], Hall [15], and Putter and van Zwet [23], etc. The most general result yielding an empirical Edgeworth expansion for symmetric statistic of i.i.d. observations was obtained by Putter and van Zwet [23]. Theorem 2.2 could be considered as an extension of their result to the simple random sample model.

If $n_\nu/N_\nu \to 0$, the simple random sample model approaches the i.i.d. situation. In the i.i.d case, Theorem 2.2 remains valid, with q_ν (respectively p_ν) replaced by 1 (respectively 0), and with τ_ν replaced by $\sqrt{n_\nu}$. In the definition of $\delta_j(T_\nu)$ (see (1.6)), one should also replace n_* by n_ν. Note that, even in this case, Theorem 2.2 differs from the corresponding i.i.d. result of [23], where the smoothness conditions are formulated in terms of variances of higher order parts of Hoeffding's decomposition of statistics. In particular, in order to verify such conditions one should be able to estimate the variances of cubic and higher order parts of the decomposition. Although our smoothness conditions (2.6) do not refer to Hoeffding's decomposition, they, in fact, control the higher order parts of the decomposition as well. An advantage of our conditions is that they are much simpler and easier to handle.

The moments δ_k are estimated in [10] for general U–statistics and smooth functions of sample means. In particular, for U–statistics, we typically have $\delta_{k,\nu} = O(\tau_\nu^{-2})$ and, for smooth functions of sample means, we have $\delta_{k,\nu} = O(n_\nu^{-1})$.

In the case where $\limsup_\nu n_\nu N_\nu^{-1} \le 1 - \delta$ for some $0 < \delta \le 1$, we have that, for large ν, $\tau_\nu^2 \le n_\nu \le \delta^{-1}\tau_\nu^2$. In view of (3.16), condition (2.6) follows from

$$\delta_{3,\nu} \le \varepsilon_\nu \tau_\nu^{-1} \qquad \text{and} \qquad \gamma_{t,\nu} \le C_3 \quad \text{for some} \quad t \ge 2 \quad \text{and} \quad C_3 > 0.$$

2.2. Consistency of jackknife estimator of variance.

Lemma 2.1. *Assume that (2.1) holds. Suppose that $\delta_{2,\nu} = o(1)$, $\tau_\nu \to \infty$, and*

$$\forall \varepsilon > 0, \qquad \beta_{2,\nu}^*(\varepsilon) = o(1) \qquad \text{as} \qquad \nu \to \infty. \tag{2.8}$$

Then

$$\forall \delta > 0, \qquad \mathbf{P}\{|\hat\sigma_\nu^2 - \sigma_{T,\nu}^2| > \delta\} = o(1) \qquad \text{as} \quad \nu \to \infty. \tag{2.9}$$

Remark 2.1. Assume that (2.1) holds. Suppose that there exists a positive sequence $\varepsilon_\nu \downarrow 0$ such that $\delta_{2,\nu} \le \varepsilon_\nu$ and

$$\beta_{2,\nu}^*(\varepsilon_\nu) \le \varepsilon_\nu, \qquad \nu = 1, 2, \ldots. \tag{2.10}$$

Then there exists a positive sequence $\psi_\nu \downarrow 0$ depending on C_1, C_2, and $\{\varepsilon_\nu\}$ only such that

$$\mathbf{P}\{|\hat\sigma_\nu^2 - \sigma_{T,\nu}^2| > \psi_\nu\} < \psi_\nu, \qquad \nu = 1, 2, \ldots. \tag{2.11}$$

Remark 2.2. Condition (2.10) is satisfied if, for some $s > 2$ and $C_3 > 0$, we have $\beta_{s,\nu} < C_3$ for all $\nu = 1, 2, \ldots$. Indeed, the inequalities

$$\beta_{2,\nu}^*(\varepsilon) \le (\tau_\nu^2 \varepsilon)^{1-s/2} \beta_{s,\nu} \le (\tau_\nu^2 \varepsilon)^{1-s/2} C_3$$

imply (2.10) with $\varepsilon_\nu = C_3^{2/s} \tau_\nu^{-2+4/s}$.

2.3 Consistency of jackknife estimators of α and κ.

Lemma 2.2. *Assume that (2.1) holds. Suppose that $\tau_\nu \to \infty$ and $\delta_{2,\nu} = o(n_\nu^{-1/3})$, and, for every $\varepsilon > 0$,*

$$\beta_{3,\nu}^*(\varepsilon) = o(1), \qquad \beta_{3,\nu} = O(1)$$

as $\nu \to \infty$. Then

$$\forall \delta > 0, \qquad \mathbf{P}\{|\hat\alpha_\nu - \alpha_\nu| > \delta\} = o(1) \qquad \text{as} \qquad \nu \to \infty. \tag{2.12}$$

Remark 2.3. Assume that (2.1) holds. Suppose that there exist an absolute constant $C_3 > 0$ and a sequence $\varepsilon_\nu \downarrow 0$ such that $\delta_{2,\nu} \le \varepsilon_\nu n_\nu^{-1/3}$ and

$$\beta_{3,\nu}^*(\varepsilon_\nu) \le \varepsilon_\nu, \qquad \text{and} \qquad \beta_{3,\nu} \le C_3 \tag{2.13}$$

for $\nu = 1, 2, \ldots$. Then there exists a positive sequence $\psi_\nu \downarrow 0$ depending on C_1, C_2, C_3, and $\{\varepsilon_\nu\}$ only such that

$$\mathbf{P}\{|\hat\alpha_\nu - \alpha_\nu| > \psi_\nu\} < \psi_\nu, \qquad \nu = 1, 2, \ldots.$$

Remark 2.4. Condition (2.13) is satisfied if, for some $s > 3$ and $C_3 > 0$, we have $\beta_{s,\nu} < C_3$ for $\nu = 1, 2, \ldots$. Indeed, the inequalities $\beta_{3,\nu} \le \beta_{s,\nu}^{3/s}$ and

$$\beta_{3,\nu}^*(\varepsilon) \le (\varepsilon\tau_\nu^2)^{1-s/3} \beta_{s,\nu} \le (\varepsilon\tau_\nu^2)^{1-s/3} C_3$$

imply (2.13) with $\varepsilon_\nu = C_3^{3/s} \tau_\nu^{-2+6/s}$.

Lemma 2.3. *Assume that (2.1) holds and $\tau_\nu \to \infty$ as $\nu \to \infty$. Suppose that (2.3) is satisfied with some $s \geq 3$ and $t \geq 2$. Assume that, for some $\varepsilon_\nu \downarrow 0$,*

$$\delta_3 \leq \varepsilon_\nu, \qquad \nu = 1, 2, \ldots.$$

Then there exists a sequence $\psi_\nu \downarrow 0$ depending on C_1, C_2, C_3, and $\{\varepsilon_\nu\}$ only such that

$$\mathbf{P}\{|\hat{\kappa}_\nu - \kappa_\nu| > \psi_\nu\} < \psi_\nu, \qquad \nu = 1, 2, \ldots. \tag{2.14}$$

3. Proofs

In order to simplify the notation we drop the subscript ν whenever this does not cause an ambiguity. By c we shall denote positive absolute constants.

For $k = 1, 2, \ldots$ and $x \in \mathbb{R}$, write

$$\Omega_k = \{1, \ldots, k\}, \qquad [x]_k = x(x-1)\ldots(x-k+1), \qquad [x]_0 = 1.$$

By A_j we shall always denote a subset of Ω_N of cardinality $|A_j| = j$.

Given a function $f : \mathcal{X} \to [0, +\infty)$, subset $A = \{i_1, \ldots, i_k\} \subset \Omega_N$, and $j \in \Omega_N \setminus A$, the following inequality holds:

$$\mathbf{E}\left(f(X_j)\big| X_{i_1}, \ldots, X_{i_k}\right) = \frac{1}{N-k} \sum_{k \in \Omega_N \setminus A} f(X_k) \leq \frac{N}{N-k} \mathbf{E} f(X_j). \tag{3.1}$$

We shall often use the inequality, which is an immediate consequence of (3.1),

$$\mathbf{E} f(X_i)f(X_j) \leq N(N-1)^{-1}(\mathbf{E} f(X_i))^2 \leq 2(\mathbf{E} f(X_i))^2.$$

Since Hoeffding's decomposition and its properties play the central role in our proofs, we collect basic facts about the decomposition in a separate Subsection 3.1. Proofs of these facts can be found in [10]. In Subsection 3.2, we prove Theorems 2.1 and 2.2. Lemmas 2.1–3 and Remarks 2.1 and 2.3 are proved in Subsection 3.3.

3.1. Let U_j denote the j-th sum in (1.2),

$$U_j = \sum_{1 \leq i_1 < \cdots < i_j \leq n} g_j(X_{i_1}, \ldots, X_{i_j}).$$

Given $1 \leq j \leq n$, the kernel g_j satisfies $\mathbf{E} g_j(X_{i_1}, \ldots, X_{i_j}) = 0$ and

$$\mathbf{E}\left(g_j(X_{i_1}, \ldots, X_{i_j})\big| X_{k_1}, \ldots, X_{k_r}\right) = 0 \qquad \text{a.s.} \tag{3.2}$$

for every $1 \le i_1 < \cdots < i_j \le n$ and every $1 \le k_1 < \cdots < k_r \le n$ such that $r < j$. In particular, we have

$$\mathbf{E}\, g_j(X_{i_1}, \ldots, X_{i_j}) g_r(X_{k_1}, \ldots, X_{k_r}) = 0 \tag{3.3}$$

By (3.3), the random variables U_i and U_j with $i \ne j$ are uncorrelated.

In the case where T is an U–statistic of arbitrary fixed degree k, that is,

$$\varphi(x_1, \ldots, x_n) = \sum_{1 \le i_1 < \cdots < i_k \le n} h(x_{i_1}, \ldots, x_{i_k})$$

for some symmetric kernel h, we have $U_j = 0$ almost surely for every $j > k$.

In the case where $n > N/2$, we have $U_j = 0$ almost surely for $j > N - n$.

For evey $j \le n_*$, the following identity holds almost surely:

$$U_j = (-1)^j U_j', \qquad U_j' = \sum_{n_*+1 \le i_1 < \cdots < i_j \le N} g_j(X_{i_1}, \ldots, X_{i_j}). \tag{3.4}$$

Therefore, by (1.2), we have almost surely (see (2.7) and (2.8) in [10])

$$T - \mathbf{E}\, T = U_1 + \cdots + U_{n_*} = (-1)U_1' + \cdots + (-1)^{n_*} U_{n_*}'. \tag{3.5}$$

Furthermore, since U_i and U_j are uncorrelated, we have

$$\sigma_T^2 = \mathbf{E}\, U_1^2 + \cdots + \mathbf{E}\, U_{n_*}^2. \tag{3.6}$$

Conditioning on $X_1 = y_1$, ..., $X_n = y_n$, we obtain from (3.5) that

$$\varphi(y_1, \ldots, y_n) = \mathbf{E}\, T + \sum_{k=1}^{n_*} \sum_{1 \le i_1 < \cdots < i_k \le n} g_k(y_{i_1}, \ldots, y_{i_k}) \tag{3.7}$$

for every $\{y_1, \ldots, y_n\} \subset \{x_1, \ldots, x_N\}$.

Given $A = \{i_1, \ldots, i_j\} \subset \Omega_N$, with $1 \le j \le n$, write $T_A = g_j(X_{i_1}, \ldots, X_{i_j})$ and denote $T_i = g_1(X_i)$ and $T_{ij} = g_2(X_i, X_j)$.

Given a class $\mathcal{H}$ of subsets $A \subset \Omega_N$ with cardinality $|A| \le n_*$, introduce the random variable $S = \sum_{A \in \mathcal{H}} T_A$. Split

$$S = U_1(S) + \cdots + U_{n_*}(S), \qquad U_j(S) = \sum_{A_j \in \mathcal{H}} T_{A_j}.$$

By (3.3), $\mathbf{E}\, S^2 = \mathbf{E}\, U_1^2(S) + \cdots + \mathbf{E}\, U_{n_*}^2(S)$. In particular, given $k < n_*$, we have

$$\mathbf{E}\, (\mathbb{D}_k T)^2 = \mathbf{E}\, U_k^2(\mathbb{D}_k T) + \cdots + \mathbf{E}\, U_{n_*}(\mathbb{D}_k T)^2. \tag{3.8}$$

For $A, B \subset \Omega$ with $|A| = |B| = j \le n$ and $|A \cap B| = m$, write

$$\sigma_j^2 = \mathbf{E}\, T_A^2, \qquad s_{j,m} = \mathbf{E}\, T_A T_B.$$

It is easy to derive from (3.2) (see Lemma 1 of [10]) that

$$s_{j,m} = (-1)^{j-m} \binom{N-j}{j-m}^{-1} \sigma_j^2, \qquad 0 \le m \le j \le n. \tag{3.9}$$

Lemma 3.1A. *(Lemma 2 of [10]). For $1 \leq i \leq j \leq n_*$, the following identities hold:*

$$\mathbf{E}\, U_j^2 = \frac{\overset{n}{\underset{j}{}} \ \overset{N-n}{\underset{j}{}}}{\underset{j}{N-j}} \, \sigma_j^2, \tag{3.10}$$

$$\mathbf{E}\, U_j^2 = 2^{-i} h_{i,j} \mathbf{E}\, U_j^2(\mathbb{D}_i T), \qquad h_{i,j} = \frac{[n]_i [N-n]_i}{[j]_i [N-j+1]_i}, \tag{3.11}$$

$$\mathbf{E}\, U_j^2 + \cdots + \mathbf{E}\, U_{n_*}^2 \leq n_*^{-(j-2)} \delta_j. \tag{3.12}$$

An application of (3.10) with $j = 1$ to the statistic $T_1 + \cdots + T_k$ gives

$$\mathbf{E}\, (T_1 + \cdots + T_k)^2 = k(N-k)(N-1)^{-1}\sigma_1^2. \tag{3.13}$$

3.2. We need the following lemma, which is proved in Section 4 below.

Lemma 3.1. *Assume that (2.1) holds. Then*

$$\zeta^2 \leq C_2, \tag{3.14}$$

$$0 \leq \sigma_T^2 - \zeta^2 \leq \delta_2 + (N-1)^{-1}C_2, \tag{3.15}$$

$$\delta_2 \leq 2^6 C_2 n_*^{-1} \gamma_2 + 2^{-1} n_*^{-1} \delta_3. \tag{3.16}$$

Proof of Theorem 2.2. The theorem follows from Theorem 2.1, Remark 2.4, and Lemma 2.3.

Proof of Theorem 2.1. Write $\Delta_\nu = \sup_x |\mathbf{P}\{U_{\{\nu\}} \leq x\zeta_\nu\} - G_\nu(x)|$, where

$$U_{\{\nu\}} = \sum_{1 \leq i \leq n_\nu} g_{1,\nu}(X_{i,\nu}) + \sum_{1 \leq i < j \leq n_\nu} g_{2,\nu}(X_{i,\nu}, X_{j,\nu}).$$

Theorem 1 of [9] provides the bound $\Delta_\nu = o(\tau_\nu^{-1})$ as $\nu \to \infty$. Analysis of their proof shows that, under conditions (2.1), (2.3), and (2.4), one can specify the dependence of the bound on the parameters C_1, C_2, C_3, s, t, and the function ϕ in the following way. There exists a sequence $\psi_\nu \downarrow 0$ depending only on C_1, C_2, C_3, s, t, and ϕ such that

$$\Delta_\nu \leq \psi_\nu \tau_\nu^{-1}, \qquad \nu = 1, 2, \ldots. \tag{3.17}$$

Write $T_\nu - \mathbf{E}\, T_\nu = U_{\{\nu\}} + R_{\{\nu\}}$. Then

$$F_\nu(x) = \mathbf{P}\{U_{\{\nu\}} \leq \zeta_\nu(x - v_1 x - v_2)\}, \qquad v_1 = 1 - \zeta_\nu^{-1}\sigma_{T,\nu}, \quad v_2 = \zeta_\nu^{-1} R_{\{\nu\}}.$$

A Slutzky's type argument yields, for every $b > 0$,

$$\sup_x |F_\nu(x) - G_\nu(x)| \leq \Delta_\nu + I_1 + I_2,$$

$$I_1 = \sup_{x \in \mathbb{R}, |r|=b} |G_\nu(x) - G_\nu(x - v_1 x - r)|, \qquad I_2 = \mathbf{P}\{|v_2| > b\}.$$

In view of (3.17), it suffices to construct a similar bound for $I_1 + I_2$.

By C^* we shall denote a constant depending on C_1, C_2, C_3, s, t, and the sequence $\{\varepsilon_\nu\}$ only. It follows from (2.1), (2.2), and (2.3) via (3.16) and (3.15) that

$$|\sigma_{T,\nu}^2 - \zeta_\nu^2| \leq C^* \tau_\nu^{-2} \qquad \text{and} \qquad \zeta_\nu^2 \geq C_1 - C^* \tau_\nu^{-2}. \qquad (3.18)$$

Combining Hölder's inequality and (3.1), we infer from (2.3) that $|\alpha_\nu| < C^*$ and $|\kappa_\nu| < C^*$. Since the function $\Phi(x)$ and its derivatives decay exponentially as $x \to \infty$, from these bounds we obtain the inequality $I_1 \leq C^* |v_1| + C^* b$. Furthermore, by (3.18), we have $|v_1| \leq C^* \tau_\nu^{-2}$. Therefore, $I_1 \leq C^* (\tau_\nu^{-2} + b)$. Finally, by Chebyshev's inequality, (3.12), and (3.18),

$$I_2 \leq b^{-2} \mathbf{E}\, v_2^2 \leq b^{-2} \zeta_\nu^{-2} \tau_\nu^{-2} \delta_{3,\nu} \leq C^* b^{-2} \tau_\nu^{-2} \delta_{3,\nu}.$$

We obtain

$$I_1 + I_2 \leq C^* (\tau_\nu^{-2} + b + b^{-2} \tau_\nu^{-2} \delta_{3,\nu}).$$

Choosing $b^{-1} = \tau_\nu \varepsilon_\nu^{1/3}$, we obtain $I_1 + I_2 \leq \psi_\nu \tau_\nu^{-1}$ with some $\psi_\nu \downarrow 0$, thus, completing the proof of the theorem.

3.3. *Proof of Remarks 2.1 and 2.3.* Let us prove Remark 2.1. Note that (2.9) is equivalent to the following statement: there exists a sequence $\psi_\nu \downarrow 0$ such that (2.11) holds. Analysis of the proof of Lemma 2.1 (see below) shows that one can choose the sequence $\{\psi_\nu\}$ so that it depends on C_1, C_2, and $\{\varepsilon_\nu\}$ only.

Remark 2.3 is obtained from Lemma 2.2 in a similar way.

Proof of Lemma 2.1. Using representation (3.7), we obtain

$$V_i = v_i + r_i, \qquad 1 \leq i \leq n+1, \qquad (3.19)$$

$$v_i = \sum_{A_1 \subset \Omega_{n+1}} \left(\mathbb{I}_{i \in A_1} - \frac{1}{n+1} \right) T_{A_1}, \qquad r_i = \sum_{j=2}^{n_*} \sum_{A_j \subset \Omega_{n+1}} \left(\mathbb{I}_{i \in A_j} - \frac{j}{n+1} \right) T_{A_j}.$$

Therefore, we can write $\hat{\sigma}_J^2 = V^2 + R$, where

$$V^2 = \sum_{i=1}^{n+1} v_i^2, \qquad R = 2W_1 + W_2^2, \qquad W_1 = \sum_{i=1}^{n+1} r_i v_i, \qquad W_2^2 = \sum_{i=1}^{n+1} r_i^2.$$

By (3.15), $\sigma_T^2 - \zeta^2 = o(1)$. Therefore, in order to prove (2.9) it suffices to show that

$$|qV^2 - \zeta^2| = o_P(1) \qquad \text{and} \qquad qR = o_P(1). \qquad (3.20)$$

To prove the first part of (3.20), write

$$V^2 = Z - Y^2/(n+1), \qquad Z = \sum_{i=1}^{n+1} T_i^2, \qquad Y = \sum_{i=1}^{n+1} T_i.$$

We have, by Lemma 6.1, $qZ - \zeta^2 = o_P(1)$. Furthermore, $Y^2/(n+1) = o_P(1)$, since, by (3.13) and (3.14), $\mathbf{E}\, Y^2/(n+1) \le \zeta^2/n \le C_2/n$.

In order to prove the second part of (3.20) it suffices to show that $q\mathbf{E}\, W_2^2 = o(1)$. Indeed, by Cauchy–Schwarz,

$$q\mathbf{E}\, |W_1| \le q\mathbf{E}\, V\, W_2 \le (q\mathbf{E}\, V^2)^{1/2}(q\mathbf{E}\, W_2^2)^{1/2},$$

where $q\mathbf{E}\, V^2 = \mathbf{E}\, U_1^2 \le C_2$. Finally, by symmetry and (4.1),

$$q\mathbf{E}\, W_2^2 = q(n+1)\mathbf{E}\, r_1^2 = (\tau^2 + q)\mathbf{E}\, r_1^2 \le \delta_2 = o(1).$$

Proof of Lemma 2.2. Write $\hat{\alpha} = WQ$, where $Q = \zeta^3 \hat{\sigma}^{-3}$. Since the bound $\beta_3 = O(1)$ implies (2.8), from (3.15) and Lemma 2.1 we obtain that $Q - 1 = o_P(1)$. In order to prove (2.12) it remains to show that

$$W - \alpha = o_P(1), \qquad \text{where} \qquad W = n^{-1} \sum_{j=1}^{n+1} V_j^3 \sigma_1^{-3}.$$

By (3.19), $W - \alpha = R_0 + 3R_1 + 3R_2 + R_3$, where

$$R_0 = w - \alpha, \quad w = n^{-1} \sum_{j=1}^{n+1} v_j^3 \sigma_1^{-3}, \qquad R_1 = n^{-1} \sum_{j=1}^{n+1} v_j^2 r_j \sigma_1^{-3},$$

$$R_2 = n^{-1} \sum_{j=1}^{n+1} v_j r_j^2 \sigma_1^{-3}, \qquad R_3 = n^{-1} \sum_{j=1}^{n+1} r_j^3 \sigma_1^{-3}.$$

We shall show that $R_i = o_P(1)$ for $0 \le i \le 3$.

Write

$$v_j = T_j - r, \qquad r = (n+1)^{-1}(T_1 + \cdots + T_{n+1}).$$

We have

$$R_0 = S_3 - 3S_2 \frac{r}{\sigma_1} + 3S_1 \frac{r^2}{\sigma_1^2} - \frac{n+1}{n} \frac{r^3}{\sigma_1^3} - \alpha, \qquad S_k = n^{-1} \sum_{i=1}^{n+1} \frac{T_i^k}{\sigma_1^k}.$$

By (3.13), $\mathbf{E}\, S_1^2 \leq n^{-1}$ and $\mathbf{E}\, r^2 \sigma_1^{-2} \leq n^{-1}$. Therefore, $S_1 = o_P(1)$ and $r/\sigma_1 = o_P(1)$. Furthermore, Lemma 6.1 gives $S_3 - \alpha = o_P(1)$ and $S_2 - 1 = o_P(1)$. Hence, we obtain $R_0 = o_P(1)$.

Let us prove that $R_i = o_P(1)$ for $i = 1, 2, 3$. By Hölder's inequality,

$$R_1 \leq w^{2/3} R_3^{1/3}, \qquad R_2 \leq w^{1/3} R_3^{2/3}.$$

Note that $w = R_0 + \alpha$ is stochastically bounded as $\nu \to \infty$. Therefore, it suffices to show that $R_3 = o_P(1)$. Given $\varepsilon > 0$, we have, by Chebyshev's inequality,

$$\mathbf{P}\{|R_3| > \varepsilon\} \leq \varepsilon^{-2/3} \mathbf{E}\, |R_3|^{2/3} \leq \varepsilon^{-2/3} n_*^{-1} \sigma_1^{-2} n^{1/3} \delta_2 = o(1).$$

Here we use the inequality $(\sum_i |r_i|^3)^{2/3} \leq \sum_i r_i^2$, the bound (4.1), and the fact that $n_* \sigma_1^2 \geq \zeta^2 = \sigma_T^2 + o(1)$ is bounded away from zero by (3.15) and (2.1). Lemma is proved.

Proof of Lemma 2.3. For brevity, we shall prove the bound $\hat{\kappa} - \kappa = o_P(1)$ only. Going along the line of the proof, one can easily see that the bound depends on C_1, C_2, C_3, s, t, and $\{\varepsilon_\nu\}$ only, i.e., that (2.14) holds.

Note that (2.3) implies $\kappa = O(1)$ as $\nu \to \infty$. Indeed, by Cauchy–Schwarz and (3.1),

$$\kappa \leq \sigma_1^{-2} \left(\mathbf{E}\, T_1^2 T_2^2\right)^{1/2} \gamma_2^{1/2} \leq (N/(N-1))^{1/2} \beta_2 \gamma_2^{1/2}.$$

Throughout the proof by $\sum'$ we denote the sum $\sum_{1 \leq i < j \leq n+2}$. Write

$$S = \sideset{}{'}\sum W_{ij} \tilde{V}_i \tilde{V}_j \qquad \text{and} \qquad s = \sideset{}{'}\sum w_{ij} \tilde{v}_i \tilde{v}_j,$$

where we denote

$$w_{ij} = T_{ij} - (n+1)^{-1}(T_i + T_j), \tag{3.21}$$

$$\tilde{v}_i = \frac{n}{n+1}(T_i - T_*), \qquad T_* = \frac{1}{n+2} \sum_{j=1}^{n+2} T_j.$$

Invoking the identity $2qn^{1/2} \hat{\sigma}_J^{-3} = 2q^2 \tau \hat{\sigma}^{-3}$, we obtain

$$\hat{\kappa} - \kappa = \hat{\sigma}^{-3}(2q^2 \tau S - \zeta^3 \kappa) + (\hat{\sigma}^{-3} \zeta^3 - 1)\kappa.$$

By (2.1), (3.15), and Lemma 2.1, we have $\hat{\sigma}^{-3} \leq C_1^{-3/2} + o(1)$ and $\hat{\sigma}^{-3} \zeta^3 - 1 = o_P(1)$. Therefore, in order to prove $\hat{\kappa} - \kappa = o_P(1)$ it suffices to show that $2q^2 \tau S - \zeta^3 \kappa = o_P(1)$. We shall prove this bound in two steps by showing that

$$q^2 \tau (S - s) = o_P(1), \tag{3.22}$$

$$2q^2 \tau s - \zeta^3 \kappa = o_P(1). \tag{3.23}$$

Proof of (3.22). Firstly, we shall show that, for every $i \in \Omega_{n+2}$, and $\{i, j\} \subset \Omega_{n+2}$,

$$\tilde{V}_i = \tilde{v}_i + \tilde{r}_i \qquad \text{and} \qquad W_{ij} = w_{ij} + r_{ij}, \tag{3.24}$$

where $\mathbf{E}\,\tilde{r}_i^2$ and $\mathbf{E}\,r_{ij}^2$ satisfy

$$\mathbf{E}\,\tilde{r}_i^2 \leq 2^{-1} n_*^{-1} \delta_2, \qquad \mathbf{E}\,r_{ij}^2 \leq c\,\frac{q\sigma_1^2}{n^3} + c\,\frac{q\sigma_2^2}{n} + c\,\frac{\delta_3}{n_*^3}. \tag{3.25}$$

Using representation (3.7) for $T_{(i,j)}$, $\{i, j\} \subset \Omega_{n+2}$, we obtain

$$\tilde{T} = \sum_{k=1}^{n_*} \frac{[n+2-k]_2}{[n+2]_2} \sum_{A_k \subset \Omega_{n+2}} T_{A_k},$$

$$\overline{T}_{(i)} = \sum_{k=1}^{n_*} \frac{n+1-k}{n+1} \sum_{A_k \subset \Omega_{n+2}} \mathbb{I}_{i \notin A_k} T_{A_k}.$$

Furthermore, invoking the identity $\mathbb{I}_{i \notin A_k} = 1 - \mathbb{I}_{i \in A_k}$, we get

$$\tilde{V}_i = \tilde{T} - \overline{T}_{(i)} = \sum_{k=1}^{n_*} \tilde{V}_{i,k}, \qquad \tilde{V}_{i,k} = \sum_{A_k \subset \Omega_{n+2}} \frac{n+1-k}{n+1}\left(\mathbb{I}_{i \in A_k} - \frac{k}{n+2}\right) T_{A_k}.$$

Note that $\tilde{V}_{i,1} = \tilde{v}_i$. Therefore, the first identity of (3.24) holds with

$$\tilde{r}_i = \tilde{V}_{i,2} + \cdots + \tilde{V}_{i,n_*}. \tag{3.26}$$

In order to prove the second identity of (3.24), we write W_{ij} in the following form (cf. [22], pp. 74–75):

$$W_{ij} = Z_1 + \cdots + Z_{n_*}, \tag{3.27}$$

$$Z_k = \sum_{A_k \subset \Omega_{n+2}} \left(\frac{[k+1]_2}{[n+2]_2} - \left(\mathbb{I}_{i \in A_k} + \mathbb{I}_{j \in A_k}\right)\frac{k}{n+1} + \mathbb{I}_{i \in A_k}\,\mathbb{I}_{j \in A_k}\right) T_{A_k}.$$

Denoting $R_0 = Z_3 + \cdots + Z_{n_*}$, we obtain $W_{ij} = Z_1 + Z_2 + R_0$. Furthermore,

$$Z_1 + Z_2 = w_{ij} + R_1 + R_2 + R_3 + R_4,$$

$$R_1 = \frac{2}{(n+1)(n+2)} \sum_{A_1 \subset \Omega_{n+2}} T_{A_1}, \qquad R_2 = \frac{6}{(n+1)(n+2)} \sum_{A_2 \subset \Omega_{n+2}} T_{A_2},$$

$$R_3 = -\frac{2}{n+1} \sum_{k \in \Omega_{n+2}\setminus\{i\}} T_{\{i,k\}}, \qquad R_4 = -\frac{2}{n+1} \sum_{k \in \Omega_{n+2}\setminus\{j\}} T_{\{j,k\}}.$$

Finally, we obtain

$$W_{ij} = w_{ij} + r_{ij}, \qquad \text{with} \qquad r_{ij} = R_0 + \cdots + R_4. \tag{3.28}$$

The bounds (3.25) are proved in Lemma 4.2 below.

Using representation (3.24), we can write

$$S - s = \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3 + \tilde{R}_4 + \tilde{R}_5,$$

$$\tilde{R}_1 = {\sum}' w_{ij}(\tilde{v}_i \tilde{r}_j + \tilde{r}_i \tilde{v}_j), \qquad \tilde{R}_2 = {\sum}' r_{ij}(\tilde{v}_i \tilde{r}_j + \tilde{r}_i \tilde{v}_j),$$

$$\tilde{R}_3 = {\sum}' w_{ij}\tilde{r}_i \tilde{r}_j, \qquad \tilde{R}_4 = {\sum}' r_{ij}\tilde{r}_i \tilde{r}_j, \qquad \tilde{R}_5 = {\sum}' r_{ij}\tilde{v}_i \tilde{v}_j.$$

Now (3.22) follows from the bounds, which are proved in Lemma 4.1 below,

$$q^2 \tau \tilde{R}_i = o_P(1), \qquad 1 \le i \le 5. \tag{3.29}$$

Proof of (3.23). Write

$$s^* = {\sum}' w_{ij} T_i T_j$$

and note that $\mathbf{E}\, q^2 \tau |s^*| = O(1)$ as $\nu \to \infty$. In particular, $q^2 \tau s^* = O_P(1)$ as $\nu \to \infty$. Indeed, by symmetry and (4.4) we have

$$\mathbf{E}\, q^2 \tau |s^*| \le q^2 \tau 2^{-1}[n+2]_2 \mathbf{E}\, |w_{1n} T_1 T_n| = O(1).$$

It is easy to see that (3.23) is a consequence of the following bounds:

$$q^2 \tau \big(s - n^2(n+1)^{-2} s^*\big) = o_P(1), \tag{3.30}$$

$$q^2 \tau (s^* - \mathbf{E}\, s^*) = o_P(1), \tag{3.31}$$

$$\Delta := 2q^2 \tau n^2 [n+2]_2^{-1} \mathbf{E}\, s^* - \zeta^3 \kappa = o(1). \tag{3.32}$$

The bounds (3.30) and (3.32) are proved in Lemma 4.1.

Let us prove (3.31). In what follows, C_* denotes a positive constant which depends on C_1, C_2, C_3, s, t, and $\{\varepsilon_\nu\}$ only.

The random variable

$$s^* - \mathbf{E}\, s^* = {\sum}' s(X_i, X_j), \qquad s(X_i, X_j) = w_{ij} T_i T_j - \mathbf{E}\, w_{ij} T_i T_j,$$

is an U-statistic of degree two of the observations $X_1, \ldots, X_{n+2}$. Hoeffding's decomposition (1.2) represents it by the sum of the linear and quadratic parts $s^* - \mathbf{E}\, s^* = L + Q$, where

$$L = \sum_{1 \leq i \leq n+2} g_1^*(X_i), \qquad Q = \sum{}' g_2^*(X_i, X_j),$$

and where, by (1.4),

$$g_1^*(X_i) = (n+1)t^*(X_i), \qquad g_2^*(X_i, X_j) = s(X_i, X_j) - t^*(X_i) - t^*(X_j),$$
$$t^*(X_i) = (N-1)(N-2)^{-1}\mathbf{E}\left(s(X_i, X_j) \big| X_i\right).$$

In order to prove (3.31) we shall show that

$$\tau q^2 L = o_P(1), \tag{3.33}$$
$$\tau q^2 Q = o_P(1). \tag{3.34}$$

Let us prove (3.33). Since $\mathbf{E}\,|t^*(X_1)|^{6/5} \leq C_* \mathbf{E}\,|w_{1n} T_1 T_n|^{6/5}$, from (4.4) we obtain the bound

$$\mathbf{E}\,|g_1^*(X_1)|^{6/5} = (n+1)^{6/5}\mathbf{E}\,|t^*(X_1)|^{6/5} = O(n^{6/5}\tau^{-6}). \tag{3.35}$$

Let $X_1^*, \ldots, X_{n+2}^*$ denote a sample drawn *with replacement* from $\mathcal{X}$. By Theorem 4 of Hoeffding (1963),

$$\mathbf{E}\,|L|^{6/5} \leq \mathbf{E}\,\Big|\sum_{1 \leq i \leq n+2} g_1^*(X_i^*)\Big|^{6/5}. \tag{3.36}$$

Furthermore, by the Marcinkievicz–Zygmund inequality,

$$\mathbf{E}\,\Big|\sum_{1 \leq i \leq n+2} g_1^*(X_i^*)\Big|^{6/5} \leq C \sum_{1 \leq i \leq n+2} \mathbf{E}\,|g_1^*(X_i^*)|^{6/5} = C(n+2)\mathbf{E}\,|g_1^*(X_1)|^{6/5}.$$

In combination with (3.36) and (3.35), this inequality yields

$$\mathbf{E}\,|L|^{6/5} = O(n^{11/5}\tau^{-6}).$$

Therefore, we obtain $\mathbf{E}\,|\tau q^2 L|^{6/5} = O(\tau^{-2/5})$, and now (3.33) follows.

Let us prove (3.34). In view of (3.4), we can further assume without loss of generality that $n \leq N/2$. In particular, $n \leq 2\tau^2$.

We first replace $g_2^*(X_i, X_j)$ by the truncated random variables

$$g_2^*(X_i, X_j)\mathbb{I}_i\mathbb{I}_j, \qquad \mathbb{I}_j = \mathbb{I}_{|T_i|<\eta}, \qquad \eta = \tau^{-1/4}.$$

Write $M = \sum' g_2^*(X_i, X_j)\mathbb{I}_i\mathbb{I}_j$. By symmetry and Chebyshev's inequality

$$\mathbf{P}\{Q \neq M\} \leq \mathbf{P}\{\max_{1\leq i\leq n+2}|T_i| > \eta\} \leq (n+2)\mathbf{P}\{|T_1| > \eta\}$$

$$\leq (n+2)\eta^{-3}\mathbf{E}\,|T_1|^3 \leq C_*\tau^{-1/4}\beta_3 = O(\tau^{-1/4}).$$

Therefore, it suffices to prove (3.34) with Q replaced by M. For this purpose, we shall show that

$$\tau q^2(M - \mathbf{E}\,M) = o_P(1) \tag{3.37}$$

and $\tau q^2\mathbf{E}\,M = o(1)$. The last bound follows from the bound (which is proved in Lemma 4.1 below)

$$\mathbf{E}\,g_2^*(X_i, X_j)\mathbb{I}_i\mathbb{I}_j = O(\tau^{1/4-6}). \tag{3.38}$$

Indeed, by symmetry and (3.38) we have

$$\tau q^2\mathbf{E}\,M = \tau q^2 2^{-1}[n+2]_2\mathbf{E}\,g_2^*(X_i, X_j)\mathbb{I}_i\mathbb{I}_j = O(\tau^{-3/4}).$$

Let us prove (3.37). Using (1.2), we decompose the statistic

$$M - \mathbf{E}\,M = \sum{}' m(X_i, X_j), \qquad m(X_i, X_j) = g_2^*(X_i, X_j)\mathbb{I}_i\mathbb{I}_j - \mathbf{E}\,g_2^*(X_i, X_j)\mathbb{I}_i\mathbb{I}_j,$$

into the sum of the linear and quadratic parts $M - \mathbf{E}\,M = \tilde{L} + \tilde{Q}$, where

$$\tilde{L} = \sum_{1\leq i\leq n+2}\tilde{g}_1(X_i), \qquad \sum{}'\tilde{g}_2(X_i, X_j).$$

Here, by (1.4),

$$\tilde{g}_1(X_i) = (n+1)\tilde{t}(X_i), \qquad \tilde{g}_2(X_i, X_j) = m(X_i, X_j) - \tilde{t}(X_i) - \tilde{t}(X_j),$$

$$\tilde{t}(X_i) = (N-1)(N-2)^{-1}\mathbf{E}\,(m(X_i, X_j)|\,X_i).$$

Clearly, (3.37) follows from the bounds

$$\tau q^2\tilde{L} = o_P(1) \qquad \text{and} \qquad \tau q^2\tilde{Q} = o_P(1).$$

The proof of the first bound is similar to that of (3.33). In order to prove the second bound we show that $\tau^2 q^4\mathbf{E}\,\tilde{Q}^2 = o(1)$. An application of (3.10) with $j = 2$ gives

$$\mathbf{E}\,\tilde{Q}^2 = \frac{\binom{n+2}{2}\binom{N-n-2}{2}}{\binom{N-2}{2}}\mathbf{E}\,\tilde{g}_2^2(X_1, X_2) \leq 2\tau^4\mathbf{E}\,\tilde{g}_2^2(X_1, X_2).$$

In the last step, we used the inequality $n^2q^2 \leq \tau^4$. Finally, the bound (which is proved in Lemma 4.1)

$$\mathbf{E}\,\tilde{g}_2^2(X_1, X_2) = O(\tau^{-7}) \tag{3.39}$$

completes the proof of the lemma.

References

[1] G. J. Babu, and K. Singh, *Edgeworth expansions for sampling without replacement from finite populations*, J. Multivar. Anal. **17** (1985), 261–278.

[2] G. J. Babu, and Z. D. Bai, *Mixtures of global and local Edgeworth expansions and their applications*, J. Multivar. Anal. **59** (1996), 282–307.

[3] P. Bertail, *Second-order properties of an extrapolated bootstrap without replacement under weak assumptions*, Bernoulli **3** (1997), 149–179.

[4] V. Bentkus, F. Götze, and W. R. van Zwet, *An Edgeworth expansion for symmetric statistics*, Ann. Statist. **25** (1997), 851–896.

[5] R. Beran, *Estimated sampling distributions: the bootstrap and competitors.*, Ann. Statist. **10** (1982), 212–225.

[6] R. N. Bhattacharya, and M. Qumsiyeh, *Second order and L^p-comparisons between the bootstrap and empirical Edgeworth expansion methodologies.*, Ann. Statist. **17** (1989), 160–169.

[7] P. J. Bickel, and D. A. Freedman, *Asymptotic normality and the bootstrap in stratified sampling*, Ann. Statist. **12** (1984), 470–482.

[8] P. J. Bickel, F. Götze, and W. R. van Zwet, *Resampling fewer than n observations: gains, losses, and remedies for losses*, Statist. Sinica **7** (1997), 1–31.

[9] M. Bloznelis, and F. Götze, *One term Edgeworth expansion for finite population U–statistics of degree two*, Acta Applicandae Mathematicae **58** (1999), 75–90.

[10] M. Bloznelis, and F. Götze, *Orthogonal decomposition of finite population statistics and its applications to distributional asymptotics*, Ann. Statist. **29** (2001).

[11] M. Bloznelis, and F. Götze, *An Edgeworth expansion for finite population symmetric statistics* (2000), Preprint 00–042, SFB 343 Universität Bielefeld.

[12] J. G. Booth, R. W. Butler, and P. Hall, *Bootstrap methods for finite populations*, J. Amer. Statist. Assoc. **89** (1994), 1282–1289.

[13] J. G. Booth, and P. Hall, *An improvement of the jackknife distribution function estimator*, Ann. Statist. **21** (1993), 1476–1485.

[14] M. T. Chao, and S. H. Lo, *A bootstrap method for finite population*, Sankhya Ser. A **47** (1985), 399–405.

[15] P. Hall, *The bootstrap and Edgeworth expansion*, Springer, New York, 1992.

[16] R. Helmers, *On the Edgeworth expansion and the bootstrap approximation for a Studentized U-statistic*, Ann. Statist. **19** (1991), 470–484.

[17] W. Hoeffding, *A class of statistics with asymptotically normal distribution*, Ann. Math. Statist. **19** (1948), 293–325.

[18] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Stat. Assoc. **58** (1963), 13–30.

[19] R. G. Miller, *The jackknife—a review*, Biometrika **61** (1974), 1–15.

[20] W. C. Parr, *Jackknifing differentiable statistical functionals*, J. Roy. Statist. Soc. Ser. B **47** (1985), 56–66.

[21] D. N. Politis, and J. P. Romano, *Large sample confidence regions based on subsamples under minimal assumptions*, Ann. Statist. **22** (1994), 2031–2050.

[22] H. Putter, *Consistency of Resampling Methods.*, PhD Thesis, Leiden University, 1994.

[23] H. Putter, and W. R. van Zwet, *Empirical Edgeworth expansions for symmetric statistics*, Ann. Statist. **26** (1998), 1540–1569.

[24] B. Rosén, *Limit theorems for sampling from finite populations*, Arch. Mat. **5** (1964), 383–424.

[25] J. Shao, *The efficiency and consistency of approximations to the jackknife variance estimators*, J. Amer. Statist. Assoc. **84** (1989), 114–119.

[26] J, Shao, and D. Tu, *The jackknife and bootstrap*, Springer, New York, 1995.

[27] J. Shao, and C.-F. J. Wu, *A general theory for jackknife variance estimation*, Ann. Statist. **17** (1989), 1176–1197.

[28] C.-F. J. Wu, *On the asymptotic properties of the jackknife histogram*, Ann. Statist. **18** (1990), 1438–1452.