The Annals of Probability
1990, Vol. 18, No. 1, 390–404

AN EDGEWORTH EXPANSION FOR *U*-STATISTICS BASED ON SAMPLES FROM FINITE POPULATIONS

BY P. N. KOKIC AND N. C. WEBER

University of Sydney

Suppose that U is a U-statistic of degree 2 based on a simple random sample of size n selected without replacement from a finite population of N elements. A bound for the difference between the distribution function of a standardized version of U and its single-term Edgeworth expansion is given. We apply these results to obtain an Edgeworth expansion for the variance estimator in a finite population. Some simulation results are reported in this case.

1. Introduction. Nandi and Sen (1963) studied the behaviour of U-statistics based on samples drawn from finite populations as part of a study of U-statistics based on dependent samples. In particular, under suitable regularity conditions, Nandi and Sen established a central limit theorem for a sequence of standardized U-statistics. In this paper we will consider a class of symmetric statistics that includes the U-statistics of degree 2 studied by Nandi and Sen and establish the validity of a single-term Edgeworth expansion of the normalized statistics.

Let $W = \{w_{ij}\}$ be an $N \times N$ array of real numbers satisfying $w_{ij} = w_{ji}$ and $w_{ii} = 0$. Let $(R_1, \ldots, R_N)$ be a vector selected at random from the set of $N!$ permutations of $(1, 2, \ldots, N)$. If $1 < n < N$, we will be interested in the statistics

$$U = \sum_{i=2}^{n} \sum_{j=1}^{i-1} W_{ij},$$

where $W_{ij} = w_{R_i R_j}$.

Note that if we have a finite population, denoted by the real numbers $a_1, a_2, \ldots, a_N$, a symmetric function h: $\mathbb{R}^2 \to \mathbb{R}$, and we set $w_{ij} = h(a_i, a_j)$, $i \neq j$, $w_{ii} = 0$, then the statistic $\binom{n}{2}^{-1} U$ is the U-statistic with kernel h based on the sample consisting of the n terms $a_{R_1}, a_{R_2}, \ldots, a_{R_n}$ drawn without replacement from the finite population.

Let $\bar{A} = n^{-1}\sum_{i=1}^{n} a_{R_i}$, $\bar{a} = N^{-1}\sum_{i=1}^{N} a_i$ and suppose that $h(x, y) = \frac{1}{2}(x - y)^2$. Then $\binom{n}{2}^{-1} U = S^2$, where

$$S^2 = (n-1)^{-1} \sum_{i=1}^{n} \left(a_{R_i} - \bar{A}\right)^2$$

is the finite population estimator of variance as defined in Cochran (1963, page 25). In sample survey theory S^2 is extensively used as an unbiased estimator of

Received February 1988; revised December 1988.
AMS 1980 *subject classification.* 60F05.
Key words and phrases. U-statistics, Edgeworth expansion, Berry–Esseen bound, sampling from a finite population.

$\sigma^2 = (N-1)^{-1}\Sigma_{i=1}^{N}(a_i - \bar{a})^2$. It would therefore be useful to know its distributional properties. Except in the simplest of cases, the exact distribution of S^2 can only be approximated. In Section 2.2 we shall derive the single-term Edgeworth expansion for S^2 and perform a numerical study to illustrate its accuracy in comparison to the simple normal approximation.

Statistical inference based on data resampling has drawn considerable attention in recent years. Wu (1986), for example, has considered variance estimation in regression analysis using the jackknife, bootstrap and other resampling techniques. The similarity of resampling and simple random sampling from a finite population is apparent. It would therefore be of interest in the context of resampling to investigate the asymptotic properties of U.

Following the work of Callaert, Janssen and Veraverbeke (1980), Bickel, Götze and van Zwet (1986) obtained an Edgeworth expansion for U-statistics based on samples of independent and identically distributed (iid) random variables. The approach adopted here is to develop a decomposition of U similar to that used for classical U-statistics, then to use techniques similar to Bickel et al. to establish a single-term Edgeworth expansion for U.

Robinson (1978) developed an Edgeworth expansion for the mean of a sample drawn without replacement from a finite population. In this paper we shall develop an Edgeworth expansion to approximate the distribution of U under a condition similar to Robinson's condition (c). Bickel and van Zwet (1978) and Babu and Singh (1985) also provide further results on Edgeworth expansions for statistics based on samples from a finite population.

The techniques used to establish the Edgeworth expansions can also be used to obtain a Berry–Esseen-type bound for U. The result, stated in Theorem 2, is more general than that given in a recent paper by Zhao and Chen (1987) and it provides an asymptotic normality result that holds under weaker assumptions than those obtained by Nandi and Sen (1963) and Zhao and Chen (1987).

In Section 2.1 we state our main theorems and relate them to results known for U-statistics based on samples of iid random variables. In Section 3 we provide details of the decomposition of U and outline the proof of the Edgeworth expansion. In Section 4 proofs are given of the technical lemmas used in Section 3. Throughout the rest of this paper we shall assume, without loss of generality, that the array (w_{ij}) satisfies

$$\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} = 0 \quad \text{and} \quad \sum_{i=2}^{N}\sum_{j=1}^{i-1} w_{ij}^2 = 1. \tag{1.1}$$

2. Main results. This section has been divided into two subsections. In the first we state our main theorems for U-statistics in general and in the second subsection we report the details of a simulation study to investigate the accuracy of the Edgeworth expansion for the variance estimator in a finite population.

2.1. *U-statistics.* Let $p = n/N$, $q = 1 - p$ and $w_i = \Sigma_{j=1}^{N} w_{ij}$. In the classical case of U-statistics based on iid random variables the normalizing constant is often the standard deviation of the projection term. By analogy, in the finite

population case we can use

$$\operatorname{Var}\left\{\sum_{i\neq j}^{n}\sum^{n} E(W_{ij}|R_i)\right\} = \left(\frac{N}{N-1}\right)^3\left(\frac{n-1}{n}\right)^2 p^3q\sum_{i=1}^{N} w_i^2.$$

However, it is more convenient to standardize our statistic using

$$\nu^2 = p^3q\sum_{i=1}^{N} w_i^2.$$

To avoid trivial cases assume that $\nu^2 > 0$.

We shall approximate the distribution of $\nu^{-1}U$ by the asymptotic expansion

$$\begin{aligned} F(x) &= \Phi(x) - H_2(x)\varphi(x) \\ &\quad\times\left\{\frac{(q-p)}{6(pq)^{1/2}}\left(\sum_{k=1}^{N} w_k^3\right)\left(\sum_{k=1}^{N} w_k^2\right)^{-3/2}\right. \\ &\quad\left.+\left(\frac{q}{p}\right)^{1/2}\left(\sum_{i=2}^{N}\sum_{j=1}^{i-1} w_{ij}w_iw_j\right)\left(\sum_{k=1}^{N} w_k^2\right)^{-3/2}\right\}, \end{aligned} \tag{2.1}$$

where $\Phi(x) = \int_{-\infty}^{x}\varphi(t)\,dt$, $\varphi(x) = (2\pi)^{-1/2}e^{-x^2/2}$ and $H_2(x) = (x^2 - 1)$. We shall show that the expansion is a valid approximation in Theorem 1 below, subject to the condition:

(C) Given $\varepsilon > 0$ and a positive sequence $\{\varepsilon_N\}$ converging to zero as $N \to \infty$, there exists $\varepsilon' > 0$ and $\delta > 0$ not depending on N such that, for all real x and all $s \in (\varepsilon b_N^{-1}, T)$, the number of indices j, for which

$$\left| w_j s\left(\sum_{k=1}^{N} w_k^2\right)^{-1/2} - x - 2r\pi \right| > \varepsilon'$$

for all $r = 0, \pm 1, \pm 2, \ldots$, is greater than δN, for all sufficiently large N, where

$$b_N = \max_{1\le i\le N}|w_i|\left(\sum_{i=1}^{N} w_i^2\right)^{-1/2}$$

and

$$T = \varepsilon_N^{-1}\left(\sum_{j=1}^{N} w_j^2\right)^{3/2}\left(\sum_{j=1}^{N}|w_j^3|\right)^{-1}.$$

Condition (C), first introduced by Albers, Bickel and van Zwet (1976), is similar to the version in Robinson (1978), and it ensures that the values of w_j do not cluster around too few values. Note, however, that the order of statements in condition (C) differs from Robinson's condition and hence is satisfied by a broader class of $\{w_j, 1 \le j \le N\}$. An alternative form of condition (C) is obtained

in Kokic (1988). The approximation at (2.1) provides more accuracy than a simple normal approximation to the distribution of $v^{-1}U$.

THEOREM 1. *Suppose*

$$(2.2) \qquad (pq)^{-1/2} \sum_{j=1}^{N} |w_j^3| \left(\sum_{j=1}^{N} w_j^2 \right)^{-3/2} \to 0 \quad \text{and} \quad v^2 \to \infty$$

as $N \to \infty$. *Furthermore, assume condition* (C) *holds for some sequence* $\{\varepsilon_N > 0\}$ *which converges to zero as* $N \to \infty$. *Then, for some constant* $C(p) > 0$, *depending only on* p, *and some sequence* $\epsilon'_N \to 0$ *as* $N \to \infty$,

$$\Delta = \left| P(v^{-1}U \le x) - F(x) \right|$$

$$\le C(p)\epsilon'_N \left\{ \left(\sum_{j=1}^{N} |w_j^3| \right) \left(\sum_{j=1}^{N} w_j^2 \right)^{-3/2} + \left(\sum_{j=1}^{N} w_j^2 \right)^{-1/2} \right\}.$$

Notice that $C(p)$ depends on n and N through p only. The bound in Theorem 1 is typically of small order $N^{-1/2}$ as is the case for the corresponding result for U-statistics based on independent and identically distributed random variables. To illustrate this point more clearly we give the following corollary, which includes the case of U-statistics with a bounded kernel.

COROLLARY 1. *Suppose condition* (C) *holds and there exists a constant* K *such that for all* N

$$\sum_{j=1}^{N} w_j^2 \ge K^{-1}N \quad \text{and} \quad \sum_{j=1}^{N} |w_j^3| \le KN.$$

Then there exists a constant $C(p)$, *depending only on* p, *and a sequence* $\{a_N\}$ *with* $a_N = o(N^{-1/2})$ *as* $N \to \infty$, *such that*

$$\sup_x \left| P(v^{-1}U \le x) - F(x) \right| \le C(p)a_N.$$

If we only wish to approximate the distribution of $v^{-1}U$ by a normal distribution function, then condition (C) may be removed. Zhao and Chen (1987) obtained a Berry–Esseen bound for U-statistics based on a sample from a finite population under the condition that p is strictly bounded away from 0 and 1. Our theorem holds universally for all p and gives a bound which is of the same or smaller order than that provided by Zhao and Chen (1987).

THEOREM 2. *There exists an absolute constant* $C > 0$ *such that*

$$\sup_x \left| P(v^{-1}U \le x) - \Phi(x) \right|$$

$$\le C \left\{ (pq)^{-1/2} \left(\sum_{j=1}^{N} |w_j^3| \right) \left(\sum_{j=1}^{N} w_j^2 \right)^{-3/2} + (pq)^{1/2} v^{-1} \right\}.$$

The following simple example indicates that the bound in Theorem 2 can be of smaller order than that of Zhao and Chen. As in Zhao and Chen, assume there exist fixed constants λ_1 and λ_2 such that $0 < \lambda_1 \le p \le \lambda_2 < 1$ and consider the $N \times N$ arrays defined by

$$w_{ij} = 2/(l(l-2))^{1/2} \quad \text{if } i \neq j,\ i \text{ and } j \text{ are odd and } i, j < l$$
$$= -2/(l(l-2))^{1/2} \quad \text{if } i \neq j,\ i \text{ and } j \text{ are even and } i, j \le l$$
$$= 0 \quad \text{otherwise,}$$

where l is the smallest even integer greater than or equal to $N^{1/2}$. It is easy to check that $\{w_{ij}\}$ satisfies (1.1) and the bound in Theorem 2 is $O(N^{-1/4})$, whereas the corresponding bound of Zhao and Chen is $O(1)$.

The proof of Theorem 2 is excluded because the techniques used are very similar to those used in Sections 3 and 4 to prove Theorem 1.

The asymptotic normality result for $v^{-1}U$ in Nandi and Sen (1963) was improved in Theorem 2 of Zhao and Chen (1987). The following result, which is obtained from Theorem 2, provides even weaker conditions than those given in Zhao and Chen.

COROLLARY 2. *Suppose as* $N \to \infty$, $(pq)^{-1/2}\sum_{j=1}^{N}|w_j^3|(\sum_{j=1}^{N}w_j^2)^{-3/2} \to 0$ *and* $(pq)^{1/2}v^{-1} \to 0$. *Then* $v^{-1}U$ *converges weakly to a standard normal distribution as* $N \to \infty$.

In fact, we do not require the restriction imposed by Zhao and Chen that the sampling fraction $p = n/N$ is strictly bounded away from 0 and 1 for all N. We see from Corollary 2 that asymptotic normality may even occur when $p \to 0$ or $p \to 1$ as $N \to \infty$. However, for the conditions of the corollary to hold we must have $Npq \to \infty$ as $N \to \infty$.

2.2. *The variance estimator.* As mentioned in the Introduction, the finite population estimator of variance S^2 is the U-statistic with kernel $h(x, y) = \frac{1}{2}(x-y)^2$. If we normalize this statistic according to (1.1), then by (2.1) a single-term Edgeworth approximation to the tail probability

$$P\Big((n-1)(Npq)^{-1/2}\big(\mu_4 - \mu_2^2\big)^{-1/2}(S^2 - \sigma^2) > x\Big) \tag{2.3}$$

is

$$1 - \Phi(x) + H_2(x)\varphi(x)N^{-1/2}\big(\mu_4 - \mu_2^2\big)^{-3/2} \times \left\{\frac{(q-p)}{6(pq)^{1/2}}\big(\mu_6 - 3\mu_2\mu_4 + 2\mu_2^3\big) - \left(\frac{q}{p}\right)^{1/2}\mu_3^2\right\} + H_2(x)\varphi(x)\left(\frac{q}{p}\right)^{1/2}N^{-3/2}\sigma^2\big(\mu_4 - \mu_2^2\big)^{-1/2},$$

where $\mu_k = N^{-1}\sum_{i=1}^{N}(a_i - \bar{a})^k$, $k \ge 1$. For details refer to Kokic (1988). Under condition (2.2) we may show that the final term of this expansion is $o(N^{-1/2})$, and so for the purposes of the present investigation it is ignored. We examined the performance of this Edgeworth expansion for three different finite populations, each of size $N = 100$.

TABLE 2.1

Estimated tail probabilities (given in parentheses) from the normalized distribution of S^2 and the single-term Edgeworth approximation to these probabilities. The normal approximation is given at the base of the table.

		x						
Population	p	**0**	**0.5**	**1**	**1.282**	**1.645**	**1.96**	**2.326**
Normal	0.1	0.461	0.283	0.159	0.111	0.067	0.041	0.021
		(0.448)	(0.278)	(0.151)	(0.104)	(0.061)	(0.036)	(0.021)
	0.2	0.479	0.294	0.159	0.106	0.059	0.034	0.016
		(0.483)	(0.296)	(0.159)	(0.108)	(0.062)	(0.031)	(0.017)
	0.4	0.495	0.305	0.159	0.101	0.052	0.027	0.011
		(0.501)	(0.309)	(0.165)	(0.111)	(0.059)	(0.030)	(0.012)
χ_8^2	0.1	0.458	0.28	0.159	0.112	0.069	0.043	0.023
		(0.442)	(0.266)	(0.162)	(0.107)	(0.064)	(0.040)	(0.021)
	0.2	0.478	0.294	0.159	0.106	0.06	0.034	0.016
		(0.469)	(0.299)	(0.168)	(0.117)	(0.063)	(0.030)	(0.015)
	0.4	0.498	0.307	0.159	0.1	0.051	0.026	0.01
		(0.491)	(0.317)	(0.175)	(0.106)	(0.054)	(0.024)	(0.01)
Lattice	0.1	0.489	0.301	0.159	0.103	0.055	0.03	0.013
		(0.482)	(0.299)	(0.172)	(0.112)	(0.063)	(0.036)	(0.017)
	0.2	0.494	0.304	0.159	0.102	0.053	0.028	0.012
		(0.489)	(0.309)	(0.169)	(0.109)	(0.058)	(0.03)	(0.01)
	0.4	0.498	0.307	0.159	0.1	0.051	0.026	0.011
		(0.505)	(0.306)	(0.163)	(0.106)	(0.056)	(0.029)	(0.012)
Normal approximation		0.5	0.309	0.159	0.1	0.05	0.025	0.01

The first population consists of 100 independent observations from a standard normal random variable, the second is 100 observations from a χ_8^2 random variable and the third consists of the integers 1 up to 100. The third population, which we shall refer to as the lattice case, although not of practical importance in its own right, is of interest because often in practice the observations a_i, $1 \leq i \leq N$, represent counts and hence are distributed on a lattice.

The probability (2.3) was estimated by producing 10,000 independent simple random samples and determining the proportion of times the left-hand quantity inside the probability exceeded x. Table 1 lists the tail probabilities, the Edgeworth approximations and normal approximations $1 - \Phi(x)$ for the standard normal, χ_8^2 and lattice populations. The values $p = 0.1, 0.2, 0.4$ and $x = 0, 0.5, 1, 1.645, 1.96, 2.326$ were chosen to cover most cases of interest.

The Edgeworth expansion provided a very good approximation to the true tail probability in all cases. In fact, the maximum absolute error of this approximation of 0.016 occurred for the χ_8^2 population when $x = 1$ and $p = 0.4$. Note that $H_2(1) = 0$, so that the normal and Edgeworth approximations correspond when $x = 1$. The normal approximation is much worse, in general, with maximum absolute error 0.058 in this study. Examining the results for each population separately, we see that the Edgeworth expansion provides greater accuracy than

the normal approximation in most cases. In fact, it is worse than the normal approximation in only 8 out of the 63 cases and then by no more than 0.005.

3. The decomposition of U and proof of Theorem 1. Given N and n, let $Y_1, \ldots, Y_N$ be iid Bernoulli random variables with expectation p, independent of $(R_1, \ldots, R_N)$. Conditional on $B_N = \sum_{i=1}^{N}(Y_i - p) = 0$, $\sum_{i=2}^{N}\sum_{j=1}^{i-1} Y_i Y_j W_{ij}$ has the same distribution as U. Let

$$\varphi_{ij} = (Y_i - p)(Y_j - p), \qquad W_i = w_{R_i} \quad \text{and} \quad V_{ij} = v_{R_i R_j},$$

where

$$v_{ij} = w_{ij} - N^{-1}w_i - N^{-1}w_j.$$

Then, conditional on $B_N = 0$,

$$\hat{U} + C_N =_{\mathscr{D}} U, \tag{3.1}$$

where

$$\hat{U} = p\sum_{i=1}^{N}(Y_i - p)W_i \quad \text{and} \quad C_N = \sum_{i=2}^{N}\sum_{j=1}^{i-1}\varphi_{ij}V_{ij},$$

since, conditional on $B_N = 0$,

$$C_m = \sum_{i=2}^{m}\sum_{j=1}^{i-1}\varphi_{ij}W_{ij} = \tfrac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\varphi_{ij}V_{ij}, \qquad 2 \le m \le N.$$

Theorem 1 is established by approximating the distribution of U by that of $\hat{U} + C_m$, conditional on $B_N = 0$, for some m. A smoothing lemma is then used to bound the difference in the conditional distribution function of $\hat{U} + C_m$ and $F(x)$ by an expression in terms of their Fourier transformations.

Let $\tau^2 = Npq$, $H_j = W_j sp\nu^{-1} + t\tau^{-1}$ and $\eta_j = w_j sp\nu^{-1} + t\tau^{-1}$, where $j = 1, 2, \ldots, N$ and $s, t \in \mathbb{R}$. Also let

$$D_N = \sum_{i=1}^{N}(Y_i - p)H_i \quad \text{and} \quad d_N = \{2\pi\tau P(B_N = 0)\}^{-1} \le \sqrt{2}/\pi, \tag{3.2}$$

by Höglund (1978). Since $\nu \to \infty$ and $\varepsilon_N \to 0$, choose N so large that $\nu^2 \ge 1$ and $T > \varepsilon b_N^{-1}$. The value of ε will be specified in the proof of Lemma 1. Now using (3.1) and a decomposition similar to that used by Callaert and Janssen (1978), Δ is bounded by

$$\begin{aligned}\sup_x \left| P\left(\nu^{-1}\left(\hat{U} + C_m\right) \le x | B_N = 0\right) - F(x)\right| \\ + P\left(\nu^{-1}|C_N - C_m| > \Delta_1 | B_N = 0\right) + O(\Delta_1), \quad \text{where } \Delta_1 > 0.\end{aligned} \tag{3.3}$$

This follows, as $\sup_x|F(x + \Delta_1) - F(x)|$ is bounded by a constant multiple of Δ_1 when $\nu^2 \ge 1$. Let ψ be the Fourier transformation of F. That is,

$$\psi(s) = e^{-s^2/2}\left\{1 + (is)^3\tfrac{1}{6}(p - q)(pq)^{-1/2}\left(\sum_{k=1}^{N} w_k^3\right)\left(\sum_{k=1}^{N} w_k^2\right)^{-3/2}\right.$$
$$\left. + (is)^3(q/p)^{1/2}\left(\sum_{k=2}^{N}\sum_{j=1}^{k-1} w_{kj}w_k w_j\right)\left(\sum_{k=1}^{N} w_k^2\right)^{-3/2}\right\}.$$

From a result in Erdös and Rényi (1959), which says for any random variable X,

$$E\big(e^{isX}|B_N = 0\big) = d_N \int_{-\pi\tau}^{\pi\tau} E\big\{\exp\big(isX + it\tau^{-1}B_N\big)\big\}\, dt,$$

and by Esseen's smoothing lemma, the first term in (3.3) is bounded by

$$\frac{2}{\pi}\int_0^T s^{-1}\left| d_N \int_{-\pi\tau}^{\pi\tau} E\big\{\exp\big(iD_N + is\nu^{-1}C_m\big)\big\}\, dt - \psi(s)\right| ds + O(T^{-1}). \tag{3.4}$$

Using the relationships

$$\begin{aligned}
&\exp\big(iD_N + is\nu^{-1}C_m\big)\\
&\quad = \big(1 + is\nu^{-1}C_N\big)\exp(iD_N) - is\nu^{-1}(C_N - C_m)\exp(iD_N)\\
&\qquad + \tfrac{1}{2}(is)^2\nu^{-2}C_m^2\exp(iD_N) + \tfrac{1}{6}(is)^2\nu^{-3}C_m^3\exp(iD_N)\\
&\qquad - \exp(iD_N)\big\{1 + is\nu^{-1}C_m + \tfrac{1}{2}(is)^2\nu^{-2}C_m^2 + \tfrac{1}{6}(is)^3\nu^{-3}C_m^3 - \exp\big(is\nu^{-1}C_m\big)\big\}
\end{aligned}$$

when $s \in (0, \varepsilon b_N^{-1})$, we may conclude via (3.3) and (3.4) that

$$\Delta \le \sum_{i=1}^{7} T_i + O(T^{-1}) + O(\Delta_1), \tag{3.5}$$

where

$$T_1 = \frac{2}{\pi}\int_0^{\varepsilon b_N^{-1}} s^{-1}\left| d_N \int_{-\pi\tau}^{\pi\tau} E\big\{\exp(iD_N)\big(1 + is\nu^{-1}C_N\big)\big\}\, dt - \psi(s)\right| ds,$$

$$T_2 = \frac{2}{\pi}\int_{\varepsilon b_N^{-1}}^{T} s^{-1}\left| d_N \int_{-\pi\tau}^{\pi\tau} E\big\{\exp\big(iD_N + is\nu^{-1}C_m\big)\big\}\, dt - \psi(s)\right| ds,$$

$$T_3 = \frac{2}{\pi}\int_0^{\varepsilon b_N^{-1}} \nu^{-1}\left| d_N \int_{-\pi\tau}^{\pi\tau} E\big\{(C_N - C_m)\exp(iD_N)\big\}\, dt\right| ds,$$

$$T_4 = \frac{1}{\pi}\int_0^{\varepsilon b_N^{-1}} s\nu^{-2}\left| d_N \int_{-\pi\tau}^{\pi\tau} E\big\{C_m^2\exp(iD_N)\big\}\, dt\right| ds,$$

$$T_5 = \frac{1}{3\pi}\int_0^{\varepsilon b_N^{-1}} s^2\nu^{-3}\left| d_N \int_{-\pi\tau}^{\pi\tau} E\big\{C_m^3\exp(iD_N)\big\}\, dt\right| ds,$$

$$\begin{aligned}
T_6 = \frac{2}{\pi}\int_0^{\varepsilon b_N^{-1}} s^{-1}\Bigg| d_N \int_{-\pi\tau}^{\pi\tau} E\Big[&\exp(iD_N)\big\{1 + is\nu^{-1}C_m + \tfrac{1}{2}(is)^2\nu^{-2}C_m^2\\
&+ \tfrac{1}{6}(is)^3\nu^{-3}C_m^3 - \exp\big(is\nu^{-1}C_m\big)\big\}\Big]\, dt\Bigg| ds,
\end{aligned}$$

$$T_7 = P\big(\nu^{-1}|C_N - C_m| > \Delta_1|B_N = 0\big)$$

and $\Delta_1 = \varepsilon_N\nu^{-1}$.

Theorem 1 may be proved by bounding the terms T_1 to T_7. We present the seven bounds in the following lemmas. Throughout, assume $K_1, K_2, \ldots$ are positive constants independent of p and $A_1, A_2, \ldots$ are expressions which depend only on p. The lemmas are proved in Section 4. In the sequel all summations will be over $k = 1, \ldots, N$ unless otherwise indicated.

LEMMA 1. *The term* T_1 *is bounded by*

$$A_1\Big[\Big(\sum w_k^4\Big)\Big(\sum w_k^2\Big)^{-2} + \Big(\sum|w_k^3|\Big)^2\Big(\sum w_k^2\Big)^{-3} + \Big(\sum|w_k^3|\Big)\Big(\sum w_k^2\Big)^{-2} + \Big(\sum w_k^4\Big)^{1/2}\Big(\sum w_k^2\Big)^{-3/2}\Big],$$

for some $A_1 > 0$.

LEMMA 2. *If* $(pq)^{-1/2}(\Sigma|w_k^3|)(\Sigma w_k^2)^{-3/2} \le 1$, $(q/p)^{1/2}(\Sigma w_k^2)^{-1/2} \le 1$, $b_N \le \varepsilon$ *and condition* (C) *holds, then* T_2 *is bounded by*

$$K_2\log(T)\Big[N^{1/2}\exp\Big\{-K_2^{-1}(N-m)pq\Big\} + \exp\Big\{-K_2^{-1}(\varepsilon_N T)^{2/3}\Big\}\Big] \quad \textit{for some } K_2.$$

LEMMA 3. *For some* K_3, K_4 *and* K_5,

$$T_3 \le K_3 pq\nu^{-1}(N-m)/N,$$

$$T_4 \le K_4(pq)^2\nu^{-2}$$

and

$$T_5 \le K_5(pq)^3\nu^{-3}.$$

LEMMA 4. *For some* K_6, $T_6 \le K_6\nu^{-4}N^{5/2}(N-m)^{-5/2}$.

LEMMA 5. *For some* K_7, $T_7 \le K_7 pq\,\Delta_1^{-2}\nu^{-2}(N-m)/N$.

PROOF OF THEOREM 1. Since condition (C) holds for all $s \in (\varepsilon b_N^{-1}, T)$, it also holds with ε_N replaced by $\varepsilon_N^* = \varepsilon_N$ if $\varepsilon_N \ge \nu^{-\alpha}$ and $\varepsilon_N^* = \nu^{-\alpha}$ otherwise, where α is a fixed constant $0 < \alpha < 1$. Hence, assume without loss of generality that $\varepsilon_N \ge \nu^{-\alpha}$.

Condition (2.2) implies that all the conditions of Lemma 2 are satisfied for sufficiently large N. For some $0 < K_8 < K_9 < 1$ choose m so that

$$K_8\varepsilon_N^3/\nu \le (N-m)/N \le K_9\varepsilon_N^3/\nu. \tag{3.6}$$

Using (2.2) and since $\Sigma w_k^4 \le (\Sigma|w_k^3|)^{4/3}$, it may be shown that T_1 is bounded by

$$A_2\Big\{o\Big(\sum|w_k|^3/\Big(\sum w_k^2\Big)^{3/2}\Big) + o\Big(\Big(\sum w_k^2\Big)^{-1/2}\Big)\Big\},$$

as $N \to \infty$, where all small order terms hold uniformly in p. Since $N - m \ge K_8 N\nu^{-1-3\alpha}$, then for some K_{10},

$$T_2 = K_{10}\log(T)\Big[(pq)^{-3}N^{-5/2}\nu^{3(1+3\alpha)} + \exp\Big\{-K_2^{-1}\sum w_k^2\Big(\sum|w_k^3|\Big)^{-2/3}\Big\}\Big]$$

$$= o\big((p/q)^{3/2}N^{-1/2}\big) + o\Big(\sum|w_k|^3/\Big(\sum w_k^2\Big)^{-3/2}\Big)$$

as $N \to \infty$, if $\alpha < 1/9$. Furthermore, by the right-hand inequality at (3.6),

$$T_3, T_4, T_5 = o(\nu^{-1})$$

as $N \to \infty$, and since $(N-m)N^{-1} \geq K_8 \nu^{-1-3\alpha}$ and $\Delta_1 = \varepsilon_N \nu^{-1}$,

$$T_6 \leq K_6 \nu^{-3/2+15\alpha/2} = o(\nu^{-1})$$

if $\alpha < 1/15$ and

$$T_7 \leq K_7 pq \varepsilon_N \nu^{-1} = o(\nu^{-1})$$

as $N \to \infty$. The theorem now follows via expression (3.5) if we choose $\alpha < 1/15$. □

4. Proofs of the lemmas.

PROOF OF LEMMA 1. Let $\beta(\eta) = qe^{-ip\eta} + pe^{iq\eta}$ be the characteristic function of $(Y_j - p)$. Recall the definition at (3.2) and also that $\eta_j = w_j sp\nu^{-1} + t\tau^{-1}$. Then, by independence and the inequality $|e^{ix} - 1 - ix| \leq |x|^2/2$,

$$\begin{aligned}
&E\{\exp(iD_N)(1 + is\nu^{-1}C_N)\} \\
&\quad = \prod_{j=1}^N \beta(\eta_j) + \tfrac{1}{2}is\nu^{-1} \sum_{k=1}^N \sum_{l=1}^N v_{kl} \prod_{j\neq k,\,l}^N \beta(\eta_j) \\
&\quad\quad \times E\left[\exp\{i\eta_k(Y_k - p) + i\eta_l(Y_l - p)\}(Y_k - p)(Y_l - p)\right] \qquad (4.1) \\
&\quad = \prod_{j=1}^N \beta(\eta_j) + \tfrac{1}{2}is\nu^{-1}(pq)^2 \\
&\quad\quad \times \sum_{k=1}^N \sum_{l=1}^N v_{kl} \prod_{j\neq k,\,l}^N \beta(\eta_j)(e^{i\eta_k q} - e^{-i\eta_k p})(e^{i\eta_l q} - e^{-i\eta_l p}) \\
&\quad = \prod_{j=1}^N \beta(\eta_j) - \tfrac{1}{2}is\nu^{-1}(pq)^2 \sum_{k=1}^N \sum_{l=1}^N v_{kl}\eta_k\eta_l \prod_{j\neq k,\,l}^N \beta(\eta_j) \\
&\quad\quad + \tfrac{1}{2}is\nu^{-1}(pq)^2 \sum_{k=1}^N \sum_{l=1}^N \prod_{j\neq k,\,l}^N \beta(\eta_j) \qquad (4.2) \\
&\quad\quad\quad \times v_{kl}\{O(\eta_k^2\eta_l^2) + O(|\eta_k|\eta_l^2) + O(|\eta_l|\eta_k^2)\},
\end{aligned}$$

where the large order terms here and elsewhere hold uniformly over s, t, p and the subscripts of any summation. Let us now estimate the final term at (4.2).

Fixing $0 < b < \pi$ and choosing $0 < a < b/2$, define

$$K = \left\{k \in [1, N]: |w_k| > \left(b \sum_{j=1}^N |w_j^3|\right) \Big/ \left(a \sum_{j=1}^N w_j^2\right)\right\}.$$

Using a technique similar to that used to bound expression (16) in Höglund (1978), we have that

$$\begin{aligned}
\prod_{j\neq k,\,l}^N |\beta(\eta_j)| &\leq \exp\left\{-\tfrac{1}{2}pq\theta(b) \sum_{\substack{j\notin K \\ j\neq k,\,l}} \eta_j^2\right\} \\
&\leq 3\exp\left[-\tfrac{1}{2}\theta(b)\left\{s^2\left(\tfrac{1}{2} - (a/b)\right) + t^2\left(\tfrac{1}{2} - (a/b)^3\right)\right\}\right],
\end{aligned}$$

where

$$\theta(b) = \left(\frac{2}{\pi}\frac{\pi - b}{\pi + b}\right)^2 > 0.$$

Since $a < b/2$, there exists a K_{11} such that

$$\prod_{j\neq k,\, l}^{N} |\beta(\eta_j)| \leq 3\exp\{-K_{11}(s^2 + t^2)\}. \tag{4.3}$$

Also, by Cauchy's inequality,

$$\begin{aligned}\sum_{k=1}^{N}\sum_{l=1}^{N} |v_{kl}|\eta_k^2\eta_l^2 &\leq 4\sum_{k=1}^{N}\sum_{l=1}^{N} |v_{kl}| \prod_{j=k,\, l}\left(w_j^2 s^2 p^2 \nu^{-2} + t^2\tau^{-2}\right)\\ &\leq 4s^4p^4\nu^{-4}\left(\sum_{k=1}^{N}\sum_{l=1}^{N} v_{kl}^2\right)^{1/2}\left(\sum w_k^4\right)\\ &\quad + 8s^2p^2\nu^{-2}t^2\tau^{-2}\left(\sum_{k=1}^{N}\sum_{l=1}^{N} v_{kl}^2\right)^{1/2}\left(N\sum w_k^4\right)^{1/2}\\ &\quad + 4t^4\tau^{-4}\left(N^2\sum_{k=1}^{N}\sum_{l=1}^{N} v_{kl}^2\right)^{1/2}.\end{aligned} \tag{4.4}$$

Using the inequalities $p^{-2}\nu^2 = pq\Sigma w_k^2 \leq pqN^{1/2}(\Sigma w_k^4)^{1/2}$ and

$$\sum_{k=1}^{N}\sum_{l=1}^{N} v_{kl}^2 \leq \sum_{k=1}^{N}\sum_{l=1}^{N} w_{kl}^2 = 2, \tag{4.5}$$

we find via (4.4) that

$$\sum_{k=1}^{N}\sum_{l=1}^{N} |v_{kl}|\eta_k^2\eta_l^2 \leq 4\sqrt{2}\,(s^2 + t^2)^2 pq^{-1}\nu^{-2}\left(\sum w_k^4\right)^{1/2}. \tag{4.6}$$

Using a similar argument, we have that

$$\sum_{k=1}^{N}\sum_{l=1}^{N} |v_{kl}|\,|\eta_k|\eta_l^2 \leq 2\sqrt{2}\,(|s| + |t|)^3(pq)^{-1/2}p^2\nu^{-2}\left(\sum w_k^4\right)^{1/2}, \tag{4.7}$$

and so by (4.3), (4.6) and (4.7), the final term in (4.2) is

$$O\left(|s|P_1(|s|, |t|)\exp\{-K_{11}(s^2 + t^2)\}p^3q\nu^{-3}\left(\sum w_k^4\right)^{1/2}\right), \tag{4.8}$$

where $P_1(x, y)$ is a polynomial in x and y with coefficients which are absolute constants. Now we estimate the second term in (4.2).

By (4.3) and since $|\beta(\eta)| \le 1$ and $|1 - \beta(\eta)| \le 2pq|\eta|$,

$$\begin{aligned}
&\left| \prod_{j \ne k,\, l}^{N} \beta(\eta_j) - \prod_{j=1}^{N} \big(\beta(\eta_j)\right| \\
&\quad = \prod_{j \ne k,\, l}^{N} \big|\beta(\eta_j)\big| \,\big|(1 - \beta(\eta_k))\beta(\eta_l) + (1 - \beta(\eta_l))\big| \qquad (4.9) \\
&\quad \le 6pq(|\eta_k| + |\eta_l|)\exp\big\{-K_{11}(s^2 + t^2)\big\}.
\end{aligned}$$

Furthermore, since $\Sigma_{k=1}^{N} v_{kl} = \Sigma_{l=1}^{N} v_{kl} = 0 = \Sigma w_k$,

$$\sum_{k=1}^{N}\sum_{l=1}^{N} v_{kl}\eta_k\eta_l = s^2p^2\nu^{-2} \sum_{k=1}^{N}\sum_{l=1}^{N} w_{kl}w_kw_l. \qquad (4.10)$$

By (4.2), (4.8), (4.9) and (4.10), and again applying (4.7), (4.1) equals

$$\begin{aligned}
&\prod_{j=1}^{N} \beta(\eta_j)\left[1 + \tfrac{1}{2}(is)^3p^4q^2\nu^{-3} \sum_{k=1}^{N}\sum_{l=1}^{N} w_{kl}w_kw_l\right] \qquad (4.11) \\
&\quad + O\Big(|s|P_2(|s|, |t|)\exp\big\{-K_{11}(s^2 + t^2)\big\}p^3q\nu^{-3}\Big(\sum w_k^4\Big)^{1/2}\Big)
\end{aligned}$$

for some polynomial P_2. To complete the proof we need to integrate (4.11) over s and t.

Using expression (8) of Höglund (1978), it may be shown that for some K_{12}, $\varepsilon_1 > 0$ and $|s| \le \varepsilon_1(\Sigma w_k^2)^{3/2}(\Sigma|w_k^3|)^{-1}$,

$$\begin{aligned}
&d_N\int_{-\pi\tau}^{\pi\tau} \left\{ \tfrac{1}{2}(is)^3p^4q^2\nu^{-3} \sum_{k=1}^{N}\sum_{l=1}^{N} w_{kl}w_kw_l \prod_{j=1}^{N} \beta(\eta_j) \right\} dt \\
&\quad = \tfrac{1}{2}(is)^3e^{-s^2/2}p^{-1/2}q^{1/2}\Big(\sum w_k^2\Big)^{-3/2} \sum_{k=1}^{N}\sum_{l=1}^{N} w_{kl}w_kw_l \qquad (4.12) \\
&\quad\quad + O\Big(|s|P_3(|s|)\exp\big\{-K_{12}s^2\big\}p^{-1}\sum|w_k^3|\Big(\sum w_k^2\Big)^{-2}\Big),
\end{aligned}$$

where $P_3(|s|)$ is a polynomial in $|s|$. Using the results of Robinson (1978), it may also be shown that

$$\begin{aligned}
&\int_0^{C'b_N^{-1}} |s|^{-1} \left| d_N\int_{-\pi\tau}^{\pi\tau} \prod_{j=1}^{N} \beta(\eta_j)\, dt - e^{-s^2/2} \right. \\
&\quad \left. \times \left\{ 1 + \frac{p - q}{6(pq)^{1/2}}(is)^3\Big(\sum w_k^3\Big)\Big(\sum w_k^2\Big)^{-3/2} \right\} \right| ds \qquad (4.13) \\
&\quad \le A_3\Big\{\Big(\sum w_k^4\Big)\Big(\sum w_k^2\Big)^{-2} + \Big(\sum w_k^3\Big)^2\Big(\sum w_k^2\Big)^{-3}\Big\}
\end{aligned}$$

for some C', $A_3 > 0$, both of which depend on p. Choosing $\varepsilon = \min(C', \varepsilon_1)$ and noting that $b_N \ge \Sigma|w_k^3|(\Sigma w_k^2)^{-3/2}$, the lemma follows from the results at (4.11), (4.12) and (4.13). □

PROOF OF LEMMA 2. Now by (3.2) and conditioning on $R_1, \ldots, R_N$,

$$T_2 \le 2\sqrt{2}\,\pi^{-2} \int_{\varepsilon b_N^{-1}}^{T} s^{-1} \int_{-\pi\tau}^{\pi\tau} E \prod_{k=m+1}^{N} |\beta(H_k)|\, dt\, ds + \frac{2}{\pi} \int_{\varepsilon b_N^{-1}}^{T} s^{-1} |\psi(s)|\, ds. \tag{4.14}$$

Let $\mathscr{A}_s = \{t \in (-\pi\tau, \pi\tau)$: for all $1 \le j \le N$, $|\eta_j| > 0\}$. Since $(-\pi\tau, \pi\tau) \setminus \mathscr{A}_s$ is a finite set it has Lebesgue measure zero, and so we may perform the inner integral on the right-hand side of (4.14) over $\mathscr{A}_s$ instead. By Theorem 4 of Hoeffding (1963), for each $t \in \mathscr{A}_s$,

$$E \prod_{k=m+1}^{N} |\beta(H_k)| = E \exp\Big\{ \sum \log|\beta(H_k)| \Big\} \le \big\{ E|\beta(H_1)| \big\}^{N-m}. \tag{4.15}$$

As $E|\beta(H_1)|^2 = 1 - 2pq\,N^{-1}\Sigma(1 - \cos\eta_k)$ and condition (C) implies that for all sufficiently large N, the number of indices j for which $|\eta_j - 2\pi r| > \varepsilon'$ for all integers r is at least δN, the first term on the right-hand side of (4.14) is bounded by

$$2\sqrt{2}\,\pi^{-2} \int_{\varepsilon b_N^{-1}}^{T} s^{-1} \int_{-\pi\tau}^{\pi\tau} \exp\{-(1 - \cos\varepsilon')\,\delta p\, q(N - m)\}\, dt\, ds \le 4\sqrt{2}\,\pi^{-1}(Npq)^{1/2} \log(T) \exp\{-(1 - \cos\varepsilon')\,\delta p\, q(N - m)\}. \tag{4.16}$$

Also, using a result similar to (16) in Robinson (1978), the second term on the right-hand side of (4.14) is bounded by a constant multiple of

$$\log(T)\varepsilon^3 b_N^{-3} \exp\Big\{-\tfrac{1}{2}\varepsilon^2 b_N^{-2}\Big\} \le 6^{3/2} e^{-3/2} \log(T) \exp\Big\{-\tfrac{1}{4}\varepsilon^2 (\varepsilon_N T)^{2/3}\Big\} \tag{4.17}$$

as $b_N^3 \le (\varepsilon_N T)^{-1}$. The result follows on applying the bounds at (4.16) and (4.17) to (4.14). □

PROOF OF LEMMA 3. Using (4.3) and the fact that

$$\big|E\big[(Y_r - p)\exp\{i(Y_r - p)H_r\}|R_1, \ldots, R_N\big]\big| \le pq|H_r|,$$

$$\begin{aligned} &\big|E\{(C_N - C_m)\exp(iD_N)\}\big| \\ &\quad \le E \sum_{i=m+1}^{N} \sum_{j=1}^{i-1} |V_{ij}| \prod_{r \ne i,\, j}^{N} |\beta(H_r)| \\ &\qquad \times \prod_{r=i,\, j} \big|E\big[(Y_r - p)\exp\{i(Y_r - p)H_r\}|R_1, \ldots, R_N\big]\big| \\ &\quad \le 3(pq)^2 \exp\{-K_{11}(s^2 + t^2)\} \sum_{i=m+1}^{N} \sum_{j=1}^{i-1} E|V_{ij}|\,|H_i H_j| \\ &\quad \le 3\sqrt{2}\,(pq)^2 \exp\{-K_{11}(s^2 + t^2)\}(N - m)N^{-1} \sum \eta_k^2 \end{aligned}$$

by (4.5) and since $m \le N$. On noting that $\Sigma\eta_k^2 = (pq)^{-1}(s^2 + t^2)$, T_3 is bounded

by a constant multiple of

$$pq\nu^{-1}(N-m)N^{-1}\int_0^\infty\int_{-\infty}^\infty (s^2+t^2)\exp\{-K_{11}(s^2+t^2)\}\,dt\,ds.$$

The bound for T_3 follows simply from this expression.

The proofs of the bounds for T_4 and T_5 are very similar to the proof of the T_3 bound and so are not included here. □

PROOF OF LEMMA 4. Since $E(C_m^4|R_1,\ldots,R_N)$ is bounded by a constant, it follows from (3.2) that T_6 is bounded by a constant multiple of

$$\begin{aligned}
&\nu^{-4}\int_0^{\varepsilon b_N^{-1}} s^3\int_{-\pi\tau}^{\pi\tau} E\prod_{k=m+1}^{N}|\beta(H_k)|\,dt\,ds\\
&\quad\le \nu^{-4}\int_0^{\varepsilon b_N^{-1}} s^3\int_{-\pi\tau}^{\pi\tau}\Big\{1-2pq\,N^{-1}\sum(1-\cos\eta_k)\Big\}^{(N-m)/2}\,dt\,ds \quad [\text{by (4.15)}]\\
&\quad\le \nu^{-4}\int_0^{\infty} s^3\int_{-\infty}^{\infty}\exp\{-K_{11}(N-m)N^{-1}(s^2+t^2)\}\,dt\,ds,
\end{aligned}$$

using steps similar to those used to bound (4.3). The result follows on evaluating the double integral in the preceding expression. □

PROOF OF LEMMA 5. By Chebyshev's inequality,

$$\begin{aligned}
P\big(\nu^{-1}|C_N-C_m|>\Delta_1|B_N=0\big) &\le \Delta_1^{-2}\nu^{-2}E\big\{(C_N-C_m)^2|B_N=0\big\}\\
&\le K_7pq\,\Delta_1^{-2}\nu^{-2}(N-m)/N.
\end{aligned}$$

The lemma follows. □

Acknowledgment. The authors are very grateful to John Robinson for his many helpful comments on the content of this paper, particularly for his help concerning the modification of condition (C).

REFERENCES

ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. *Ann. Statist.* **4** 108–156.

BABU, G. J. and SINGH, K. (1985). Edgeworth expansions for sampling without replacement from finite populations. *J. Multivariate Anal.* **17** 261–278.

BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1986). The Edgeworth expansion for U-statistics of degree two. *Ann. Statist.* **14** 1463–1484.

BICKEL, P. J. and VAN ZWET, W. R. (1978). Asymptotic expansions for the power of distribution free tests in the two-sample problem. *Ann. Statist.* **6** 937–1004.

CALLAERT, H. and JANSSEN, P. (1978). The Berry–Esseen theorem for U-statistics. *Ann. Statist.* **6** 417–421.

CALLAERT, H., JANSSEN, P. and VERAVERBEKE, N. (1980). An Edgeworth expansion for U-statistics. *Ann. Statist.* **8** 299–312.

COCHRAN, W. G. (1963). *Sampling Techniques*, 2nd ed. Wiley, New York.

ERDÖS, P. and RÉNYI, A. (1959). On the central limit theorem for samples from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci.* **4** 49–61.

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

HÖGLUND, T. (1978). Sampling from a finite population. A remainder term estimate. *Scand. J. Statist.* **5** 69–71.

KOKIC, P. N. (1988). Limit theorems for U-statistics in finite populations. Ph.D. dissertation, Univ. Sydney.

NANDI, H. K. and SEN, P. K. (1963). Unbiased estimation of the parameters of a finite population. *Calcutta Statist. Assoc. Bull.* **12** 124–148.

ROBINSON, J. (1978). An asymptotic expansion for samples from a finite population. *Ann. Statist.* **6** 1005–1011.

WU, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.* **14** 1261–1295.

ZHAO, L. and CHEN, X. (1987). Berry–Esseen bounds for finite-population U-statistics. *Sci. Sinica Ser. A* **30** 113–127.

DEPARTMENT OF MATHEMATICAL STATISTICS
UNIVERSITY OF SYDNEY
N.S.W. 2006
AUSTRALIA