EDGEWORTH EXPANSIONS FOR STUDENTIZED VERSIONS OF SYMMETRIC FINITE POPULATION STATISTICS

M. Bloznelis

Abstract. We show the validity of the one-term Edgeworth expansion for Studentized asymptotically linear statistics based on samples drawn without replacement from finite populations. Replacing the moments defining the expansion by their estimators, we obtain an empirical Edgeworth expansion. We show the validity of the empirical Edgeworth expansion in probability.

1. Introduction and Results

Here we give complete proofs of the results announced in the First Baltic-Nordic conference on survey sampling and presented in [5].

1. Let $T = t(X_1, \dots, X_n)$ be a real-valued statistic based on the sample $X_1, \dots, X_n$ drawn without replacement from a finite population $\mathcal{X} = \{x_1, \dots, x_N\}$. Write

$$p = n/N, \qquad q = 1 - p, \qquad n_* = \min\{n,\, N-n\}.$$

Let σ_T^2 denote the variance of T and let

$$S^2 = S^2(T) = q \sum_{j=1}^{n+1} \bigl(T_{(j)} - \overline{T}\bigr)^2, \qquad \overline{T} = \frac{1}{n+1} \sum_{j=1}^{n+1} T_{(j)},$$

denote the jackknife estimator of variance based on sample $X_1, \dots, X_{n+1}$ (of size $n+1$) drawn without replacement from $\mathcal{X}$. Here $T_{(j)} = t(X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_{n+1})$.

In the simplest case of a linear statistic $L = g(X_1) + \dots + g(X_n)$ (here $g : \mathcal{X} \to R$), the asymptotic normality as $n_* \to \infty$ of $(L - \mathbf{E}\, L)/S(L)$ follows from the central limit theorem combined with the law of large numbers. Edgeworth expansions were constructed by Babu and Singh [1], see also Sugden, Smith and Jones [15].

1991 *Mathematics Subject Classification.* Primary 62E20; secondary 60F05.

Key words and phrases. sampling without replacement, finite population, U-statistic, Studentized statistic, asymptotic expansion, Edgeworth expansion, jackknife.

Many important statistics are asymptotically linear as n increases. Consequently, their Studentized versions $(T - \mathbf{E}\,T)/S$ are asymptotically standard normal. In order to treat general asymptotically linear statistics, we use linearization by means of the orthogonal decomposition. This kind of the decomposition of statistics was first used by Hoeffding [10] in the case of *independent and identically distributed* observations. The orthogonal decomposition of symmetric statistics based on samples *drawn without replacement* was studied by Bloznelis and Götze [7], see also Zhao and Chen [16].

We shall assume, in what follows, that T is symmetric. That is, the kernel t is invariant under permutations of its arguments, i.e., $t(y_1, \dots, y_n) = t(y_{\pi(1)}, \dots, y_{\pi(n)})$ for any permutation π of indices $1, 2, \dots, n$. Note that the sample mean, sample variance, sample quantiles, U-statistics, L-statistics, and many other statistics are symmetric. Given a symmetric statistic T, it is decomposed into the sum of centered and *uncorrelated* $U-$statistics of increasing order

$$T = \mathbf{E}\,T + \sum_{1\le i\le n} g_1(X_i) + \sum_{1\le i<j\le n} g_2(X_i, X_j) + \dots. \tag{1.1}$$

The first sum

$$L = \sum_{1\le i\le n} g_1(X_i), \qquad g_1(X_i) = \frac{N-1}{N-n}\,\mathbf{E}\,(T - \mathbf{E}\,T | X_i), \tag{1.2}$$

is called the linear part of T. The second sum $Q = \sum_{i<j} g_2(X_i, X_j)$ is called the quadratic part. Here, for $i \neq j$,

$$g_2(X_i, X_j) = \frac{N-3}{N-n-1}\Big(\frac{N-2}{N-n}\,\mathbf{E}\,(T - \mathbf{E}\,T \big| X_i, X_j) - g_1(X_i) - g_1(X_j)\Big). \tag{1.3}$$

The random variables $g_1(X_k)$ and $g_2(X_i, X_j)$ are centered and uncorrelated for arbitrary $1 \le i, j, k \le n$, $i \neq j$. For a detailed description of the decomposition we refer to [7].

We shall assume that the linear part does not vanish, i.e., $\sigma_1^2 > 0$, where $\sigma_1^2 = \mathbf{Var}\, g_1(X_1)$. Note that the variance of the linear part equals

$$\mathbf{Var}\, L = \tau^2 \sigma_1^2 N/(N-1), \qquad \text{where} \qquad \tau^2 = Npq.$$

Furthermore, $n_*/2 \le \tau^2 \le n_*$.

If, for large n, the linear part dominates the statistic, we call T asymptotically linear. In this case, Bloznelis and Götze [8] showed that the one-term Edgeworth expansion

$$G(x) = \Phi(x) - \frac{(q-p)\alpha + 3\kappa}{6\tau}\,\Phi'(x)(x^2 - 1) \tag{1.4}$$

approximates the distribution function $F(x) = \mathbf{P}\{T - \mathbf{E}\,T \le \sigma_T x\}$ up to the error $o(n_*^{-1/2})$. The moments

$$\alpha = \sigma_1^{-3}\mathbf{E}\,g_1^3(X_1), \qquad \kappa = \sigma_1^{-3}\tau^2\mathbf{E}\,g_2(X_1, X_2)g_1(X_1)g_1(X_2) \tag{1.5}$$

refer to the linear and the quadratic part of the decomposition.

We shall show in Theorem 2 below that the one-term Edgeworth expansion

$$H(x) = \Phi(x) + \frac{(q - p + (q+1)x^2)\alpha + 3(x^2+1)\kappa}{6\tau}\,\Phi'(x) \tag{1.6}$$

approximates the distribution function of the Studentized statistic

$$F_S(x) = \mathbf{P}\{T - \mathbf{E}\,T \le S(T)x\}$$

up to the error $o(n_*^{-1/2})$.

Note that in order to write expansion (1.6), one does not need to evaluate all terms of decomposition (1.1), but the moments (1.5) only. Furthermore, these moments can be estimated.

Let us define the jackknife estimators. In what follows, $\{X_1, \dots, X_m\}$, for $m = n, n+1, n+2$, 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 the first m variables of the random permutation $(X_1, \dots, X_N)$ of the ordered set $(x_1, \dots, x_N)$. For $1 \le k \le n+1$, $1 \le i, j, r \le n+2$, $i \ne j$, 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

$$\overline{T}_{(r)} = \frac{1}{n+1}\sum_{1 \le j \le n+2,\, j \ne r} T_{(r,j)}, \qquad \tilde{T} = \frac{1}{\binom{n+2}{2}}\sum_{1 \le i < j \le n+2} T_{(i,j)}.$$

Here $T_{(i,j)}$ denotes the value of t at the sample $\{X_1, \dots, X_{n+2}\} \setminus \{X_i, X_j\}$. Write

$$\hat{\alpha}_J = \frac{\sqrt{n}}{\hat{\sigma}_J^3}\sum_{k=1}^{n+1} V_k^3, \qquad \hat{\kappa} = q\,\frac{2\sqrt{n}}{\hat{\sigma}_J^3}\sum_{1 \le i < j \le n+2} W_{ij}\tilde{V}_i\tilde{V}_j, \tag{1.7}$$

where $\hat{\sigma}_J^2 = \sum_{k=1}^{n+1} V_k^2$.

Bloznelis [4] showed that $\hat{\alpha}$, $\hat{\kappa}$, and $S^2(T)\,(= q\hat{\sigma}_J^2)$ are consistent estimators of α, κ, and σ_T^2 as $n \to \infty$. Using this fact, we show in Theorem 3 below that the empirical Edgeworth expansion

$$\hat{H}(x) = \Phi(x) + \frac{(q - p + (q+1)x^2)\hat{\alpha} + 3(x^2+1)\hat{\kappa}}{6\tau}\,\Phi'(x) \tag{1.8}$$

approximates $F_S(x)$ up to the error $o(n_*^{-1/2})$ in probability.

One-term Edgeworth expansions for Studentized U-statistics based on *independent and identically distributed* observations were constructed by Helmers [9]. Expansions for Studentized versions of general symmetric statistics were obtained by Putter and van Zwet [14], see also Putter [13], Bentkus, Götze and van Zwet [3]. Empirical Edgeworth expansions that use jackknife estimators were studied by Beran [3]. Putter and van Zwet [14] constructed such expansions for general symmetric statistics and their Studentized versions.

One-term Edgeworth expansion for U-statistics of degree two based on samples *drawn without replacement* was constructed by Kokic and Weber [12]. Bloznelis and Götze [8] established the validity of the one-term Edgeworth expansion for general symmetric finite population statistics. The corresponding empirical Edgeworth expansions were constructed by Bloznelis [4]. Since often the variance of the underlying statistic (estimator) is unknown, it is important, for practical purposes, to have such approximations for Studentized versions of statistics too. This question is addressed in the present paper.

2. Results. In order to formulate asymptotic results for finite population statistics, we introduce a sequence of populations $\mathcal{X}_\nu = \{x_{\nu\,1}.\dots,x_{\nu\,N_\nu}\}$, $\nu = 1, 2, \dots$, and a sequence of symmetric statistics $T_\nu = t_\nu(X_{\nu\,1},\dots,X_{\nu\,n_\nu})$. Here $\{X_{\nu\,1},\dots,X_{\nu\,n_\nu}\}$ denotes a sample drawn without replacement from $\mathcal{X}_\nu$. The orthogonal decomposition expands T_ν into the sum of uncorellated centered U-statistics (see [7])

$$(1.9)\quad T_\nu = \mathbf{E}\,T_\nu + U_{\nu\,1} + \cdots + U_{\nu\,n_{\nu *}}, \qquad U_{\nu\,k} = \sum_{1\le i_1<\dots<i_k\le n_\nu} g_{\nu\,k}(X_{\nu\,i_1},\dots X_{\nu\,i_k}).$$

Here $n_{\nu\,*} = \min\{N_\nu - n_\nu,\, n_\nu\}$. Let $L_\nu = U_{\nu\,1}$ and $Q_\nu = U_{\nu\,2}$ denote the linear and the quadratic part, respectively. Furthermore, let us denote $\sigma_{\nu\,1}^2 = \mathbf{Var}\,g_{\nu\,1}^2(X_{\nu\,1})$,

$$(1.10)\qquad \beta_{\nu\,s} = \sigma_{\nu\,1}^{-s}\mathbf{E}\,|g_{\nu\,1}(X_{\nu\,1})|^s, \qquad \gamma_{\nu\,s} = \sigma_{\nu\,1}^{-s}\tau_\nu^{2s}\mathbf{E}\,|g_{\nu\,2}(X_{\nu\,1},X_{\nu\,2})|^s,$$

where $s > 0$ and $\tau_\nu = N_\nu p_\nu q_\nu$, $p_\nu = n_\nu/N_\nu$, $q_\nu = 1 - p_\nu$. Let α_ν, κ_ν, and $\hat\alpha_\nu$, $\hat\kappa_\nu$ denote the moments defined by (1.5) and their jackknife estimators (1.7), respectively. Furthermore, let $S_\nu^2 = S_\nu^2(T_\nu)$ denote the jackknife estimator of the variance $\sigma_{\nu\,T}^2$ of T_ν. Write $\Psi_\nu(t) = \mathbf{E}\,\exp\{it\sigma_{\nu\,1}^{-1}g_{\nu\,1}(X_{\nu\,1})\}$.

We shall assume that $n_{\nu\,*}$ tends to infinity as $\nu \to \infty$ and construct bounds for

$$\Delta_\nu = \sup_x |F_{\nu\,S}(x) - H_\nu(x)| \qquad \text{and} \qquad \hat\Delta_\nu = \sup_x |F_{\nu\,S}(x) - \hat H_\nu(x)|.$$

Here $F_{\nu\,S}(x) = \mathbf{P}\{(T_\nu - \mathbf{E}\,T_\nu)/S_\nu(T_\nu) \le x\}$. The functions $H_\nu(x)$ and $\hat H_\nu(x)$ are defined by (1.6) and (1.8), but with α_ν, κ_ν and $\hat\alpha_\nu$, $\hat\kappa_\nu$ instead of α, κ etc.

Firstly, we consider a special case of U-statistics. We write $T_\nu = U_\nu$, where

$$U_\nu = \sum_{1\le i<j\le n_\nu} h_\nu(X_{\nu\, i}, X_{\nu\, j}). \tag{1.11}$$

In this case, decomposition (1.9) reduces to $U_\nu = \mathbf{E}\, U_\nu + L_\nu + Q_\nu$. The kernels defining the linear and quadratic parts are obtained from (1.2) and (1.3):

$$\begin{aligned} g_{\nu\,1}(x) &= (n_\nu - 1)\,\frac{N_\nu - 1}{N_\nu - 2}\,\mathbf{E}\,\Big(h_\nu(X_{\nu\,1}, X_{\nu\,2}) - \mathbf{E}\, h_\nu(X_{\nu\,1}, X_{\nu\,2})\Big| X_{\nu\,1} = x\Big),\\ g_{\nu\,2}(x,y) &= h_\nu(x,y) - \mathbf{E}\, h_\nu(X_{\nu\,1}, X_{\nu\,2}) - (n_\nu - 1)^{-1}(g_{\nu\,1}(x) + g_{\nu\,1}(y)). \end{aligned}$$

Theorem 1. *Let T_ν be a U-statistic of the form (1.11). Assume that $n_{\nu\,*}\to\infty$ as $\nu\to\infty$. Assume that there exist absolute constants $s>6$, $C_1>0$, a positive continuous function ϕ on $(0,+\infty)$, and sequences $\{\xi_\nu\}\uparrow\infty$ and $\{\eta_\nu\}\uparrow\infty$ such that for $\nu = 1, 2, \dots,$*

$$\beta_{\nu\, s} \le C_1, \qquad \gamma_{\nu\, s}\le C_1, \tag{1.12}$$

$$|\Psi(t)| \le 1 - \phi(|t|), \qquad \text{for} \quad 0<|t|\le \eta_\nu, \tag{1.13}$$

$$n_\nu \le N_\nu - \xi_\nu N_\nu^{2/3}. \tag{1.14}$$

Then there exists a sequence $\{\psi_\nu\}\downarrow 0$ depending only on C_1, ϕ, $\{\xi_\nu\}$, and $\{\eta_\nu\}$ such that, for every $\nu = 1, 2\dots,$

$$\Delta_\nu \le \psi_\nu \tau_\nu^{-1}, \tag{1.15}$$

$$\mathbf{P}\big\{\hat\Delta_\nu > \psi_\nu\tau_\nu^{-1}\big\} \le \psi_\nu. \tag{1.16}$$

Remark 1. Under the moment condition (1.12), the nonlattice condition (1.13) and (1.14), inequality (1.15) (respectively, (1.16)) establishes the bound as $n_{\nu *}\to\infty$

$$\Delta_\nu = o(n_{\nu *}^{-1/2}) \qquad (\text{respectively,} \qquad \hat\Delta_\nu = o_P(n_{\nu *}^{-1/2})). \tag{1.17}$$

The latter bound holds in probability. Here $n_{\nu *}$ plays the same role as the sample size does in the i.i.d. situation, see [8].

Remark 2. Let us note that the nonlattice condition (1.13) is the weakest possible smoothness condition. Condition (1.14) is very mild. The moment condition (1.12) is far from the optimal one. Here one would expect the uniform integrability of $\beta_{\nu\,3}$ and $\gamma_{\nu\,5/3}$, for $\nu = 1, 2, \dots,$ instead of (1.12). In the proof, no effort was made to obtain result (1.17) under the optimal moment conditions. A modification of our proof involving truncation would probably reduce the moment condition (1.12) to $\beta_{\nu\,s} < C_1$ and $\gamma_{\nu\,t} < C_1$, for $s > 3$ and $t > 2$, cf. [6].

Let us consider a general symmetric statistic T_ν. Using (1.9), we write

$$T_\nu = U_\nu + R_\nu, \tag{1.18}$$

where $U_\nu = \mathbf{E}\,T_\nu + L_\nu + Q_\nu$ is a finite population U-statistic of degree two and where the remainder $R_\nu = \sum_{j\ge 3} U_{\nu\,j}$. For a typical standardized asymptotically linear statistic $\sigma_{\nu T}^{-1} T_\nu$ admitting one-term expansion we have as $n_{\nu *} \to \infty$

$$\sigma_{\nu T}^{-1} R_\nu = o_P(n_{\nu *}^{-1/2}) \qquad \text{and} \qquad S_\nu(T_\nu)/S_\nu(U_\nu) = 1 + o_P(n_{\nu *}^{-1/2}) \tag{1.19}$$

in probability. Consequently, one-term expansions for $(T_\nu - \mathbf{E}\,T_\nu)/S_\nu(T_\nu)$ and $(U_\nu - \mathbf{E}\,U_\nu)/S_\nu(U_\nu)$ are, in fact, the same.

Bloznelis and Götze [7] introduced simple conditions that ensure the validity of approximations like (1.19). These conditions are formulated in terms of moments of differences. Recall that $(X_{\nu\,1}, \dots, X_{\nu\,N_\nu})$ denotes a random permutation of the ordered population $(x_{\nu\,1}, \dots, x_{\nu\,N_\nu})$. For $j < n_{\nu\,*}$, define

$$D^j T_\nu = t_\nu(X_{\nu 1}, \dots, X_{\nu n}) - t_\nu(X_{\nu 1}, \dots, X_{\nu j-1},\, X_{\nu j+1}, \dots, X_{\nu n},\, X_{\nu n+j}).$$

Higher order differences are defined recursively: $D^{ij} T_\nu = D^j\big(D^i T_\nu\big)$, for $i \ne j$; $D^{ijk} T_\nu = D^k\big(D^j(D^i T_\nu)\big)$, for $i \ne j \ne k$; $\dots$. Write

$$\delta_{\nu\,j} = \delta_{\nu\,j}(T_\nu) = \mathbf{E}\,\big(n_{\nu *}^{(j-1)} \mathbb{D}_j T_\nu\big)^2, \qquad \mathbb{D}_j T_\nu := D^{12\dots j} T_\nu.$$

Let us note that for U-statistics, and statistics which are smooth functions of sample means, the moments $\delta_{\nu\,j}$ can be easily estimated, see [7]. In particular, for a typical asymptotically standard normal statistic one can expect $\delta_{\nu\,j} = O(n_*^{-1})$, $1 \le j \le k$, for some $k \ge 2$.

Theorem 2. *The statement (1.15) holds true for a sequence of general symmetric statistics $\{T_\nu\}$ if, in addition to (1.12), (1.13), and (1.14), we assume that*

$$\delta_{\nu\,3} \le \varepsilon_\nu n_{\nu\,*}^{-2/3} \sigma_{\nu\,T}^2, \tag{1.20}$$

for some decreasing sequence $\{\varepsilon_\nu\} \downarrow 0$ as $\nu \to \infty$.

Theorem 3. *The statement (1.16) holds true for a sequence of general symmetric statistics $\{T_\nu\}$ if, in addition to (1.12), (1.13), and (1.14), we assume that*

$$\delta_{\nu\,2} \le \varepsilon_\nu n_\nu^{-1/3} \sigma_{\nu\,T}^2, \qquad \delta_{\nu\,3} \le \varepsilon_\nu n_{\nu\,*}^{-2/3} \sigma_{\nu\,T}^2,$$

for some decreasing sequence $\{\varepsilon_\nu\} \downarrow 0$ as $\nu \to \infty$.

Theorems 2 and 3 establish bounds (1.17) for general symmetric statistics.

2. Proofs

This section is organized as follows. Firstly, we prove Theorem 1. Then we prove Theorem 2. Theorem 3 is a simple consequence of Theorem 2 and the consistency results for $S^2_\nu(T_\nu)$, $\hat\alpha_\nu$, and $\hat\kappa_\nu$ established in Lemmas 2.1-2.3 (see [4]) Bloznelis (2001). Auxiliary results (Lemmas 1-5) are formulated at the end of the section.

In the proofs we shall use the variance decomposition (see formula (2.6) [7])

$$\mathbf{Var}\, T_\nu = \sum_{k=1}^{n_{\nu *}} \mathbf{Var}\, U_{\nu\, k}, \qquad \mathbf{Var}\, U_{\nu k} = \frac{\binom{n_\nu}{k}\binom{N_\nu - n_\nu}{k}}{\binom{N_\nu - k}{k}}\,\sigma^2_{\nu\, k}. \tag{2.1}$$

Here we denote $\sigma^2_{\nu\, k} = \mathbf{Var}\, g_{\nu k}(X_{\nu\, 1}, \dots, X_{\nu\, k})$.

A sequence of random variables $\{M_\nu,\ \nu = 1, 2, \dots\}$ is said to satisfy the condition (2.2) if, for $\nu = 1, 2, \dots,$

$$\mathbf{P}\{|M_\nu| > \tilde\psi_\nu \tau_\nu^{-1}\} \le \tilde\psi_\nu \tau_\nu^{-1}, \tag{2.2}$$

for some nonrandom sequence $\{\tilde\psi_\nu\} \downarrow 0$ which depends only on C_1 and $\{\xi_\nu\}$ in the proof of Theorem 1 (respectively, C_1, $\{\xi_\nu\}$, and $\{\varepsilon_\nu\}$ in the proof of Theorem 2).

In order to simplify the notation, we shall write $o(n^s_{\nu *})$ to denote the sequence $\tilde\psi_\nu n^s_{\nu *}$, where $\{\tilde\psi_\nu\} \downarrow 0$ is a sequence depending only on the constant C_1, the function ϕ, and the sequences $\{\xi_\nu\}$ and $\{\eta_\nu\}$ (and also the sequence $\{\varepsilon_\nu\}$ in the proof of Theorem 2). Furthermore, we shall drop the subscript ν whenever this does not cause an ambiguity. By c we denote a constant which depends only on C_1. Write, for brevity, $Y_{\nu i} = g_{\nu 1}(X_{\nu i})$ and $Y_{\nu ij} = g_{\nu 2}(X_{\nu i}, X_{\nu j})$.

Proof of Theorem 1. Let us denote $\sigma^2_{\nu U} = \mathbf{Var}\, U_\nu$ and assume without loss of generality that $\mathbf{E}\, U_\nu = 0$ and $\mathbf{Var}\, U_\nu = 1$ for $\nu = 1, 2, \dots$.

Note that (1.16) follows from (1.15) and the following consistency results established in Lemmas 2.1-3 of [4]: as $\nu \to \infty$

$$|S^2_\nu - \sigma^2_{\nu U}| = o_P(1), \quad |\hat\alpha_\nu - \alpha_\nu| = o_P(1), \quad \text{and} \quad |\hat\kappa_\nu - \kappa_\nu| = o_P(1).$$

Let us mention that Lemma 2.2 (ibidem), which establishes the consistency result for $\hat\alpha_\nu$, assumes, in addition, that $\delta_2(U_\nu) = o(n_\nu^{-1/3})$. For U statistics of degree two (and such that $\mathbf{Var}\, U = 1$) this relation is implied by the second inequality of (1.12), see Lemma 4.

Let us prove (1.15). It follows from (2.1) that $\sigma^2_U q = \mathbf{Var}\, L + \mathbf{Var}\, Q$,

$$\mathbf{Var}\, L = \frac{\binom{n}{1}\binom{N-n}{1}}{\binom{N-1}{1}}\,\sigma_1^2, \qquad \mathbf{Var}\, Q = \frac{\binom{n}{2}\binom{N-n}{2}}{\binom{N-2}{2}}\,\sigma_2^2. \tag{2.3}$$

By assumption that $\sigma_U^2 = 1$, we have $\mathbf{Var}\, L \le 1$. Therefore, (2.3) implies $\sigma_1^2 \le \tau^{-2}$. Invoking Hölder's inequality $\gamma_2 \le \gamma_6^{1/3}$, we obtain from (1.12)

$$\sigma_2^2 = \gamma_2 \sigma_1^2 \tau^{-4} \le C_1^{1/3} \sigma_1^2 \tau^{-4} \le c\tau^{-6}. \tag{2.4}$$

The identities $\sigma_U^2 = 1$ and (2.3) combined with (2.4) show that

$$1 - c\tau^{-2} \le \mathbf{Var}\, L \le 1 \qquad \text{and} \qquad \tau^{-2} \ge \sigma_1^2 \ge \tau^{-2} - c\tau^{-4}. \tag{2.5}$$

We split the remaining part of the proof into three parts.

Step 1. In this step, we show that, for every $\nu = 1, 2, \dots,$

$$S_\nu^2 = 1 + L_S + M_\nu, \qquad L_S = \sum_{j=1}^{n_\nu+1} f(X_{\nu i}), \tag{2.6}$$

$$f(X_{\nu i}) = q_\nu \bigl(1 - \frac{1}{n_\nu + 1}\bigr)(Y_{\nu i}^2 - \sigma_{\nu 1}^2) + 2 q_\nu n_\nu \, \frac{N_\nu - 1}{N_\nu - 2}\, \mathbf{E}\, (Y_{\nu j} Y_{\nu i j} | X_{\nu i}).$$

Here the sequence of random variables $\{M_\nu\}$ satisfies (2.2).

To show (2.6) and (2.2), fix k and split $V_k = \overline{U} - U_{(k)} = Z_k + W_k$, where

$$Z_k = \sum_{j=1}^{n+1} Y_j \bigl(\mathbb{I}_{\{j=k\}} - \frac{1}{n+1}\bigr), \quad W_k = \sum_{1 \le i < j \le n+1} Y_{ij} \bigl(\mathbb{I}_{\{k \in \{i,j\}\}} - \frac{2}{n+1}\bigr),$$

and write

$$S^2 \ (= S_\nu^2) = q \sum_k V_k^2 = q \sum_k Z_k^2 + 2q \sum_k Z_k W_k + q \sum_k W_k^2. \tag{2.7}$$

A calculation shows

$$\sum_k Z_k^2 = \bigl(1 - \frac{1}{n+1}\bigr) \sum_{k=1}^{n+1} Y_k^2 - \frac{2}{n+1} \sum_{1 \le i < j \le n+1} Y_i Y_j,$$

$$\sum_k Z_k W_k = \sum_{1 \le i < j \le n+1} (Y_i + Y_j) Y_{ij} - \frac{2}{n+1}\, H_1 H_2,$$

$$\sum_k W_k^2 = 2 \sum_{1 \le i < j \le n+1} Y_{ij}^2 + 2H_3 - \frac{8}{n+1}\, H_2^2 + \frac{4}{(n+1)^2}\, H_2^2.$$

Here $H_1 = \sum_{1 \le k \le n+1} Y_k$ and $H_2 = \sum_{1 \le i < j \le n+1} Y_{ij}$, and

$$H_3 = \sum_{1 \le i \le n+1} \ \sum_{1 \le j < k \le n+1} Y_{ij} Y_{ik} \mathbb{I}_{\{i \notin \{j,k\}\}}.$$

Finally, from (2.7) we obtain (2.6) with $M_\nu = m_1 + q(m_2 + \cdots + m_8)$. Here

$$m_1 = nq\sigma_1^2 - 1, \qquad m_2 = -\frac{2}{n+1} \sum_{1\le i<j\le n+1} Y_i Y_j,$$
$$m_3 = -\frac{4}{n+1} H_1 H_2, \qquad m_4 = 2(H - L_H),$$

where

$$H = \sum_{1\le i<j\le n+1} (Y_i + Y_j) Y_{ij} \qquad \text{and} \qquad L_H = \sum_{1\le i\le n+1} n\,\frac{N-1}{N-2}\,\mathbf{E}\,(Y_j Y_{ij} | X_i),$$

and

$$m_5 = 2 \sum_{1\le i<j\le n+1} Y_{ij}^2, \quad m_6 = 2H_3, \quad m_7 = -\frac{8}{n+1} H_2^2, \quad m_8 = \frac{4}{(n+1)^2} H_2^2.$$

In order to prove that $\{M_\nu\}$ satisfies (2.2) we show that

$$\begin{aligned} &\mathbf{E}\,H_1^2 < 2, \qquad \mathbf{E}\,H_2^2 \le c\tau^{-2}, \qquad |\mathbf{E}\,H_3| \le cq^{-2}\tau^{-2}, \\ &|m_1| \le c\tau^{-2}, \qquad |\mathbf{E}\,m_2| \le \tau^{-2} \qquad \mathbf{E}\,m_5 \le cq^{-2}\tau^{-2}, \end{aligned} \tag{2.8}$$
$$\begin{aligned} &\mathbf{Var}\,m_2 \le cn^{-2}, \qquad \mathbf{Var}\,m_5 \le cq^{-2}\tau^{-6}, \\ &\mathbf{Var}\,(H - L_H) \le c\tau^{-4}, \qquad \mathbf{Var}\,H_3 \le cq^{-2}\tau^{-4}. \end{aligned} \tag{2.9}$$

It is easy to show that bounds (2.8) and (2.9) imply (2.2), provided that $q_\nu \tau_\nu \to \infty$ as $\nu \to \infty$. The latter condition is equivalent to (1.14).

The bounds for expectations (2.8) are simple consequences of (2.1), (2.3), (2.4), and (2.5). In order to bound variances (2.9), we decompose the random variables m_2, m_5, $H - L_H$ (which are U-statistics of degree two), and H_3 (which is a U-statistic of degree three) by means of (1.9) and use identity (2.1), see Lemma 5 below.

Step 2. Set $\tilde L_S = \sum_{i=1}^n f(X_i)$ and write $\tilde S^2 = 1 + \tilde L_S$ for $\tilde L_S > -1$. Put $\tilde S^2 = 2$ for $\tilde L_S \le -1$. Let us denote $\tilde\Delta = \sup_x \tilde\Delta(x)$, where $\tilde\Delta(x) = |\mathbf{P}\{U/\tilde S \le x\} - H(x)|$. Using Chebyshev's inequality, we obtain

$$\begin{aligned} \mathbf{P}\{|\tilde S - 1| > 1/2\} &\le \mathbf{P}\{|\tilde L_S| > 1/2\} \le 4\mathbf{E}\,\tilde L_S^2 \\ &= 4\tau^2 N(N-1)^{-1}\mathbf{E}\,f^2(X_1) \\ &\le c\,n_*^{-1}. \end{aligned} \tag{2.10}$$

In particular, we have $\tilde S^2 > 0$ with probability $1 - o(n_*^{-1/2})$. Now, (2.6) and the identity $\sigma_U^2 = 1$ imply

$$S_\nu/\tilde S = \sqrt{1 + (M_\nu + f(X_{n+1}))/\tilde S^2} = 1 + \tilde M_\nu. \tag{2.11}$$

By the Lagrange mean-value theorem, from (2.10) and the fact that $\{M_\nu\}$ satisfies (2.2) we obtain that $\{\tilde M_\nu\}$ satisfies (2.2). Since $xH'(x)$ is bounded, this implies

$$\Delta = \sup_x |\mathbf{P}\{\frac{U}{\tilde S} \le x \frac{S_\nu}{\tilde S}\} - H(x)| \le \tilde\Delta + o(n_*^{-1/2}).$$

In what follows, we construct the bound $\tilde\Delta = o(n_*^{-1/2})$.

The bound $\sup\{\tilde\Delta(x) : |x| \ge \log n_*\} = o(n_*^{-1/2})$ is a consequence of the following result (see Theorem 2 of [7], or Theorem 2.1 of [4])

$$\sup_x |P\{\frac{U}{\sigma_U} \le x\} - G(x)| = o(n_*^{-1/2}) \tag{2.12}$$

and the fact that $|G(x)|$ and $|H(x)|$ decay exponentially as $|x| \to \infty$. Here we also use $\sigma_U^2 = 1$ and (2.10).

Let $|x| < \log n_*$. For $|\tilde L_S| \le 4/5$ we use the inequalities $A \le (1+\tilde L_S)^{1/2} \le B$, where $A = 1 + \tilde L_S/2 - \tilde L_S^2/4$ and $B = 1 + \tilde L_S/2$. Since, by (2.10), $\mathbf{P}\{|\tilde L_S| > 4/5\} = o(n_*^{-1/2})$, we obtain $\tilde\Delta(x) \le \max\{\Delta_A, \Delta_B\} + o(n_*^{-1/2})$, where

$$\Delta_A = \sup_{|x|\le \log n_*} |\mathbf{P}\{U \le Ax\} - H(x)|, \qquad \Delta_B = \sup_{|x|\le \log n_*} |\mathbf{P}\{U \le Bx\} - H(x)|.$$

Using Chebyshev's inequality one can show, see Lemma 2 below, that

$$\mathbf{P}\{|A-B| > n_*^{-8/15}\} \le c n_*^{-8/15}. \tag{2.13}$$

Since $xH'(x)$ is bounded, (2.13) implies $\Delta_A \le \Delta_B + o(n_*^{-1/2})$.

Step 3. Here we show that $\Delta_B = o(n_*^{-1/2})$. Write $\mathbf{P}\{U \le Bx\} = \mathbf{P}\{U_x \le x\}$, where $U_x = U - 2^{-1}x\tilde L_S$ is a U-statistic. We are going to apply (2.12) to U_x.

Decompose U_x by means of (1.9), $U_x = L_x + Q_x$, where

$$L_x = \sum_{i=1}^n \varphi(X_i), \qquad \varphi(X_i) = Y_i - 2^{-1}xf(X_i),$$

is the linear part and $Q_x = \sum_{1\le i<j\le n} Y_{ij}$ is the quadratic part. By means of (2.1), we decompose the variance $\sigma_x^2 := \mathbf{Var}\, U_x$ as follows:

$$\begin{aligned}\sigma_x^2 &= \mathbf{Var}\, L_x + \mathbf{Var}\, Q_x = \sigma_U^2 + (\mathbf{Var}\, L_x - \mathbf{Var}\, L) \\ &= 1 + \tau^2 \frac{N}{N-1}(\sigma_{x1}^2 - \sigma_1^2).\end{aligned} \tag{2.14}$$

Here L denotes the linear part of U and $\sigma_{x\,1}^2 := \mathbf{E}\,\varphi^2(X_1)$. We have

$$\sigma_{x\,1}^2 - \sigma_1^2 = -x\mathbf{E}\, Y_i f(X_i) + 4^{-1}x^2 \mathbf{E}\, f^2(X_i) = -q\sigma_1^3\alpha_3 x - 2\sigma_1^3\kappa x + R. \tag{2.15}$$

Using (1.12), (2.5), and the inequality $|x| \le \log n_*$, we bound the remainder $R = o(n_*^{-3/2})$. This bound, in combination with (2.14) and (2.5), yields

$$\sigma_x^2 = 1 - \tau^{-1}(q\alpha_3 + 2\kappa)x + o(n_*^{-1/2}) \tag{2.16}$$

uniformly in $|x| \le \log n_*$.

Let α_x, κ_x, and $\beta_{s,x}$, $\gamma_{s,x}$ denote moments (1.5) and (1.10) corresponding to the statistic U_x. Let G_x denote expansion (1.4) of U_x (replace α, κ by α_x, κ_x in (1.4)). It follows from (1.12) and (2.15) that $\alpha_x = \alpha + o(1)$ and $\kappa_x = \kappa + o(1)$ uniformly in $|x| \le \log n_*$ as $n_* \to \infty$. This yields

$$\sup_{|x| \le \log n_*} \sup_{y \in R} |G_x(y) - G(y)| = o(n_*^{-1/2}). \tag{2.17}$$

Furthermore, Theorem 2.1 of [4] shows that

$$\sup_{y \in R} |\mathbf{P}\{U_x \le y\sigma_x\} - G_x(y)| = o(n_*^{-1/2}) \tag{2.18}$$

uniformly in $|x| \le \log n_*$, provided that for some δ and $\nu_0 > 0$:

(i) the moments $\beta_{3+\delta,x}$ and $\gamma_{2+\delta,x}$ are uniformly bounded for $|x| \le \log n_*$ and $\nu \ge \nu_0$;

(ii) for some sequence $\{\tilde\eta_\nu\} \uparrow \infty$ and positive continuous function $\tilde\phi$ on $(0, +\infty)$, the characteristic function $\Psi_x(t) = \mathbf{E}\exp\{it\sigma_{x\,1}^{-1}\varphi(X_i)\}$ satisfies $|\Psi_x(t)| \le 1 - \tilde\phi(|t|)$ for $0 < |t| \le \tilde\eta_\nu$ and $|x| \le \log n_*$ for every $\nu \ge \nu_0$.

Note that condition (i), for $\delta = (s-6)/2 > 0$, is implied by (1.12).

In order to verify (ii) for some $\{\tilde\eta_\nu\}$ (with $\tilde\eta_\nu \le \eta_\nu$), we show that

$$\sup_{|x| \le \log n_*} |\Psi_x(t) - \Psi(t)| \le c|t| n_*^{-1/2} \log n_*. \tag{2.19}$$

This inequality and (1.13) imply (ii). To show (2.19), write $|\Psi_x(t) - \Psi(t)| \le A + B$,

$$A = \mathbf{E}\,|\exp\{it\sigma_{x\,1}^{-1}\varphi(X_1)\} - \exp\{it\sigma_1^{-1}\varphi(X_1)\}| \le |t|\,|\sigma_{x\,1}^{-1} - \sigma_1^{-1}|\,\mathbf{E}\,|\varphi(X_1)|,$$
$$B = \mathbf{E}\,|\exp\{it\sigma_1^{-1}\varphi(X_1)\} - \exp\{it\sigma_1^{-1}Y_1\}| \le |t|\sigma_1^{-1}\mathbf{E}\,|\varphi(X_1) - Y_1|.$$

Using the inequality $\mathbf{E}\,|\varphi(X_1)| \le \sigma_{x\,1}$ and (2.15), we obtain

$$A \le |t||1 - \sigma_{x\,1}/\sigma_1| \le c|t|(|x| + o(1))n_*^{-1/2} \le c|t|n_*^{-1/2}\log n_*.$$

Invoking the simple bound $\mathbf{E}\,|\varphi(X_1) - Y_1| = 2^{-1}|x|\mathbf{E}\,|f(X_1)| \le c|x|n_*^{-1}$, we obtain $B \le c|t|n_*^{-1/2}\log n_*$, thus completing the proof of (2.19).

Finally, using (2.16) we expand

$$\begin{aligned} G\Big(\frac{x}{\sigma_x}\Big) &= G\big(x + x\big(\frac{1}{\sigma_x} - 1\big)\big) \\ &= G(x) + \Phi'(x)x\big(\frac{1}{\sigma_x} - 1\big) + o(n_*^{-1/2}) \\ &= G(x) + \Phi'(x)\,\frac{x^2}{2}\;\frac{q\alpha + 2\kappa}{\tau} + o(n_*^{-1/2}) \\ &= H(x) + o(n_*^{-1/2}) \end{aligned}$$

uniformly in $|x| \le \log n_*$. Therefore, $\sup_{|x|\le \log n_*} |G(x/\sigma_x) - H(x)| = o(n_*^{-1/2})$. This bound, in combination with (2.17) and (2.18), shows that $\Delta_B = o(n_*^{-1/2})$.

The following chain of inequalities completes the proof of Theorem 1

$$\Delta \le \tilde{\Delta} + o(n_*^{-1/2}) \le \Delta_B + o(n_*^{-1/2}) = o(n_*^{-1/2}). \tag{2.20}$$

Theorem 1 is proved.

Proof of Theorem 2. We assume, without loss of generality, that $\mathbf{E}\,T_\nu = 0$ and $\mathbf{E}\,U_\nu^2 = 1$, where U_ν denotes the U-statistic from decomposition (1.18).

The proof is similar to that of Theorem 1. The only difference occurs in the first step of the proof. Here, using (1.18), we replace the probability distribution function $F_S(x)$ by $\tilde{F}_S(x) := \mathbf{P}\{U_\nu \le S_\nu(T_\nu)x\}$ and then replace $S_\nu(T_\nu)$ by $\tilde{S}$. In particular, we show that

$$\sup_{x\in R} |F_S(x) - H(x)| \le \sup_{x\in R} |\tilde{F}_S(x) - H(x)| + o(n_*^{-1/2}). \tag{2.21}$$

This bound in combination with the bound

$$\sup_{x\in R} |\tilde{F}_S(x) - H(x)| = o(n_*^{-1/2}). \tag{2.22}$$

proves the theorem.

In order to prove (2.22) we show that

$$S_\nu^2(T_\nu) = 1 + L_S + M_\nu + M_\nu', \tag{2.23}$$

where the random variables L_S, M_ν are defined by (2.6). The random variable M_ν' is defined in (2.30) below. In order to derive (2.22) from (2.23), we show in Lemma 1 below that $\{M_\nu'\}$ satisfies (2.2). Using this fact and proceeding as in the proof of (2.11), we obtain from (2.23) that

$$S_\nu(T_\nu)/\tilde{S} = 1 + \tilde{M}_\nu, \tag{2.24}$$

where $\{\tilde{M}_\nu\}$ satisfies (2.2). Therefore, the same argument as that used in Step 2 of the proof of Theorem 1 shows that $\sup_x |\tilde{F}_S(x) - H(x)| \le \tilde{\Delta} + o(n_*^{-1/2})$. Finally, invoking the bound (see (2.20)) $\tilde{\Delta} = o(n_*^{-1/2})$, we obtain (2.22).

It remains to prove (2.21) and (2.23). It follows from (1.18), by the orthogonality, that $\sigma_{\nu T}^2 = \mathbf{E}\, U_\nu^2 + \mathbf{E}\, R_\nu^2$. In view of the identity $\mathbf{E}\, U_\nu^2 = 1$, we obtain

$$1 = \mathbf{E}\, U_\nu^2 \le \sigma_{\nu T}^2 = 1 + \mathbf{E}\, R_\nu^2. \tag{2.25}$$

Invoking the bound, see Theorem 1 of [7], $\mathbf{E}\, R_\nu^2 \le n_{\nu *}^{-1}\delta_{\nu 3}$, we obtain from (1.20) that

$$\mathbf{E}\, R_\nu^2 \le \varepsilon_\nu n_{\nu *}^{-5/3} \sigma_{\nu T}^2. \tag{2.26}$$

In particular, for a sufficiently large integer ν_0 (depending only on the sequence $\{\varepsilon_\nu\} \downarrow 0$) we have $\mathbf{E}\, R_\nu^2 < \sigma_{\nu T}^2/2$ for $\nu > \nu_0$. In what follows, we assume that $\nu > \nu_0$. Then (2.25) and (2.26) imply

$$1 \le \sigma_{\nu T}^2 \le 2, \qquad \mathbf{E}\, R_\nu^2 = o(n_{\nu *}^{-5/3}), \qquad \text{and} \qquad \mathbf{E}\, R_\nu^2/\sigma_{\nu T}^2 = o(n_{\nu *}^{-5/3}). \tag{2.27}$$

Write $\Omega_k = \{1, 2, \dots, k\}$ for $k = 1, 2, \dots$. By A_j we shall denote subsets of Ω_N of size $|A_j| = j$. Furthermore, given $A_j = \{i_1, \dots, i_j\} \subset \Omega_N$, denote $T_{A_j} = g_j(X_{i_1}, \dots, X_{i_j})$. Recall that $(X_1, \dots, X_N)$ denotes a random permutation of the ordered population $\mathcal{X}$. Decomposition (1.9) can be written in the form

$$T = \mathbf{E}\, T + \sum_{j=1}^{n_*} \sum_{A_j \subset \Omega_n} T_{A_j}, \tag{2.28}$$

where the second sum is taken over all $j-$subsets A_j of Ω_n.

Let us prove (2.23). Using (2.28) we can decompose

$$V_k = \overline{T} - T_{(k)} = Z_k + W_k + \tilde{R}_k,$$

where Z_k and W_k are defined in Step 1 of the proof of Theorem 1 and where

$$\tilde{R}_k = \sum_{j=3}^{n_*} \varkappa_{k,j}, \qquad \varkappa_{k,j} = \sum_{A_j \subset \Omega_{n+1}} T_{A_j} \big(\mathbb{I}_{\{k \in A_j\}} - \frac{j}{n+1} \big). \tag{2.29}$$

Then

$$V_k^2 = Z_k^2 + W_k^2 + \tilde{R}_k^2 + 2Z_k W_k + 2Z_k \tilde{R}_k + 2W_k \tilde{R}_k.$$

This identity, in combination with (2.7) and (2.6), shows (2.23) with

$$M'_\nu = q\sum_{k=1}^{n+1} \tilde R_k^2 + 2q\sum_{k=1}^{n+1} Z_k\tilde R_k + 2q\sum_{k=1}^{n+1} W_k\tilde R_k =: q(J_{\nu 1} + 2J_{\nu 2} + 2J_{\nu 3}). \tag{2.30}$$

Let us prove (2.21). Write $F_S(x) = \mathbf{P}\{U_\nu S_\nu^{-1}(T_\nu) + Q_\nu \le x\}$, where $Q_\nu = R_\nu S_\nu^{-1}(T_\nu)$. Since $H'(x)$ is bounded, inequality (2.21) would follow if we show that $\mathbf{P}\{|Q_\nu| > 4\tau_\nu^{-8/7}\} = o(\tau_\nu^{-1})$. This bound is a consequence of the bounds

$$\mathbf{P}\{|S_\nu(T_\nu) - 1| > 3/4\} = o(\tau_\nu^{-1}), \qquad \text{and} \qquad \mathbf{P}\{|R_\nu| > \tau_\nu^{-8/7}\} = o(\tau_\nu^{-1}).$$

The first bound follows from (2.10) and (2.24). The second bound follows from (2.27) via Chebyshev's inequality. Theorem 2 is proved.

Lemma 1. *Under conditions of Theorem 2, the sequence $\{M'_\nu\}$ defined in (2.30) satisfies (2.2).*

Proof of Lemma 1. In view of (2.30), it suffices to show (2.2) for the sequences $\{qJ_{\nu i},\ \nu = 1, 2, \dots\}$, $i = 1, 2, 3$.

In order to verify (2.2) for the sequence $\{qJ_{\nu 1}\}$, we show that

$$\mathbf{E}\,|qJ_{\nu 1}| = o(\tau_\nu^{-10/3}). \tag{2.31}$$

By symmetry,

$$J_{\nu 1} = (n+1)\mathbf{E}\,\tilde R_{n+1}^2. \tag{2.32}$$

It follows from (2.29), by the orthogonality, that

$$\mathbf{E}\,\tilde R_{n+1}^2 = \mathbf{E}\,\varkappa_{n+1,3}^2 + \cdots + \mathbf{E}\,\varkappa_{n+1,n_*}^2. \tag{2.33}$$

Proceeding as in the proof of the inequality (4.33) of [4], we obtain

$$\mathbf{E}\,\varkappa_{n+1,j}^2 \le 2b_j d'_j \mathbf{E}\,U_j^2,$$

where $b_j = j/(n+1)$ and $d'_j = (N-j+1)/(N-n)$. Invoking the identity (see (3.11) of [4])

$$\mathbf{E}\,U_j^2 = 2^{-3}h_{3,j}\mathbf{E}\,U_j^2(\mathbb{D}_3 T), \qquad h_{3,j} = [n]_3[N-n]_3/[j]_3[N-j+1]_3,$$

(here $[x]_3 := x(x-1)(x-2)$) and using the simple bound $b_j d'_j h_{3,j} \le n_*^2$, we obtain

$$\mathbf{E}\,\varkappa_{n+1,j}^2 \le \tau^4\mathbf{E}\,U_j^2(\mathbb{D}_3 T). \tag{2.34}$$

Here $U_j(\mathbb{D}_3 T)$ denotes the j-th summand of the orthogonal decomposition (1.9) for $\mathbb{D}_3 T$. In particular, we have (see (3.8) in [4])

$$\mathbf{E}\,(\mathbb{D}_3 T)^2 = \mathbf{E}\,\big(U_3(\mathbb{D}_3 T)\big)^2 + \cdots + \mathbf{E}\,\big(U_{n_*}(\mathbb{D}_3 T)\big)^2.$$

This identity, in combination with (2.33) and (2.34), shows that $\mathbf{E}\,\tilde{R}^2_{n+1} \le \tau^4 \mathbf{E}\,(\mathbb{D}_3 T)^2$. Since $\mathbf{E}\,(\mathbb{D}_3 T)^2 = n_*^{-4}\delta_{\nu 3}$, we have, under condition (1.20) and in view of (2.27),

$$\mathbf{E}\,\tilde{R}^2_{n+1} \le \varepsilon_\nu \tau_\nu^{-16/3}. \tag{2.35}$$

This bound in combination with (2.32) implies (2.31).

In order to verify (2.2) for the sequence $\{qJ_{\nu 2}\}$, we split

$$J_{\nu 2} = J_{2,1} - J_{2,2}, \qquad J_{2,1} = \sum_{k=1}^{n+1} Y_k \tilde{R}_k, \qquad J_{2,2} = \overline{Y}\sum_{k=1}^{n+1} \tilde{R}_k,$$

where we denote $\overline{Y} = (Y_1 + \cdots + Y_{n+1})/(n+1)$. We shall show that

$$\mathbf{E}\,|J_{2,2}| \le 2\varepsilon_\nu^{1/2}\tau_\nu^{-8/3}, \qquad \mathbf{E}\,|qJ_{2,1}|^{3/2} \le c\,\varepsilon_\nu^{3/4}\tau_\nu^{-5/2}. \tag{2.36}$$

These bounds imply (2.2) for $\{qJ_{\nu 2}\}$ via Chebyshev's inequality.

Let us prove the first bound. By symmetry and Cauchy-Schwarz,

$$\mathbf{E}\,|J_{2,2}| = (n+1)\mathbf{E}\,|\overline{Y}\tilde{R}_{n+1}| \le (n+1)(\mathbf{E}\,\overline{Y}^2)^{1/2}(\mathbf{E}\,\tilde{R}^2_{n+1})^{1/2}.$$

The first bound of (2.36) follows from (2.35) and the simple bound

$$\mathbf{E}\,\overline{Y}^2 = \frac{N-n-1}{(n+1)(N-1)}\,\mathbf{E}\,Y_1^2 \le \frac{\tau^2}{n(n+1)}\,\mathbf{E}\,Y_1^2 \le \frac{1}{n(n+1)}$$

(use (2.5) in the last step).

Let us prove the second bound of (2.36). Put $s = 3/2$. Hölder's inequality applied to a sequence of positive numbers $a_1, \dots, a_k$ shows that $(\sum a_i)^s \le k^{s-1}\sum a_i^s$. Using this inequality and the symmetry, we obtain

$$\mathbf{E}\,|J_{2,1}|^s \le (n+1)^{s-1}\sum |Y_k\tilde{R}_k|^s = (n+1)^s\mathbf{E}\,|Y_{n+1}\tilde{R}_{n+1}|^s.$$

Letting $K_1 = \tau_\nu Y_{n+1}$ and $K_2 = \tau_\nu^{16/3}\varepsilon_\nu^{-1}\tilde{R}^2_{n+1}$, we obtain

$$\mathbf{E}\,|J_{2,1}|^s = (n+1)^s\tau_\nu^{-s}\tau_\nu^{-8s/3}\varepsilon_\nu^{s/2}\mathbf{E}\,|K_1 K_2^{1/2}|^s.$$

Finally, (2.36) follows from the bound $\mathbf{E}\,|K_1 K_2^{1/2}|^s \le c$. This bound follows from (1.12) and (2.35) via the inequality $ab \le a^u + b^v$ (here $a, b, u, v > 0$ and $u^{-1} + v^{-1} = 1$). We apply this inequality to $a = |K_1|^s$, $b = |K_2^{s/2}|$, and $v = 4/3$.

In order to verify (2.2) for the sequence $\{qJ_{\nu 3}\}$, we write, by Cauchy-Schwarz,

$$q|J_{\nu 3}| \le (q\tilde{J}_{\nu 3})^{1/2}(qJ_{\nu 1})^{1/2}, \quad \tilde{J}_{\nu 3} = \sum W_k^2,$$

and show that $\{q\tilde{J}_{\nu 3}\}$ satisfies (2.2). In view of (2.31) this implies (2.2) for $\{qJ_{\nu 3}\}$. Using the identity $\tilde{J}_{\nu 3} = m_5 + m_6 + m_7 + m_8$ and the fact that qm_j satisfies (2.2) for $j = 5, 6, 7, 8$, see Step 1 of the proof of Theorem 1, we conclude that $\{q\tilde{J}_{\nu 3}\}$ satisfies (2.2). Lemma 1 is proved

Lemma 2. *Assume that (1.12) holds and $\sigma_U^2 = 1$. Then there exists a constant c depending only on C_1 such that (2.13) holds for every $\nu = 1, 2, \ldots$.*

Proof. Write $\tilde{L} = \sum_{i=1}^{n_*} f(X_i)$. It follows from $\mathbf{E}\, f(X_1) = 0$ that $\tilde{L}_S = -\sum_{k=n+1}^{N} f(X_k)$ and, therefore, the distributions of $|\tilde{L}_S|$ and $|\tilde{L}|$ are the same. Split $\tilde{L} = L_1 + L_2$, where $L_j = \sum_{i=1}^{n_*} f_j(X_i)$, for $j = 1, 2$, and where

$$f_1(X_i) = q\big(1 - \frac{1}{n+1}\big)(Y_i^2 - \sigma_1^2), \qquad f_2(X_i) = 2qn\,\frac{N-1}{N-2}\,\mathbf{E}\,(Y_j Y_{ij}|X_i).$$

We have, by Chebyshev's inequality,

$$\begin{aligned}\mathbf{P}\{|A - B| > n_*^{-8/15}\} &= \mathbf{P}\{|\tilde{L}| > 2n_*^{-4/15}\}\\ &\le \mathbf{P}\{|L_1| > n_*^{-3/10}\} + \mathbf{P}\{|L_2| > n_*^{-4/15}\}\\ &\le n_*^{9/10}\mathbf{E}\,|L_1|^3 + n_*^{48/75}\mathbf{E}\,|L_2|^{12/5}.\end{aligned}$$

We complete the proof of (2.13) by showing the bounds

$$\mathbf{E}\,|L_1|^3 \le cn_*^{-3/2}, \qquad \mathbf{E}\,|L_2|^{12/5} \le cn_*^{-6/5}. \tag{2.37}$$

For this purpose we replace the sums L_1 and L_2 by the corresponding sums of i.i.d. random variables using Theorem 4 of Hoeffding [11] and then apply Rosenthal's inequality. Given $i = 1, 2$, let $z_1, \ldots, z_{n_*}$ be independent copies of $f_i(X_1)$. By Theorem 4 of [11], for every $t \ge 1$ we have $\mathbf{E}\,|L_i|^t \le \mathbf{E}\,|z_1 + \cdots + z_{n_*}|^t$. Rosenthal's inequality shows

$$\mathbf{E}\,|z_1 + \cdots + z_{n^*}|^t \le cn_*\mathbf{E}\,|f_i(X_1)|^t + c\big(n_*\mathbf{E}\,f_i^2(X_1)\big)^{t/2}. \tag{2.38}$$

For $i = 1$ we choose $t = 3$ and obtain the first bound of (2.37) using the first condition of (1.12). For $i = 2$ we choose $t = 12/5$ and obtain the second bound of (2.37). In this case the application of (1.12) is not straightforward. We have

$$\mathbf{E}\,|f_2(X_i)|^t \le c n_*^t \mathbf{E}\,|Y_j Y_{ij}|^t \le c n_*^{-t} \mathbf{E}\, a^t b^t,$$

where $a = |\sigma_1^{-1} Y_1|$ and $b = |\sigma_1^{-1}\tau^2 Y_{12}|$. The inequality $a^t b^t = (b^4)^{t/4}(a^6)^{1-t/4} \le b^4 + a^6$ combined with (1.12) shows $\mathbf{E}\, a^t b^t \le 2C_1$, thus proving $\mathbf{E}\,|f_2(X_i)|^t \le c n_*^{-t}$. The bound $\mathbf{E}\, f_2^2(X_i) \le c n_*^{-2}$ is simpler. Substitution to (2.38) shows the second bound of (2.37), thus proving (2.13).

Lemma 3. *There exists an absolute constant c_0 such that for every n and N we have*

$$\mathbf{E}\,\Big(\sum_{1\le i<j\le n_*} Y_{ij}\Big)^4 \le c_0 n_*^4 \mathbf{E}\, Y_{12}^4, \tag{2.39}$$

$$\mathbf{E}\,\Big(\sum_{1\le j\le n_*} Y_{j\,n+1}\Big)^4 \le c_0 n_*^2 \mathbf{E}\, Y_{12}^4. \tag{2.40}$$

Proof of Lemma 3. For random variables X and Y, write $X \simeq Y$ if $\mathbf{E}\, X = \mathbf{E}\, Y$. Let $H = \sum_{1\le i<j\le n_*} Y_{ij}$ and $y = \mathbf{E}\, Y_{12}^4$. For two real numbers A and B we write $A \prec B$ if $|A| \le c_1|B|$. Here c_1 is a sufficiently large absolute constant.

The bound (2.39) follows from the identities

$$H^4 \simeq \binom{n_*}{2} Y_{12} H^3, \qquad Y_{12}H^3 \simeq H_1 + 2\binom{n_*-2}{1} H_2 + \binom{n_*-2}{2} H_3,$$
$$H_1 = Y_{12}^2 H^2, \quad H_2 = Y_{12}Y_{13}H^2, \quad H_3 = Y_{12}Y_{34}H^2,$$

and the bounds $\mathbf{E}\, H_1 \prec n_*^2 y$, $\mathbf{E}\, H_2 \prec n_* y$, and $\mathbf{E}\, H_3 \prec y$.

Let us show that $\mathbf{E}\, H_1 \prec n_*^2 y$. Write, by symmetry,

$$H_1 \simeq Y_{12}^3 H + 2\binom{n_*-2}{1} Y_{12}^2 Y_{13} H + \binom{n_*-2}{2} Y_{12}^2 Y_{34} H,$$

where

$$Y_{12}^3 H \simeq Y_{12}^4 + 2\binom{n_*-2}{1} Y_{12}^3 Y_{13} + \binom{n_*-2}{2} Y_{12}^3 Y_{34},$$

$$\begin{aligned} Y_{12}^2 Y_{13} H \simeq\ & Y_{12}^3 Y_{13} + Y_{12}^2 Y_{13}^2 + Y_{12}^2 Y_{13} Y_{23} \\ & + \binom{n_*-3}{1}\Big(Y_{12}^2 Y_{13} Y_{14} + Y_{12}^2 Y_{13} Y_{24} + Y_{12}^2 Y_{13} Y_{34}\Big) \\ & + \binom{n_*-3}{2} Y_{12}^2 Y_{13} Y_{45}, \end{aligned}$$

$$\begin{aligned} Y_{12}^2 Y_{34} H \simeq\ & Y_{12}^3 Y_{34} + Y_{12}^2 Y_{34}^2 \\ & + Y_{12}^2 Y_{34} Y_{13} + Y_{12}^2 Y_{34} Y_{14} + Y_{12}^2 Y_{34} Y_{23} + Y_{12}^2 Y_{34} Y_{24} \\ & + 2\binom{n_*-4}{1} Y_{12}^2 Y_{34} Y_{15} + 2\binom{n_*-4}{1} Y_{12}^2 Y_{34} Y_{35} + \binom{n_*-4}{2} Y_{12}^2 Y_{34} Y_{56}. \end{aligned}$$

In a product like $\prod = Y_{12} Y_{23}^2 Y_{24} Y_{56} Y_{46}^3$ we call the indices 1 and 5 free (respective factors Y_{12} and Y_{56} have the power 1 and these indices are not present in other multipliers). Using the identity $\mathbf{E}\,(Y_{12}|X_2) = 0$, we write

$$\begin{aligned} \mathbf{E}\,(Y_{12}|X_2, X_3, \ldots, X_6) &= \frac{1}{N-5} \sum_{j\in\Omega_N\setminus\{2,\ldots,6\}} Y_{j2} \\ &= \frac{1}{N-5} \sum_{j\in\Omega_N\setminus\{2\}} Y_{j2} - \frac{1}{N-5} \sum_{j=3}^{6} Y_{j2} \\ &= -\frac{1}{N-5} \sum_{j=3}^{6} Y_{j2}. \end{aligned}$$

Therefore, the expectation $\mathbf{E}\,p$ splits into a sum of expectations of similar products, but with the additional factor $-(N-5)^{-1} = O(N^{-1})$. In this way the "free" index 1 produces the factor $O(N^{-1})$. We call this procedure "free index argument". Using this argument, we obtain

$$\mathbf{E}\,Y_{12}^3 H \prec y, \quad \mathbf{E}\,Y_{12}^2 Y_{13} H \prec y, \quad \mathbf{E}\,Y_{12}^2 Y_{34} H \prec y.$$

Therefore, $\mathbf{E}\,H_1 \prec n_*^2 y$.

The proof of $H_2 \prec n_* y$ and $\mathbf{E}\,H_3 \prec y$ is similar.

Let us prove (2.40). Write $Y_{i\,n+1} = Z_i$, for $1 \le i \le n_*$, and $S = Z_1 + \cdots + Z_{n_*}$. By symmetry,

$$S^4 \simeq n_* Z_1 S^3, \qquad Z_1 S^3 = Z_1^2 S^2 + (n_*-1) Z_1 Z_2 S^2.$$

In order to prove (2.40), we show that

$$\mathbf{E}\, Z_1^2 S^2 \prec n_* y, \qquad \mathbf{E}\, Z_1 Z_2 S^2 \prec y. \tag{2.41}$$

Write, by symmetry,

$$Z_1^2 S^2 \simeq Z_1^3 S + (n_* - 1) Z_1^2 Z_2 S$$

and

$$Z_1^3 S \simeq Z_1^4 + (n_* - 1) Z_1^3 Z_2, \qquad Z_1^2 Z_2 S = Z_1^2 Z_2 (Z_1 + Z_2) + (n_* - 2) Z_1^2 Z_2 Z_3.$$

The "free" index argument shows that $\mathbf{E}\, Z_1^3 S \prec y$ and $\mathbf{E}\, Z_1^2 Z_2 S \prec y$. This implies the first bound of (2.41).

To prove the second bound of (2.41), write, by symmetry,

$$Z_1 Z_2 S^2 \simeq Z_1^2 Z_2 S + Z_1 Z_2^2 S + (n_* - 2) Z_1 Z_2 Z_3 S.$$

We have already shown that the expectations of the first two terms $\prec y$. In order to show that $\mathbf{E}\, Z_1 Z_2 Z_3 S \prec n_*^{-1} y$, we write, by symmetry,

$$Z_1 Z_2 Z_3 S \simeq Z_1 Z_2 Z_3 (Z_1 + Z_2 + Z_3) + (n_* - 3) Z_1 Z_2 Z_3 Z_4$$

and invoke "free" index argument. We obtain $\mathbf{E}\, Z_1 Z_2 S^2 \prec y$, thus showing the second bound of (2.41). The lemma is proved.

Lemma 4. *Assume that $\{U_\nu\}$ is a sequence of U statistics of degree two. Assume that for some $s > 3$ and $c_1 > 0$ we have for $\nu = 1, 2, \ldots$*

$$\mathbf{Var}\, U_\nu = 1 \qquad \beta_{\nu\, s} < c_1, \qquad \gamma_{\nu\, 4} < c_1.$$

Then there exists a sequence $\{\psi_\nu\} \downarrow 0$ such that $\mathbf{P}\{|\hat\alpha_\nu - \alpha_\nu| > \psi_\nu\} < \psi_\nu$ as $n_{\nu *} \to \infty$.

Note that the condition $\beta_{\nu\, s} < c_1$ for some $s > 3$, can be reduced to the uniform integrability condition for $\beta_{\nu\, 3}$, see condition (2.13) of [4].

Proof of Lemma 4. The proof if almost the same as that of Lemma 2.2 of [4]. The only difference appears in the proof of the relation $R_3 = o_P(1)$ (in probability), where

$$R_3 = n^{-1} \sum_{k=1}^{n+1} r_k^3 \sigma_1^{-3}, \qquad r_k = \sum_{1 \le i < j \le n+1} \Big(\mathbb{I}_{k \in \{i,j\}} - \frac{2}{n+1} \Big) Y_{ij}.$$

Therefore, we shall only show that conditions of Lemma 4 imply $\mathbf{E}\, |R_3| = o(1)$.

By symmetry,

$$\mathbf{E}\,|R_3| \le (n+1)n^{-1}\sigma_1^{-3}\mathbf{E}\,|r_{n+1}|^3.$$

In view of the relation (which follows from $\mathbf{Var}\,U_\nu = 1$ and $\gamma_{\nu 4} < c_1$) $\sigma_1^2 = \tau^{-2} + O(\tau^{-4})$, see (2.5), it suffices to show that $\mathbf{E}\,|r_{n+1}|^3 = o(\tau^{-3})$. For this purpose, we show $\mathbf{E}\,r_{n+1}^4 = o(\tau^{-4})$.

Write

$$r_{n+1} = r_A - 2(n+1)^{-1} r_B, \tag{2.42}$$
$$r_A = \sum_{j=1}^{n} Y_{j\,n+1}, \qquad r_B = \sum_{1\le i<j\le n+1} Y_{ij}.$$

We shall show, for an absolute constant $c_0 > 0$,

$$\mathbf{E}\,r_A^4 \le c_0 n_*^{-4}\gamma_4, \qquad \mathbf{E}\,r_B^4 \le c_0 n_*^{-2}\gamma_4. \tag{2.43}$$

If $n = n_*$, these bounds follow from Lemma 3. For $n > n_*$ (i.e., $n_* = N - n$) we use $\mathbf{E}\,(Y_{ij}|X_j) = 0$ and $\mathbf{E}\,Y_{ij} = 0$ to show that

$$r_A = -\sum_{j=n+2}^{N} Y_{j\,n+1} \qquad \text{and} \qquad r_B = \sum_{n+2\le i<j\le N} Y_{ij}.$$

Now (2.43) follows from Lemma 3 again. Finally, combining (2.42) and (2.43), we obtain $\mathbf{E}\,r_{n+1}^4 \le c n_*^{-4}\gamma_4$, thus completing the proof.

Lemma 5. *For random variables defined in Step 1 of the proof of Theorem 1, we have*

$$\mathbf{E}\,H_1^2 = \frac{n+1}{n}\,\frac{N-n-1}{N-n}\,\mathbf{Var}\,L \le 2\mathbf{Var}\,L, \tag{2.44}$$
$$\mathbf{E}\,H_2^2 = \frac{n+1}{n-1}\,\frac{N-n-2}{N-n}\,\mathbf{Var}\,Q \le 3\mathbf{Var}\,Q. \tag{2.45}$$

Assume that $\sigma_U^2 = 1$. *Assume that (1.12) holds. Then there exists a constant* $c > 0$ *depending only on* C_1 *such that*

$$\mathbf{E}\,H_1^2 \le 2, \qquad \mathbf{E}\,H_2^2 \le c\tau^{-2}, \tag{2.46}$$
$$\mathbf{E}\,m_5 \le cq^{-2}\tau^{-2} \qquad \mathbf{Var}\,m_5 \le cq^{-2}\tau^{-6}, \tag{2.47}$$
$$|\mathbf{E}\,m_2| \le \tau^{-2}, \qquad \mathbf{Var}\,m_2 \le cn^{-2}, \tag{2.48}$$
$$\mathbf{E}\,(H - L_H)^2 \le c\tau^{-4}, \tag{2.49}$$
$$|\mathbf{E}\,H_3| \le cq^{-2}\tau^{-2}, \qquad \mathbf{Var}\,H_3 \le cq^{-2}\tau^{-4}. \tag{2.50}$$

Proof of Lemma 5. Note that (2.44), (2.45), and (2.46) are immediate consequences of (2.1), (2.3), and (2.4). The first bound of (2.47) follows from (2.4).

Before proving the remaining inequalities, we introduce some notation. Given a centered U-statistics W based on the sample $X_1, \dots, X_{n+1}$ drawn without replacement form the population $\mathcal{X}$, we decompose by (1.9) $W = L_W + Q_W$, where

$$L_W = \sum_{1 \le i \le n+1} W_i, \qquad Q_W = \sum_{1 \le i < j \le n+1} W_{ij} \tag{2.51}$$

denote the linear and quadratic parts. It follows from (2.1) and (2.3) that

$$\mathbf{E}\, W^2 = \mathbf{E}\, L_W^2 + \mathbf{E}\, Q_W^2 \le c_0 \tau^2 \mathbf{E}\, W_1^2 + c_0 \tau^4 \mathbf{E}\, W_{12}^2. \tag{2.52}$$

In order to prove the second bound of (2.47), write

$$m_5 = n(n+1)\sigma_2^2 + 2W, \qquad W = \sum_{1 \le i < j \le n+1} (Y_{ij}^2 - \sigma_2^2).$$

Clearly, $\mathbf{Var}\, m_5 = 4\mathbf{E}\, W^2$. A simple calculation shows that the summands of the linear and quadratic parts, see (1.11) and (2.51), are

$$W_i = n\, \frac{N-1}{N-2}\, \mathbf{E}\, (Y_{ij}^2 - \sigma_2^2 | X_i), \qquad W_{ij} = Y_{ij}^2 - \sigma_2^2 - (W_i + W_j)/n.$$

It follows from (1.12) that $\mathbf{E}\, W_i^2 \le c n^2 \tau^{-12}$ and $\mathbf{E}\, W_{ij}^2 \le c \tau^{-12}$. These bounds in combination with (2.52) yield the second inequality of (2.47).

Let us prove (2.48). Write

$$m_2 = -\frac{2}{n+1} \binom{n+1}{2} \mathbf{E}\, Y_1 Y_2 - \frac{2}{n+1}\, W, \qquad W = \sum_{1 \le i < j \le n+1} (Y_i Y_j - \mathbf{E}\, Y_i Y_j).$$

Here $\mathbf{E}\, W = 0$. Using $\mathbf{E}\, Y_j = 0$ we obtain for $i \ne j$

$$\mathbf{E}\, (Y_i Y_j | X_i) = -Y_i^2/(N-1) \qquad \text{and} \qquad \mathbf{E}\, Y_i Y_j = -\sigma_1^2/(N-1).$$

Furthermore, invoking (2.5) we get

$$\mathbf{E}\, m_2 = \sigma_1^2 n/(N-1) \qquad \text{and} \qquad |\mathbf{E}\, m_2| \le \tau^{-2},$$

thus proving the first bound of (2.48). To prove the second bound, we apply (2.51) and (2.52). A simple calculations shows that the summands of the linear and quadratic parts are

$$W_i = n\, \frac{N-1}{N-2}\, \mathbf{E}\, \bigl(Y_i Y_j + \frac{\sigma_1^2}{N-1} \,\big|\, X_i\bigr) = \frac{n}{N-2}\, (\sigma_1^2 - Y_i^2),$$
$$W_{ij} = Y_i Y_j + (N-1)^{-1} \sigma_1^2 - n^{-1}(W_i + W_j).$$

Using (1.12) we show that $\mathbf{E}\, W_1^2 \le c\tau^{-4}$ and $\mathbf{E}\, W_{12}^2 \le c\tau^{-4}$. Invoking (2.52) we obtain $\mathbf{E}\, W^2 \le c$. This implies the second bound of (2.48).

Let us prove (2.49). Note that $H - L_H$ is a centered U-statistic. Calculation shows that the linear part of the statistic $H - L_H$ vanishes. Its orthogonal decomposition (1.9) contains only the quadratic part and, therefore, reduces to

$$H - L_H = Q_H, \qquad Q_H = \sum_{1\le i<j\le n+1} H_{ij},$$
$$H_{ij} = Y_{ij}(Y_i + Y_j) - \frac{N-1}{N-2}\Big(\mathbf{E}\,(Y_j Y_{ij}|X_i) + \mathbf{E}\,(Y_i Y_{ij}|X_j)\Big).$$

It follows from (1.12) that $\mathbf{E}\, H_{ij}^2 \le c\tau^{-8}$. The second identity of (2.3) shows that

$$\mathbf{E}\, Q_H^2 = \frac{\binom{n+1}{2}\binom{N-n-1}{2}}{\binom{N-2}{2}}\,\mathbf{E}\, H_{ij}^2 \le c_0\tau^4 \mathbf{E}\, H_{ij}^2 \le c\tau^{-4},$$

thus proving (2.49).

In order to prove (2.50), we write H_3 in the form of U-statistic of degree three

$$H_3 = \sum_{1\le i<j<k\le n+1} \Psi_{ijk}, \qquad \Psi_{ijk} = Y_{ij}Y_{ik} + Y_{ki}Y_{kj} + Y_{ji}Y_{jk}.$$

Calculation shows that $\mathbf{E}\, Y_{12}Y_{13} = -\sigma_2^2/(N-2)$. We obtain, by symmetry, that $\mathbf{E}\,\Psi_{ijk} = 3\mathbf{E}\, Y_{12}Y_{13} = -3\sigma_2^2/(N-2)$. Therefore,

$$\mathbf{E}\, H_3 = \binom{n+1}{3}\mathbf{E}\,\Psi_{123} = -3\,\frac{\binom{n+1}{3}}{N-2}\,\sigma_2^2.$$

This implies the first bound of (2.50). In order to prove the second bound, decompose, by means of (1.9), $H_3 = \mathbf{E}\, H_3 + U_1 + U_2 + U_3$, where

$$U_1 = \sum_{1\le i\le n+1}\binom{n}{2}\tilde\Psi_i, \quad U_2 = \sum_{1\le i<j\le n+1}\binom{n-1}{1}\tilde\Psi_{ij}, \quad U_3 = \sum_{1\le i<j<k\le n+1}\tilde\Psi_{ijk}.$$

Here we write for $r \ne i \ne j$

$$\tilde\Psi_i = \frac{3(N-1)}{(N-2)(N-3)}\Big(\sigma_2^2 - \mathbf{E}\,(Y_{ri}^2|X_i)\Big),$$
$$\tilde\Psi_{ij} = -\frac{2}{N-4}\,Y_{ij}^2 + \frac{N-2}{N-4}\,\mathbf{E}\,(Y_{ri}Y_{rj}|X_i,X_j) + \frac{3}{N-4}\,\sigma_2^2$$
$$+\frac{N-2}{N-4}\Big(3\,\frac{N-1}{(N-2)^2}\,(\mathbf{E}\,(Y_{ri}^2|X_i) + \mathbf{E}\,(Y_{rj}^2|X_j)) - 6\,\frac{N-1}{(N-2)^2}\,\sigma_2^2\Big),$$
$$\tilde\Psi_{ijk} = \Psi_{ijk} - \mathbf{E}\,\Psi_{ijk} - \tilde\Psi_i - \tilde\Psi_j - \tilde\Psi_k - \tilde\Psi_{ij} - \tilde\Psi_{ik} - \tilde\Psi_{jk}.$$

Using (1.12) one can show that

$$\mathbf{E}\,\tilde{\Psi}_i^2 \le c\tau^{-12}N^{-2}, \qquad \mathbf{E}\,\tilde{\Psi}_{12}^2 \le c\tau^{-12}, \qquad \mathbf{E}\,\tilde{\Psi}_{123}^2 \le c\tau^{-12}.$$

Finally, invoking (2.1) we obtain

$$\begin{aligned}\mathbf{E}\,\mathbf{Var}\,H_3 &= \mathbf{E}\,U_1^2 + \mathbf{E}\,U_2^2 + \mathbf{E}\,U_3^2 \\ &\le c\tau^2\binom{n}{2}^2\mathbf{E}\,\tilde{\Psi}_1^2 + c\tau^4\binom{n-1}{1}^2\mathbf{E}\,\tilde{\Psi}_{12}^2 + c\tau^6\mathbf{E}\,\tilde{\Psi}_{123}^2 \\ &\le cn^2\tau^{-8}.\end{aligned}$$

This proves (2.50).

References

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

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

[3]. Beran, R., *Jackknife approximations to bootstrap estimates*, Ann. Statist. **12** (1984), 101–118.

[4]. Bloznelis, M., *Empirical Edgeworth expansion for finite population statistics.I, II*, Liet. Matem. Rink. **41** (2001), 154–171, 263–276.

[5]. Bloznelis, M., *An Edgeworth expansion for Studentized finite population statistics*, *Submitted to* Acta Applicandae Mathematicae (2002).

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

[7]. Bloznelis, M. and Götze, F., *Orthogonal decomposition of finite population statistics and its applications to distributional asymptotics*, Ann. Statist. **29** (2001), 899–917.

[8]. Bloznelis, M. and Götze, F., *An Edgeworth expansion for finite population symmetric statistics*, Ann. Probab. **30** (2002).

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

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

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

[12]. Kokic, P.N. and Weber, N.C., *An Edgeworth expansion for U-statistics based on samples from finite populations*, Ann. Probab. **18** (1990), 390–404.

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

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

[15]. Sugden, R. A., Smith, T. M. F. and Jones, R. P., *Cochran's rule for simple random sampling*, J.R. Statist. Soc. B **62** (2000), 787–793.

[16]. Zhao Lin Cheng and Chen Xi Ru, *Normal approximation for finite-population U-statistics*, Acta mathematicae applicatae Sinica **6** (1990), 263–272.