The Annals of Probability
1978, Vol. 6, No. 2, 231-249

AN L_p BOUND FOR THE REMAINDER IN A COMBINATORIAL CENTRAL LIMIT THEOREM[1]

BY SOO-THONG HO AND LOUIS H. Y. CHEN

University of Singapore

For $n \geqq 2$ let X_{nij}, $i, j = 1, \cdots, n$, be a square array of independent random variables with finite variances and let $\pi_n = (\pi_n(1), \cdots, \pi_n(n))$ be a random permutation of $(1, \cdots, n)$ independent of the X_{nij}'s. By using Stein's method, a bound is obtained for the L_p norm $(1 \leqq p \leqq \infty)$ with respect to the Lebesgue measure of the difference between the distribution function of $(W_n - EW_n)/(\text{Var } W_n)^{\frac{1}{2}}$ and the standard normal distribution function where $W_n = \sum_{i=1}^n X_{ni\pi_n(i)}$. This result generalizes and improves a number of known results. In particular, it provides bounds for Motoo's combinatorial central limit theorem as well as the central limit theorem.

0. Introduction. For $n \geqq 2$ let X_{nij}, $i, j = 1, \cdots, n$, be a square array of independent random variables with finite variances and let $\pi_n = (\pi_n(1), \cdots, \pi_n(n))$ be a random permutation of $(1, \cdots, n)$ independent of the X_{nij}'s. This paper is concerned with the normal approximation to the distribution of $W_n = \sum_{i=1}^n X_{ni\pi_n(i)}$. A special case of W_n is the statistic $\xi_n = \sum_{i=1}^n c_{ni\pi_n(i)}$ where c_{nij}, $i, j = 1, \cdots, n$, is a square array of real numbers. A further special case is the statistic $\eta_n = \sum_{i=1}^n a_{ni} b_{n\pi_n(i)}$ where a_{ni} and b_{ni}, $i = 1, \cdots, n$, are two sequences of real numbers. Both statistics ξ_n and η_n arise in permutation tests in nonparametric inference. (See, for example, Fraser (1957) and Puri and Sen (1971).)

The literature concerning the limiting behavior of ξ_n and η_n dates back to 1944 when Wald and Wolfowitz first established the asymptotic normality of η_n with some strong sufficient conditions. These were weakened by Noether (1949) and later simplified by Hoeffding (1951) who also considered the more general statistic ξ_n. Motoo (1957) showed that a Lindeberg-type condition is sufficient for the asymptotic normality of ξ_n. The same condition was also shown to be necessary in the case of η_n by Hájek (1961). More recently Robinson (1972) obtained necessary and sufficient conditions for the moments of η_n to converge to those of a normal distribution. Kolchin and Chistyakov (1973, 1974) considered a different η_n where π_n is no longer uniform but attributes equal probabilities to only those permutations with one cycle.

It seems that so far only limit theorems have been proved for the statistics ξ_n and η_n. In this paper we use Stein's method (1972) to obtain an L_p bound, where

Received February 25, 1976; revised March 18, 1977.

[1] This paper is based on a part of Soo-Thong Ho's M. Sc. thesis written at the University of Singapore in 1975 under the supervision of Louis H. Y. Chen.

AMS 1970 *subject classifications.* Primary 60F05, 62E20; Secondary 62G99.

Key words and phrases. Normal approximation, Stein's method, combinatorial central limit theorem, L_p bound, Berry-Esseen bound, permutation tests.

$1 \leqq p \leqq \infty$, for the difference between the distribution of $(W_n - EW_n)/(\text{Var } W_n)^{\frac{1}{2}}$ and the standard normal distribution. It is interesting to note that our result contains bounds for the remainder in the central limit theorem as well as that in Motoo's limit theorem (1957), where the nature of dependence in each case bears no relationship with the other.

The notion of an L_p bound for the normal approximation was first introduced by Erickson (1973). Since $\|\cdot\|_p^p \leqq \|\cdot\|_\infty^{p-1}\|\cdot\|_1$, it suffices to consider only the L_∞ and the L_1 bounds. Our way of obtaining an L_∞ bound is inspired by Stein's proof of the Berry–Esseen theorem for i.i.d. random variables, which the second author learned from Professor Charles Stein in 1970. Since this proof has never been published, we shall present it (with some simplification) in the next section. (We wish to point out that this proof differs from that in Stein's 1972 paper.)

In the sequel all notations will be the same as in the preceding sections.

1. Stein's proof of the Berry–Esseen theorem. Let $X_1, X_2, \cdots, X_n$ be a sequence of independent and identically distributed random variables such that $EX_i = 0$, $EX_i^2 = 1/n$ and $\beta = n^{\frac{3}{2}}E|X_i|^3 < \infty$ for $i = 1, 2, \cdots, n$. The Berry–Esseen theorem states that for every real z,

$$|F(z) - \Phi(z)| \leqq C\beta/n^{\frac{1}{2}} \tag{1.1}$$

where F is the distribution function of $\sum_{i=1}^n X_i$, Φ the standard normal distribution function and C an absolute constant.

Stein's proof proceeds as follows. Let $W_n = \sum_{i=1}^n X_i$, $W_{n-1} = \sum_{i=1}^{n-1} X_i$ and μ be the common distribution of X_i's. Let $\mathscr{A}$ be the set of real valued functions defined on the real line such that if f belongs to $\mathscr{A}$, then either (1) $f(w) = w$ or (2) f is a bounded function which is the indefinite integral of a bounded measurable function f'. Then, for any function $f \in \mathscr{A}$, we have

$$EW_nf(W_n) = \sum_{i=1}^n E[X_if(\sum_{j\neq i} X_j + X_i)] ,$$

which by independence and symmetry

$$= nE[X_nf(\sum_{j=1}^{n-1} X_j + X_n)] = nE \int sf(W_{n-1} + s)\, d\mu(s) ,$$

which again by independence and the fact that $EX = 0$

$$= nE \int s[f(W_{n-1} + s) - f(W_{n-1})]\, d\mu(s)$$
$$= -nE \int_{-\infty}^{0+} s \int_{s+}^{0+} f'(W_{n-1} + t)\, dt\, d\mu(s) + nE \int_{0+}^{\infty} s \int_{0+}^{s+} f'(W_{n-1} + t)\, dt\, d\mu(s) ,$$

which by Fubini's theorem

$$= -nE \int_{-\infty}^{0+} f'(W_{n-1} + t) \int_{-\infty}^{t-} s\, d\mu(s)\, dt + nE \int_{0+}^{\infty} f'(W_{n-1} + t) \int_{t-}^{\infty} s\, d\mu(s)\, dt$$
$$= E \int f'(W_{n-1} + t)K(t)\, dt ,$$

where

$$K(t) = n \int_{t-}^{\infty} s\, d\mu(s) \qquad t > 0 \tag{1.2}$$
$$= -n \int_{-\infty}^{t-} s\, d\mu(s) \quad t \leqq 0 .$$

Hence we obtain the identity

$$EW_n f(W_n) = E \int f'(W_{n-1} + t)K(t)\,dt\,. \tag{1.3}$$

It is clear that $K(t)$ is a nonnegative function such that $K(-\infty) = 0$ and $K(+\infty) = 0$. By letting $f(w) = w$, (1.3) yields

$$\int K(t)\,dt = EW_n^{\,2} = 1 \tag{1.4}$$

showing that K is a probability density function.

To do the approximation, we choose f to be the unique bounded solution f_z of the differential equation

$$f'(w) - wf(w) = h_z(w) - \Phi(z) \tag{1.5}$$

where h_z is the indicator function of the set $(-\infty, z]$. Then (1.3) yields

$$F(z) - \Phi(z) = E \int [f_z'(W_n) - f_z'(W_{n-1} + t)]K(t)\,dt\,. \tag{1.6}$$

What remains now is to bound the right-hand side of (1.6).

We need a few lemmas. First we note that f_z is given by

$$\begin{aligned} f_z(w) &= \Phi(w)[1 - \Phi(z)]/\phi(w) \qquad \text{if} \quad w \leqq z \\ &= \Phi(z)[1 - \Phi(w)]/\phi(w) \qquad \text{if} \quad w > z \end{aligned} \tag{1.7}$$

where ϕ is the standard normal density and that

$$\int |t|K(t)\,dt = \beta/2n^{\frac{1}{2}} \tag{1.8}$$

and

$$\int |s|\,d\mu(s) \leqq \beta/n^{\frac{1}{2}}\,. \tag{1.9}$$

The following lemma can be found in Stein (1972) and is therefore stated without proof.

LEMMA 1.1. *Let f_z be defined as in* (1.7). *Then for all real w and z*, $0 \leqq f_z(w) \leqq 1$ *and* $|f_z'(w)| \leqq 1$.

Let a and b be two real numbers such that $a < b$. We define, for every real $x > 0$,

$$\begin{aligned} g_x(w) &= -\tfrac{1}{2}(b - a) - x & w &\leqq a - x \\ &= w - \tfrac{1}{2}(a + b) & a - x &\leqq w \leqq b + x \\ &= \tfrac{1}{2}(b - a) + x & b + x &\leqq w\,. \end{aligned} \tag{1.10}$$

Clearly g_x is the indefinite integral of the function $g_x'(w) = I(a - x \leqq w \leqq b + x)$ and hence belongs to $\mathscr{A}$ for every $x > 0$.

Now we prove a concentration inequality using the identity (1.3).

LEMMA 1.2. *For all real a and b such that $a < b$, we have*

$$EI(a \leqq W_{n-1} \leqq b) \leqq b - a + 2\beta/n^{\frac{1}{2}}\,.$$

PROOF. First we deduce a simple inequality. We have, by (1.8),

$$\int_{|t|>\beta/n^{\frac{1}{2}}} K(t)\,dt \leqq (n^{\frac{1}{2}}/\beta) \int_{|t|>\beta/n^{\frac{1}{2}}} |t|K(t)\,dt \leqq (n^{\frac{1}{2}}/\beta) \int |t|K(t)\,dt = \tfrac{1}{2}\,.$$

This and (1.4) yield

$$\int_{|t|\leqq\beta/n^{\frac{1}{2}}} K(t)\,dt = \int K(t)\,dt - \int_{|t|>\beta/n^{\frac{1}{2}}} K(t)\,dt \geqq 1 - \tfrac{1}{2} = \tfrac{1}{2}\,. \tag{1.11}$$

Let $f(w) = g_{\beta/n^{\frac{1}{2}}}(w)$ with $f'(w) = I(a - \beta/n^{\frac{1}{2}} \leqq w \leqq b + \beta/n^{\frac{1}{2}})$ where $g_{\beta/n^{\frac{1}{2}}}$ is defined in (1.10). Then we have

$$\begin{aligned} E \int f'(W_{n-1} + t)K(t)\,dt &= E \int I(a - \beta/n^{\frac{1}{2}} \leqq W_{n-1} \leqq b + \beta/n^{\frac{1}{2}})K(t)\,dt \\ &\geqq E \int I(a \leqq W_{n-1} \leqq b)I(|t| \leqq \beta/n^{\frac{1}{2}})K(t)\,dt\,, \end{aligned}$$

which by independence and (1.11)

$$\geqq \tfrac{1}{2}EI(a \leqq W_{n-1} \leqq b)\,.$$

This together with (1.3) yield

$$EI(a \leqq W_{n-1} \leqq b) \leqq 2EW_nf(W_n) \leqq 2E|W_nf(W_n)| \leqq b - a + 2\beta/n^{\frac{1}{2}}$$

where it is noted that $|f| \leqq \frac{1}{2}(b - a) + \beta/n^{\frac{1}{2}}$ and $E|W_n| \leqq (EW_n^2)^{\frac{1}{2}} = 1$.

The next lemma is simple and is stated without proof.

LEMMA 1.3. *Let f_z be as in* (1.7). *Then*

$$\begin{aligned} |f_z'(w + s) - f_z'(w + t)| \leqq (|t| + |s|)(|w| + 1) & \\ + I(z - t \leqq w \leqq z - s)I(s \leqq t) & \\ + I(z - s \leqq w \leqq z - t)I(s > t)\,. & \end{aligned}$$

We are now in position to prove (1.1). By (1.6), Lemma 1.3 and independence, we have

$$\begin{aligned} \sup_z |F(z) - \Phi(z)| \leqq \iint (|t| + |s|)(E|W_{n-1}| + 1)\,d\mu(s)K(t)\,dt & \\ + \iint I(s \leqq t)EI(z - t \leqq W_{n-1} \leqq z - s)\,d\mu(s)K(t)\,dt & \\ + \iint I(s > t)EI(z - s \leqq W_{n-1} \leqq z - t)\,d\mu(s)K(t)\,dt & \end{aligned}$$

which by Lemma 1.2, (1.4), (1.8), (1.9) and independence

$$\leqq 6\tfrac{1}{2}\beta/n^{\frac{1}{2}}\,.$$

Hence the theorem.

2. L_∞ versus L_1 bounds. In Stein's proof of the Berry–Esseen theorem, a crucial step is the derivation of a concentration inequality of the correct order (Lemma 1.2). One would hope that this method could easily be extended to cover the independent and nonidentically distributed case. Unfortunately this is not so. We have not been able to obtain the correct Berry–Esseen bound for this case by Stein's method. However, a somewhat weaker concentration inequality can be obtained for independent but nonidentically distributed random variables with second moments. Using this inequality, one could obtain an L_∞ bound of the form

$$C \inf_{\varepsilon>0} \{\varepsilon + \textstyle\sum_{i=1}^n EX_i^2I(|X_i| > \varepsilon)\} \tag{2.1}$$

for the normal approximation where I denotes the indicator function. This result is implied by more general results known in the literature (see, for example, Osipov (1966) and Feller (1968)). In our present problem, which is more general than the independent and nonidentically distributed case, we can only expect to obtain an L_∞ bound similar to (2.1).

In obtaining an L_1 bound, the question of concentration inequality does not arise. As a result, Stein's method works more smoothly. This fact has been pointed out by Erickson (1974). In Ho (1975), the following L_1 bound in the normal approximation is obtained for independent and nonidentically distributed random variables $X_1, X_2, \cdots, X_n$,

$$\inf_{\varepsilon>0}\{4 \textstyle\sum_{i=1}^n EX_i^2 I(|X_i| > \varepsilon) + 4\frac{1}{2} \sum_{i=1}^n E|X_i|^3 I(|X_i| \leqq \varepsilon)\} . \tag{2.2}$$

It has been pointed out in Loh (1975) that uniform truncation at arbitrary ε (in fact at 1) is as general as arbitrary truncation considered by Feller (1968).

Note that the absolute constants in (2.2) are considerably less than those in Erickson (1973). This is due, in part, to the following improvement of a lemma due to Erickson (1974) who obtained an upper of 3 instead of 1 for the second inequality. The following lemma will also be needed in the next section.

LEMMA 2.1. *Let f_z be as in* (1.7). *Then for all real w, we have*

$$\int |f_z(w)|\, dz = 1 \tag{2.3}$$

and

$$\int |f_z'(w)|\, dz \leqq 1 . \tag{2.4}$$

PROOF. (2.3) follows immediately from (1.7) by direct computation. For (2.4), let $L(w) = \int |f_z'(w)|\, dz$. It can be shown that $L(w) = 2G(w)H(w)$ where $G(w) = w\Phi(w) + \phi(w)$ and $H(w) = 1 - w[1 - \Phi(w)]/\phi(w)$. Since $L(w) = L(-w)$, one may without loss of generality assume $w \geqq 0$. For $w > 0$, we have $1 - \Phi(w) = w^{-1}\phi(w) - \int_w^\infty t^{-2}\phi(t)\, dt \geqq w^{-1}\phi(w) - w^{-2}[1 - \Phi(w)]$ and so we have the inequality $[1 - \Phi(w)]/\phi(w) \geqq w(1 + w^2)^{-1}$. By differentiation and this inequality, we can show that for $w \geqq 0$, $H(w) \leqq (1 + w^2)^{-1}$, $G'(w) \geqq 0$, and $H'(w) \leqq 0$, and that for $w \geqq 1$, $[G(w)/(1 + w^2)]' \leqq 0$. These imply that $L(w) \leqq 2G(x + 0.1)H(x) \leqq 1$ for $x \leqq w \leqq x + 0.1$ where $x = 0, 0.1, 0.2, \cdots, 1.5$ and that $L(w) \leqq 2G(w)/(1 + w^2) \leqq 2G(1.5)/(1 + 1.5^2) \leqq 1$ for $w \geqq 1.5$. Hence the lemma.

3. Statement of the main theorem and corollaries. From now on we shall drop the subscript n for brevity but shall pick it up whenever we need it. Throughout this paper, a random permutation of $(1, 2, \cdots, n)$ is an n-dimensional random vector which takes on each permutation of $(1, 2, \cdots, n)$ with probability $1/n!$.

Let X_{ij}, $i, j = 1, 2, \cdots, n$, be a square array of independent random variables such that $EX_{ij} = c_{ij}$ and $\operatorname{Var} X_{ij} = \sigma_{ij}^2 < \infty$. Also let $\pi = (\pi(1), \pi(2), \cdots, \pi(n))$

be a random permutation of $(1, 2, \cdots, n)$ which is independent of the X_{ij}'s. Further, let $W = \sum_{i=1}^n X_{i\pi(i)}$.

Define

$$c_{i-} = \frac{1}{n} \sum_{j=1}^n c_{ij}\,, \qquad c_{-j} = \frac{1}{n} \sum_{i=1}^n c_{ij}\,, \qquad c_{--} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n c_{ij}\,,$$

$$d_{ij} = c_{ij} - c_{i-} - c_{-j} + c_{--}\,,$$

$$d^2 = \frac{1}{n-1} \sum_{i=1}^n \sum_{j=1}^n d_{ij}^2 \qquad \text{and} \qquad \sigma^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^2\,.$$

Then $EW = nc_{--}$ and it will be shown in Lemma 4.3 that $\operatorname{Var} W = d^2 + \sigma^2$. Further, define

$$Y_{ij} = (X_{ij} - n^{-1}c_{--})/(d^2 + \sigma^2)^{\frac{1}{2}}$$

and

$$L_n(\varepsilon) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n EY_{ij}^2 I(|Y_{ij}| > \varepsilon)$$

$$L_{2,\text{out},n} = L_n(1)$$

$$L_{3,\text{in},n} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n E|Y_{ij}|^3 I(|Y_{ij}| \leqq 1)\,.$$

Now we state the main result.

THEOREM 3.1. *For every* $0 < \varepsilon \leqq 1$, $1 \leqq p \leqq \infty$ *and* $n \geqq 2$, *we have*

$$\|F - \Phi\|_p \leqq 24\{\varepsilon + 3L_n(\varepsilon)\} + \left(\frac{18}{n} + 40\varepsilon\right) \frac{\sigma}{(d^2 + \sigma^2)^{\frac{1}{2}}}\,,$$

where F *is the distribution function of* $(W - nc_{--})/(d^2 + \sigma^2)^{\frac{1}{2}}$ *and* Φ *the standard normal distribution function.*

The following corollaries are simple consequences of the main theorem. Unless otherwise stated, all notations will be the same as defined above.

COROLLARY 3.1. *For every* $1 \leqq p \leqq \infty$ *and* $n \geqq 2$, *we have*

$$\|F - \Phi\|_p \leqq 18/n + 96(2)^{\frac{1}{2}} L_{3,\text{out},n}^{\frac{1}{2}} + 72 L_{2,\text{in},n}\,.$$

COROLLARY 3.2. *If* $X_{ij} = EX_{ij} = c_{ij}$, *then for every* $0 < \varepsilon \leqq 1$, $1 \leqq p \leqq \infty$ *and* $n \geqq 2$, *we have*

$$\|F - \Phi\|_p \leqq 24 \left\{\varepsilon + \frac{3}{n} \sum_{i=1}^n \sum_{j=1}^n e_{ij}^2 I(|e_{ij}| > \varepsilon)\right\}$$

where $e_{ij} = (c_{ij} - nc_{--})/d$.

COROLLARY 3.3. *Suppose that for every* $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n EY_{nij}^2 I(|Y_{nij}| > \varepsilon) = 0\,.$$

Then, for every $1 \leqq p \leqq \infty$, *we have*

$$\lim_{n\to\infty} \|F_n - \Phi\|_p = 0\,.$$

To obtain Corollary 3.1 from the main theorem, let $\varepsilon = xL_{3,\mathrm{in},n}^{\frac{1}{3}}$ and use Chebyshev's inequality to get $L_n(\varepsilon) \leqq \varepsilon^{-1}L_{3,\mathrm{in},n} + L_{2,\mathrm{out},n}$. Then minimize the resulting bound with respect to x. The setting of Corollary 3.2 is due to Hoeffding (1951) who proved that $\lim_{n\to\infty} (1/n) \sum_{i=1}^n \sum_{j=1}^n |e_{ij}|^r = 0$ for $r > 2$ is sufficient for the asymptotic normality of W_n. Motoo (1957) weakened Hoeffding's condition to the Lindeberg-type condition

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n e_{nij}^2 I(|e_{nij}| > \varepsilon) = 0 .$$

Hájek (1961) showed that the Lindeberg-type condition in the case where $e_{nij} = a_{ni}b_{nj}$ is both necessary and sufficient. The sufficiency of Motoo's Lindeberg-type condition follows from Corollaries 3.2 or 3.3. It also follows from Corollary 3.2 that the bound $48\gamma^{\frac{1}{2}}$ can be obtained for $\|F - \Phi\|_p$ by letting $\varepsilon = \gamma$ where $\gamma = (3/n) \sum_{i=1}^n \sum_{j=1}^n |e_{ij}|^3$. For $p = \infty$ and under some appropriate conditions, the Lindeberg-type condition in Corollary 3.3 is shown to be also necessary in Chen [2] (Corollary 5.1). Finally we wish to thank Professor Charles Stein for suggesting to the second author in 1970 the possibility of obtaining an L_∞ bound in the special case of Corollary 3.2 using his technique.

4. Proof of Theorem 3.1. In applying Stein's method, an appropriate identity for W has to be derived. To this end, we use the following construction due to Chen (1975) who has considered the Poisson counterpart of this problem: I, J, K, L, M are random variables each uniformly distributed on $\{1, 2, \cdots, n\}$, and

$$\pi = (\pi(1), \pi(2), \cdots, \pi(n)) ,$$
$$\rho = (\rho(1), \rho(2), \cdots, \rho(n)) ,$$

and

$$\tau = (\tau(1), \tau(2), \cdots, \tau(n))$$

are random permutations of $(1, 2, \cdots, n)$ such that

(4.1) $\{I, J, K, L, M, \pi, \rho, \tau\}$ is independent of X_{ij}'s ,

(4.2) (I, K) and (L, M) are uniformly distributed on $\{(i, k) : i \neq k, i, k = 1, 2, \cdots, n\}$,

(4.3) J , (I, K) , (L, M) and τ are mutually independent,

(4.4) J , (I, K) and ρ are mutually independent,

(4.5) I and π are mutually independent,

$$\begin{aligned} \rho(\alpha) &= \tau(\alpha) && \alpha \neq I, K, \tau^{-1}(L), \tau^{-1}(M) \\ &= L && \alpha = I \\ &= M && \alpha = K \\ &= \tau(I) && \alpha = \tau^{-1}(L) \\ &= \tau(K) && \alpha = \tau^{-1}(M) \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} \pi(\alpha) &= \rho(\alpha) \quad \alpha \neq I, \rho^{-1}(J) \\ &= J \qquad \alpha = I \\ &= \rho(I) \quad \alpha = \rho^{-1}(J) \end{aligned} \tag{4.7}$$

where $\rho(\rho^{-1}(\alpha)) = \rho^{-1}(\rho(\alpha)) = \alpha$ and $\tau(\tau^{-1}(\alpha)) = \tau^{-1}(\tau(\alpha)) = \alpha$.

The reason for the introduction of all the notation (4.1) to (4.7) is the very important fact that V (defined below) is "nearly" conditionally independent of X_{IM} given τ. Such conditional independence will be used in Lemmas 4.1, 4.9 and 4.10.

Now let $(\Omega, \mathscr{B}, P)$ be the probability space on which all the above random vectors are defined and let

$\mathscr{F}$ be the σ-algebra generated by π and X_{ij}'s,
$\mathscr{G}$ the σ-algebra generated by ρ and X_{ij}'s,

and

$\mathscr{H}$ the σ-algebra generated by the X_{ij}'s.

Also let

$$\begin{aligned} Z &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij}, \\ W &= \sum_{i=1}^{n} X_{i\pi(i)}, \qquad W^* = \sum_{i\neq I} X_{i\pi(i)}, \\ U &= \sum_{i=1}^{n} X_{i\rho(i)}, \qquad U^* = \sum_{i\neq I} X_{i\rho(i)}, \\ V &= \sum_{i=1}^{n} X_{i\tau(i)}, \qquad V^{**} = \sum_{i\neq I,K} X_{i\rho(i)}, \\ \Delta V &= V - V^{**} \\ &= X_{I\tau(I)} + X_{K\tau(K)} + X_{\tau^{-1}(L)L} + X_{\tau^{-1}(M)M} - X_{\tau^{-1}(L)\tau(I)} - X_{\tau^{-1}(M)\tau(K)}. \end{aligned}$$

By using the properties of conditional expectations, it can be shown that

$$nE^{\mathscr{G}} X_{IJ} = E^{\mathscr{H}} W = Z. \tag{4.8}$$

Also, using the fact that ρ and $\{X_{ij}\}$ have the same joint distribution as π and $\{X_{ij}\}$, we have, for every $f \in \mathscr{A}$, where $\mathscr{A}$ is defined as in Section 1,

$$EZf(W) = EZf(U). \tag{4.9}$$

Now, let $f \in \mathscr{A}$. Then, using (4.8), (4.9) and the basic properties of conditional expectations, we have

$$\begin{aligned} &E[(W - Z)f(W)] \\ &\quad = EWf(W) - EZf(U) \\ &\quad = nE\{[E^{\mathscr{F}} X_{I\pi(I)}]f(W)\} - nE\{[E^{\mathscr{G}} X_{IJ}]f(U)\} \\ &\quad = nE[X_{I\pi(I)} f(W)] - nE[X_{IJ} f(U)] \\ &\quad = nE\{X_{IJ}[f(W^* + X_{IJ}) - f(U)]\} \\ &\quad = nE\{X_{IJ}[f(W^* + X_{IJ}) - f(U)][I(\rho^{-1}(J) = I) + I(\rho^{-1}(J) \neq I)]\}; \end{aligned}$$

which by the fact that $\rho^{-1}(J) = I$ implies $\pi = \rho$

$$= nE\{X_{IJ}[f(W^* + X_{IJ}) - f(U)]I(\rho^{-1}(J) \neq I)\};$$

which by (4.7)

$$\begin{aligned}
&= nE\{X_{IJ}[f(\textstyle\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} + X_{IJ} + X_{\rho^{-1}(J)\rho(I)}) \\
&\quad - f(\textstyle\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} + X_{I\rho(I)} + X_{\rho^{-1}(J)J})]I(\rho^{-1}(J) \neq I)\} \\
&= n(n-1)E\{X_{IJ}[f(\textstyle\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} + X_{IJ} + X_{\rho^{-1}(J)\rho(I)}) \\
&\quad - f(\textstyle\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} + X_{I\rho(I)} + X_{\rho^{-1}(J)J})] \\
&\quad \times [E^{I,J,\rho,\mathscr{X}} I(\rho^{-1}(J) \neq I)I(\rho^{-1}(J) = K)]\} \\
&= n(n-1)E\{X_{IJ}[f(\textstyle\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} + X_{IJ} + X_{\rho^{-1}(J)\rho(I)}) \\
&\quad - f(\textstyle\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} + X_{I\rho(I)} + X_{\rho^{-1}(J)J})][I(\rho^{-1}(J) \neq I)I(\rho^{-1}(J) = K)]\};
\end{aligned}$$

which by noting that $I \neq K$ and that $\sum_{\alpha\neq I,\rho^{-1}(J)} X_{\alpha\rho(\alpha)} = V^{**}$ on $\{\rho^{-1}(J) = K\}$

$$\begin{aligned}
&= n(n-1)E\{X_{IJ}[f(V^{**} + X_{IJ} + X_{K\rho(I)}) \\
&\quad - f(V^{**} + X_{I\rho(I)} + X_{KJ})]I(\rho^{-1}(J) = K)\};
\end{aligned}$$

which by (4.6)

$$\begin{aligned}
&= n(n-1)E\{X_{IJ}[f(V^{**} + X_{IJ} + X_{KL}) - f(V^{**} + X_{IL} + X_{KJ})]I(J = M)\} \\
&= n(n-1)E\{X_{IM}[f(V^{**} + X_{IM} + X_{KL}) - f(V^{**} + X_{IL} + X_{KM})]I(J = M)\}.
\end{aligned}$$

Since $V^{**} = V - \Delta V$, which depends only on τ, I, K, L, M and X_{ij}'s, it follows from (4.3) that J is independent of (V^{**}, I, K, L, M). Thus,

$$\begin{aligned}
(4.10)\qquad E[(W - Z)f(W)] &= (n-1)E\{X_{IM}[f(V^{**} + X_{IM} + X_{KL}) \\
&\quad - f(V^{**} + X_{IL} + X_{KM})]\};
\end{aligned}$$

which by interchanging I and K, L and M

$$\begin{aligned}
&= (n-1)E\{X_{KL}[f(V^{**} + X_{KL} + X_{IM}) - f(V^{**} + X_{KM} + X_{IL})]\} \\
&= \frac{(n-1)}{2} E\{(X_{IM} + X_{KL})[f(V^{**} + X_{IM} + X_{KL}) - f(V^{**} + X_{IL} + X_{KM})]\};
\end{aligned}$$

which by interchanging I and K

$$\begin{aligned}
&= \frac{(n-1)}{2} E\{(X_{KM} + X_{IL})[f(V^{**} + X_{KM} + X_{IL}) - f(V^{**} + X_{KL} + X_{IM})]\} \\
&= \frac{(n-1)}{4} E\{(X_{IM} + X_{KL} - X_{IL} - X_{KM})[f(V^{**} + X_{IM} + X_{KL}) \\
&\quad - f(V^{**} + X_{IL} + X_{KM})]\} \\
&= E \textstyle\int f'(V^{**} + t)K(t)\,dt;
\end{aligned}$$

where

$$(4.11)\qquad K(t) = \frac{(n-1)}{4}(X_{IM} + X_{KL} - X_{IL} - X_{KM})\psi(t, X_{IL} + X_{KM}, X_{IM} + X_{KL})$$

and

$$\begin{aligned} \psi(t, c, d) &= 1\,, && \text{if } \; c < t \leqq d\,; \\ &= -1\,, && \text{if } \; d < t \leqq c\,; \\ &= 0\,, && \text{otherwise.} \end{aligned}$$

It is clear that $K(t)$ is a nonnegative function. Hence, we obtain the identity

$$E[(W - Z)f(W)] = E \int f'(V^{**} + t)K(t)\, dt\,. \tag{4.12}$$

For the remaining part of the proof of the theorem, we shall break it up into twelve lemmas.

LEMMA 4.1. *For every $f \in \mathscr{A}$, we have*

$$E[(Z - EZ)f(W)] = \frac{1}{n} E[(W - E^{\pi}W)f(W)]\,.$$

PROOF. We have

$$\begin{aligned} &E[(Z - EZ)f(W)] \\ &\quad = \frac{1}{n} E[\textstyle\sum_{i=1}^n \sum_{j=1}^n (X_{ij} - c_{ij})f(W)] \\ &\quad = \frac{1}{n} E[\textstyle\sum_{i=1}^n \sum_{j\neq\pi(i)} (X_{ij} - c_{ij})f(W)] + \frac{1}{n} E[\sum_{i=1}^n (X_{i\pi(i)} - c_{i\pi(i)})f(W)] \\ &\quad = \frac{1}{n} E\{\textstyle\sum_{i=1}^n [E^{\pi} \sum_{j\neq\pi(i)} (X_{ij} - c_{ij})]f(W)\} + \frac{1}{n} E[(W - E^{\pi}W)f(W)] \end{aligned}$$

where it is noted that for every i, $\sum_{j\neq\pi(i)} X_{ij}$ is conditionally independent of W given π and that the first term on the right-hand side of the last equality vanishes by virtue of the fact $E^{\pi} \sum_{j\neq\pi(i)} (X_{ij} - c_{ij}) = 0$. Hence the lemma.

LEMMA 4.2. *We have*

$$E[(W - E^{\pi}W)W] = \sigma^2\,.$$

PROOF. From Lemma 4.1 we have, by letting $f(w) = w$,

$$\begin{aligned} E[(W - E^{\pi}W)W] &= nE[(Z - EZ)W] \\ &= nE[E^{\mathscr{X}}(Z - EZ)W] \\ &= nE[(Z - EZ)E^{\mathscr{X}}W] \\ &= nE[(Z - EZ)Z] \\ &= n \operatorname{Var} Z = \sigma^2\,. \end{aligned}$$

Hence the lemma.

LEMMA 4.3. *We have* Var $W = d^2 + \sigma^2$.

PROOF. By letting $f(w) = w$ in (4.12), we have

$$EW^2 - EZW = E \int K(t)\, dt\,. \tag{4.13}$$

Also, by (4.8), we have

$$EZ = E[E^{\mathscr{X}} W] = EW \tag{4.14}$$

and

$$EZW = E[E^{\mathscr{X}} ZW] = E[ZE^{\mathscr{X}} W] = EZ^2 . \tag{4.15}$$

Combining (4.13), (4.14) and (4.15), we have

$$\begin{aligned} \operatorname{Var} W &= E \int K(t)\,dt + \operatorname{Var} Z \\ &= E \int K(t)\,dt + \frac{\sigma^2}{n} . \end{aligned} \tag{4.16}$$

Now, by letting $f(w) = w$ in (4.10), we obtain

$$E \int K(t)\,dt = (n-1)E[X_{IM}(X_{IM} + X_{KL} - X_{IL} - X_{KM})] \tag{4.17}$$

which by independence

$$\begin{aligned} &= (n-1)[E(X_{IM}^2 - c_{IM}^2) + Ec_{IM}(c_{IM} + c_{KL} - c_{IL} - c_{KM})] \\ &= (n-1)\left[\frac{\sigma^2}{n} + \frac{d^2}{(n-1)}\right] . \end{aligned}$$

Combining (4.16) and (4.17), we prove the lemma.

From now on, we shall assume without loss of generality that $EW = 0$ and $\operatorname{Var} W = 1$.

In this case $Y_{ij} = X_{ij}$ and we obtain from Lemmas 4.1, 4.2, (4.14) and (4.16) the following:

$$EZf(W) = \frac{1}{n} E(W - E^{\mathscr{X}} W)f(W) ; \tag{4.18}$$

$$E|W - E^{\mathscr{X}} W| \leqq (E(W - E^{\mathscr{X}} W)^2)^{\frac{1}{2}} = (E(W - E^{\mathscr{X}} W)W)^{\frac{1}{2}} = \sigma ; \tag{4.19}$$

and

$$1 = E \int K(t)\,dt + \frac{\sigma^2}{n} . \tag{4.20}$$

We shall need (4.18), (4.19) and (4.20) later. In particular, Lemma 4.3 and (4.20) imply

$$\sigma^2 \leqq 1 \qquad \text{and} \qquad E \int K(t)\,dt \leqq 1 .$$

The next lemma is a concentration inequality which we shall need in obtaining the L_∞ bound.

LEMMA 4.4. *Let a and b be real numbers such that $a < b$. Then, for every $\varepsilon > 0$, we have*

$$E \int_{|t| \leq 2\varepsilon} I(a \leqq V^{**} \leqq b) K(t)\,dt \leqq \tfrac{1}{2}(b - a) + 2\varepsilon .$$

PROOF. Let $f(w) = g_{2\varepsilon}(w)$ as defined in (1.10) so that

$$f'(w) = I(a - 2\varepsilon \leqq w \leqq b + 2\varepsilon) .$$

Then we have

$$E \int f'(V^{**} + t)K(t)\,dt = E \int I(a - 2\varepsilon \leqq V^{**} + t \leqq b + 2\varepsilon)K(t)\,dt$$
$$\geqq E \int_{|t| \leqq 2\varepsilon} I(a \leqq V^{**} \leqq b)K(t)\,dt\,.$$

Thus, by (4.12), we obtain

$$E \int_{|t| \leqq 2\varepsilon} I(a \leqq V^{**} \leqq b)K(t)\,dt \leqq E[(W - Z)f(W)]$$
$$\leqq E|(W - Z)f(W)|\,,$$

which by the definition of f

$$\leqq [\tfrac{1}{2}(b - a) + 2\varepsilon]E|W - Z|$$
$$\leqq \tfrac{1}{2}(b - a) + 2\varepsilon\,,$$

where by (4.15) it is noted that $E|W - Z| \leqq (EW^2 - EZ^2)^{\frac{1}{2}} \leqq (EW^2)^{\frac{1}{2}} = 1$.

LEMMA 4.5. *Let c and d be two real numbers such that $c = c_1 + c_2$ and $d = d_1 + d_2$. Then for every $\varepsilon > 0$, we have*

$$\int_{|t| > 2\varepsilon} (d - c)\psi(t, c, d)\,dt \leqq 8[c_1^2 I(|c_1| > \varepsilon) + c_2^2 I(|c_2| > \varepsilon)$$
$$+ d_1^2 I(|d_1| > \varepsilon) + d_2^2 I(|d_2| > \varepsilon)]$$

where $\psi(t, c, d)$ is defined as in (4.11).

PROOF. Consider three defined ranges of values $(-\infty, -2\varepsilon)$, $[-2\varepsilon, 2\varepsilon]$ and $(2\varepsilon, \infty)$ for each of c and d and evaluate the integral in the left over each of the nine possible regions in which (c, d) lies. This yields

$$\int_{|t| > 2\varepsilon} (d - c)^2\psi(t, c, d)\,dt \leqq 2[c^2 I(|c| > 2\varepsilon) + d^2 I(|d| > 2\varepsilon)]\,.$$

By direct computation it can be shown that for every pair of real numbers (u, v),

$$(u + v)^2 I(|u + v| > 2\varepsilon) \leqq 4[u^2 I(|u| > \varepsilon) + v^2 I(|v| > \varepsilon)]\,.$$

Combining these two inequalities, we prove Lemma 4.5.

LEMMA 4.6. *For every $\varepsilon > 0$, we have*

$$E \int_{|t| > 2\varepsilon} K(t)\,dt \leqq 8L_n(\varepsilon)\,.$$

PROOF. By (4.11), we have

$$E \int_{|t| > 2\varepsilon} K(t)\,dt = \frac{n-1}{4} E \int_{|t| > 2\varepsilon} (X_{IM} + X_{KL} - X_{IL} - X_{KM})$$
$$\times\ \psi(t, X_{IL} + X_{KM}, X_{IM} + X_{KL})\,dt$$

which by Lemma 4.5

$$\leqq 2(n - 1)E[X_{IM}^2 I(|X_{IM}| > \varepsilon) + X_{KL}^2 I(|X_{KL}| > \varepsilon)$$
$$+ X_{IL}^2 I(|X_{IL}| > \varepsilon) + X_{KM}^2 I(|X_{KM}| > \varepsilon)]$$
$$= 8(n - 1)E[X_{IM}^2 I(|X_{IM}| > \varepsilon)]$$
$$\leqq 8L_n(\varepsilon)\,.$$

Hence the lemma.

LEMMA 4.7. *For every* $\varepsilon > 0$, *we have*

$$E \int_{|t|\leq 2\varepsilon} K(t)\, dt\, I(|\Delta V| > 6\varepsilon) \leqq 28L_n(\varepsilon) .$$

PROOF. By (4.11), we have

$$\begin{aligned}
E \int_{|t|\leq 2\varepsilon} & K(t)\, dt\, I(|\Delta V| > 6\varepsilon) \\
&\leqq (n-1)\varepsilon E[|X_{IM} + X_{KL} - X_{IL} - X_{KM}| I(|\Delta V| > 6\varepsilon)] \\
&\leqq (n-1)\varepsilon E[(|X_{IM}| + |X_{KL}| + |X_{IL}| + |X_{KM}|) I(|\Delta V| > 6\varepsilon)] \\
&= (n-1)\varepsilon E[(|X_{IM}| + |X_{KM}|) I(|\Delta V| > 6\varepsilon)] \\
&\quad + (n-1)\varepsilon E[(|X_{KL}| + |X_{IL}|) I(|\Delta V| > 6\varepsilon)]
\end{aligned}$$

which by interchanging I and K, L and M in the second term and noting that ΔV remains invariant

$$\begin{aligned}
&= 2(n-1)\varepsilon E[(|X_{IM}| + |X_{KM}|) I(|\Delta V| > 6\varepsilon)] \\
&= 2(n-1)\varepsilon E\{[|X_{IM}|(I(|X_{IM}| > \varepsilon) + I(|X_{IM}| \leqq \varepsilon)) \\
&\quad + |X_{KM}|(I(|X_{KM}| > \varepsilon) + I(|X_{KM}| \leqq \varepsilon))] I(|\Delta V| > 6\varepsilon)\} \\
&\leqq 2(n-1)\varepsilon E[|X_{IM}| I(|X_{IM}| > \varepsilon) + |X_{KM}| I(|X_{KM}| > \varepsilon)] \\
&\quad + 4(n-1)\varepsilon^2 E[I(|\Delta V| > 6\varepsilon)] \\
&\leqq 4(n-1) E[X_{IM}^2 I(|X_{IM}| > \varepsilon) + 4(n-1)\varepsilon^2 E[I(|X_{I\tau(I)}| > \varepsilon) \\
&\quad + I(|X_{K\tau(K)}| > \varepsilon) + I(|X_{\tau^{-1}(L)L}| > \varepsilon) + I(|X_{\tau^{-1}(M)M}| > \varepsilon) \\
&\quad + I(|X_{\tau^{-1}(L)\tau(I)}| > \varepsilon) + I(|X_{\tau^{-1}(M)\tau(K)}| > \varepsilon)]] .
\end{aligned}$$

We can show that each of the six pairs $(I, \tau(I))$, $(K, \tau(K))$, $(\tau^{-1}(L), L)$, $(\tau^{-1}(M), M)$, $(\tau^{-1}(L), \tau(I))$ and $(\tau^{-1}(M), \tau(K))$ are uniformly distributed on $\{1, 2, 3, \cdots, n\}^2$. Hence

$$\begin{aligned}
E \int_{|t|>2\varepsilon} & K(t)\, dt\, I(|\Delta V| > 6\varepsilon) \\
&\leqq 4(n-1) E X_{IM}^2 I(|X_{IM}| > \varepsilon) + 24(n-1)\varepsilon^2 E I(|X_{IM}| > \varepsilon) \\
&\leqq 28(n-1) E X_{IM}^2 I(|X_{IM}| > \varepsilon) \\
&\leqq 28 L_n(\varepsilon) ,
\end{aligned}$$

and this proves the lemma.

LEMMA 4.8. *For* f_z *defined in* (1.7), *we have*

$$|E[Zf_z(W)]| \leqq \frac{\sigma}{n} \tag{4.21}$$

and

$$\int |E[Zf_z(W)]|\, dz \leqq \frac{\sigma}{n} . \tag{4.22}$$

PROOF. By (4.18), we have

$$|E[Zf_z(W)]| = \frac{1}{n} |E[(W - E^{\tau} W) f_z(W)]|$$

which by Lemma 1.1 and (4.19)

$$\leqq \frac{1}{n} E|W - E^{\pi}W| \leqq \frac{\sigma}{n} .$$

This proves (4.21). Next,

$$\int |E[Zf_z(W)]|\, dz = \frac{1}{n} \int |E[(W - E^{\pi}W)f_z(W)]|\, dz$$
$$\leqq \frac{1}{n} E[|W - E^{\pi}W| \cdot \int |f_z(W)|\, dz] ,$$

which by Lemma 2.1

$$\leqq \frac{1}{n} E|W - E^{\pi}W| \leqq \frac{\sigma}{n} ,$$

and this proves (4.22). Hence the lemma.

For the next two lemmas, we let

$$A = \{\tau(I) \neq L, \tau(K) \neq M, \tau(I) \neq M, \tau(K) \neq L\} ,$$
$$H = c_{IM} + c_{KL} - c_{IL} - c_{KM}$$

and

$$G = X_{IM} + X_{KL} - X_{IL} - X_{KM}$$

so that we have $\int K(t)\, dt = ((n-1)/4)G^2$.

LEMMA 4.9. *For every $\varepsilon > 0$, we have*

$$E[|V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt] \leqq 1 + (1 + 4\varepsilon)\sigma .$$

PROOF. First we write

$$\begin{aligned} &E[|V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt] \\ &\leqq E[|V - E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt] + E[|E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt] \\ &= E[|V - E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt\, I(A)] + E[|V - E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt\, I(A^c)] \\ &\quad + E[|E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt] . \end{aligned} \tag{4.23}$$

Next we bound each of the three terms on the extreme right of (4.23). Since $\int K(t)\, dt$ depends on $(X_{IM}, X_{KL}, X_{IL}, X_{KM})$ it follows that $V - E^{\tau}V$ is conditionally independent of $I(A)$ and $\int K(t)\, dt$ given τ. Thus,

$$\begin{aligned} E[|V - E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt\, I(A)] &\leqq E[|V - E^{\tau}V| \int K(t)\, dt\, I(A)] \\ &= E\{[E^{\tau}|V - E^{\tau}V|][E^{\tau} \int K(t)\, dt\, I(A)]\} \\ &\leqq E\{[E^{\tau}|V - E^{\tau}V|][E^{\tau} \int K(t)\, dt]\} \\ &\leqq E|V - E^{\tau}V| = E|W - E^{\pi}W| \leqq \sigma \end{aligned} \tag{4.24}$$

where the fourth step follows from $E^{\tau} \int K(t)\, dt = E \int K(t)\, dt \leqq 1$, the fifth from the fact that (π, W) has the same distribution as (τ, V) and the last from (4.19). By (4.11), we have

$$\begin{aligned} &E[|V - E^{\tau}V| \int_{|t| \leqq 2\varepsilon} K(t)\, dt\, I(A^c)] \\ &\qquad \leqq (n-1)\varepsilon E[|V - E^{\tau}V|\, |G| I(A^c)] \\ &\qquad \leqq (n-1)\varepsilon (E(V - E^{\tau}V)^2)^{\frac{1}{2}} (EG^2 I(A^c))^{\frac{1}{2}} \\ &\qquad = (n-1)\varepsilon (E(W - E^{\pi}W)W)^{\frac{1}{2}} (E[G^2 E^{I,K,L,M} I(A^c)])^{\frac{1}{2}} . \end{aligned}$$

Now

$$E^{I,K,L,M}I(A^c) = 1 - E^{I,K,L,M}I(A) = 1 - \frac{(n-2)(n-3)[(n-2)]!}{n!}$$

$$= \frac{4n-6}{n(n-1)} \leqq \frac{4}{n}$$

and

$$\frac{n-1}{4} EG^2 = E \int K(t)\,dt \leqq 1 .$$

Thus, by Lemma 4.2, we have

$$E[|V - E^\tau V| \textstyle\int_{|t|\leqq 2\varepsilon} K(t)\,dt\, I(A^c)] \leqq 4\varepsilon\sigma . \tag{4.25}$$

By the independence of τ and $\int K(t)\,dt$, we have

$$E[|E^\tau V| \textstyle\int_{|t|\leqq 2\varepsilon} K(t)\,dt] \leqq E|E^\tau V| E \int K(t)\,dt \leqq E|V| E \int K(t)\,dt \leqq 1 \tag{4.26}$$

where it is noted that $E|V| = E|W| \leqq (EW^2)^{\frac{1}{2}} = 1$ and that $E \int K(t)\,dt \leqq 1$.

The combination of (4.23), (4.24), (4.25) and (4.26) proves the lemma.

LEMMA 4.10. *For f_z defined in* (1.7), *we have*

$$|E[f_z'(V)]E \textstyle\int K(t)\,dt - Ef_z'(V) \int K(t)\,dt| \leqq \frac{16\sigma}{n} \tag{4.27}$$

and

$$\textstyle\int |E[f_z'(V)]E \int K(t)\,dt - Ef_z'(V) \int K(t)\,dt|\,dz \leqq \frac{16\sigma}{n} . \tag{4.28}$$

PROOF. We have

$$\begin{aligned} &|Ef_z'(V)E \textstyle\int K(t)\,dt - Ef_z'(V) \int K(t)\,dt| \\ &\quad = \frac{n-1}{4} |Ef_z'(V)EG^2 - Ef_z'(V)G^2| \\ &\quad = \frac{n-1}{4} |E[f_z'(V)E^\tau G^2[I(A) + I(A^c)]] \\ &\qquad - E[f_z'(V)G^2[I(A) + I(A^c)]]| \\ &\quad \leqq \frac{n-1}{4} |E[f_z'(V)E^\tau G^2 I(A)] - E[f_z'(V)G^2 I(A)]| \\ &\qquad + \frac{n-1}{4} |E[f_z'(V)E^\tau G^2 I(A^c)] - E[f_z'(V)G^2 I(A^c)]| \end{aligned} \tag{4.29}$$

where it is noted that $EG^2 = E^\tau G^2$ by independence of τ and G.

Since V is conditionally independent of $I(A)$ and G given τ, we have

$$E[f_z'(V)G^2 I(A)] = E\{[E^\tau f_z'(V)][E^\tau G^2 I(A)]\} = E[f_z'(V)E^\tau G^2 I(A)] .$$

This implies that the first term on the extreme right of (4.29) vanishes.

By the conditional independence of V and $H^2 I(A^c)$ given τ, we have

$$E[f_z'(V)H^2 I(A^c)] = E\{[E^\tau f_z'(V)][E^\tau H^2 I(A^c)]\} = E[f_z'(V)E^\tau H^2 I(A^c)] .$$

Thus

$$|Ef_z'(V)E \int K(t)\,dt - Ef_z'(V) \int K(t)\,dt|$$

(4.30)
$$\leqq \frac{n-1}{4} |E[f_z'(V)E^{\tau}G^2 I(A^c)] - E[f_z'(V)G^2 I(A^c)]|$$

$$= \frac{n-1}{4} |E[f_z'(V)E^{\tau}(G^2 - H^2)I(A^c)] - E[f_z'(V)(G^2 - H^2)I(A^c)]|$$

which by Lemma 1.1

$$\leqq \frac{n-1}{2} E[|G^2 - H^2|I(A^c)] = \frac{n-1}{2} E[|G^2 - H^2|E^{I,K,L,M}I(A^c)]$$

$$\leqq \frac{2(n-1)}{n} E|G^2 - H^2| \leqq \frac{2(n-1)}{n} (E(G-H)^2 E(G+H)^2)^{\frac{1}{2}},$$

where, as before, we noted that $E^{I,K,L,M}I(A^c) \leqq 4/n$.

Now,

$$E(G-H)^2 \leqq E[E^{I,K,L,M}(G-H)^2]$$
$$= E[\mathrm{Var}^{I,K,L,M}G] = E[\sigma_{IM}^2 + \sigma_{KL}^2 + \sigma_{IL}^2 + \sigma_{KM}^2] = 4\sigma^2/n$$

and

$$E(G+H)^2 = E(G^2 + 3H^2) = \frac{4}{n-1} E \int K(t)\,dt + 3EH^2$$
$$\leqq \frac{4}{n-1} + 3EH^2 \leqq \frac{16}{n-1}$$

noting that (4.11) implies $EG^2 = (4/(n-1))E \int K(t)\,dt$ and that H is a special case of G. These two inequalities and (4.30) prove (4.27).

Next, by (4.30), we have

$$\int |Ef_z'(V)E \int K(t)\,dt - Ef_z'(V) \int K(t)\,dt|\,dz$$
$$\leqq \frac{n-1}{4} \int |E[f_z(V)E^{\tau}(G^2 - H^2)I(A^c)] - E[f_z'(V)(G^2 - H^2)I(A^c)]|\,dz$$
$$\leqq \frac{n-1}{4} E[\int |f_z'(V)|\,dz\,|E^{\tau}(G^2 - H^2) - (G^2 - H^2)|I(A^c)]$$

which by Lemma 2.1 and a consequence in (4.30)

$$\leqq \frac{n-1}{2} E[|G^2 - H^2|I(A^c)] \leqq \frac{16\sigma}{n}.$$

This proves (4.28). Hence the lemma.

The next lemma is a simple consequence of Lemmas 1.1 and 2.1 and is therefore stated without proof.

LEMMA 4.11. *Let h_z and f_z be defined as in* (1.5) *and* (1.7) *respectively. Then for all real s and t such that $|s| \leqq 6\varepsilon$ and $|t| \leqq 2\varepsilon$ we have*

(4.31)
$$|wf_z(w) - (w+t-s)f_z(w+t-s)| \leqq 8\varepsilon(|w|+1)$$
$$|h_z(w+s) - h_z(w+t)| \leqq I(z - 6\varepsilon \leqq w \leqq z + 6\varepsilon)$$

and

$$\int |f_z'(w) - f_z'(w + t - s)|\, dz \leqq 8\varepsilon(|w| + 2)\,. \tag{4.32}$$

Now we prove the last lemma.

LEMMA 4.12. *For all* $\varepsilon > 0$, *we have*

$$E \int |f_z'(V) - f_z'(V^{**} + t)|K(t)\, dt \leqq 24[\varepsilon + 3L_n(\varepsilon)] + 8\varepsilon(1 + 4\varepsilon)\sigma \tag{4.33}$$

and

$$E \iint |f_z'(V) - f_z'(V^{**} + t)|\, dz\, K(t)\, dt \\ \leqq 24[\varepsilon + 3L_n(\varepsilon)] + 8\varepsilon(1 + 4\varepsilon)\sigma\,. \tag{4.34}$$

PROOF. Since $V = V^{**} + \Delta V$, we have, using (1.5),

$$\begin{aligned} &E \int |f_z'(V) - f_z'(V^{**} + t)|K(t)\, dt \\ &\quad = E \int_{|t|>2\varepsilon} |f_z'(V) - f_z'(V^{**} + t)|K(t)\, dt \\ &\qquad + E \int_{|t|\leqq 2\varepsilon} |f_z'(V) - f_z'(V^{**} + t)|I(|\Delta V| > 6\varepsilon)K(t)\, dt \\ &\qquad + E \int_{|t|\leqq 2\varepsilon} |Vf_z(V) - (V + t - \Delta V)f_z(V + t - \Delta V)|I(|\Delta V| \leqq 6\varepsilon)K(t)\, dt \\ &\qquad + E \int_{|t|\leqq 2\varepsilon} |h_z(V^{**} + \Delta V) - h_z(V^{**} + t)|I(|\Delta V| \leqq 6\varepsilon)K(t)\, dt\,, \end{aligned}$$

which by Lemma 1.1 and (4.31)

$$\begin{aligned} &\leqq 2E \int_{|t|>2\varepsilon} K(t)\, dt + 2E \int_{|t|\leqq 2\varepsilon} I(|\Delta V| > 6\varepsilon)K(t)\, dt \\ &\quad + 8\varepsilon E \int_{|t|\leqq 2\varepsilon} (|V| + 1)K(t)\, dt \\ &\quad + E \int_{|t|\leqq 2\varepsilon} I(z - 6\varepsilon \leqq V^{**} \leqq z + 6\varepsilon)K(t)\, dt\,, \end{aligned}$$

which by Lemmas 4.4, 4.6, 4.7 and 4.9

$$\leqq 24[\varepsilon + 3L_n(\varepsilon)] + 8\varepsilon(1 + 4\varepsilon)\sigma\,.$$

This proves (4.33).

Next, we have

$$\begin{aligned} &E \int [\int |f_z'(V) - f_z'(V^{**} + t)|\, dz]K(t)\, dt \\ &\quad \leqq E \int_{|t|>2\varepsilon} [\int |f_z'(V) - f_z'(V^{**} + t)|\, dz]K(t)\, dt \\ &\qquad + E \int_{|t|\leqq 2\varepsilon} [\int |f_z'(V) - f_z'(V^{**} + t)|\, dz]I(|\Delta V| > 6\varepsilon)K(t)\, dt \\ &\qquad + E \int_{|t|\leqq 2\varepsilon} [\int |f_z'(V) - f_z'(V^{**} + t)|\, dz]I(|\Delta V| \leqq 6\varepsilon)K(t)\, dt\,, \end{aligned}$$

which by Lemma 2.1 and (4.32)

$$\begin{aligned} &\leqq 2E \int_{|t|>2\varepsilon} K(t)\, dt + 2E[\int_{|t|\leqq 2\varepsilon} K(t)\, dt\, I(|\Delta V| > 6\varepsilon)] \\ &\quad + 8\varepsilon E \int_{|t|\leqq 2\varepsilon} (|V| + 2)K(t)\, dt\,, \end{aligned}$$

which by Lemmas 4.6, 4.7 and 4.9 again

$$\leqq 24[\varepsilon + 3L_n(\varepsilon)] + 8\varepsilon(1 + 4\varepsilon)\sigma\,.$$

This proves (4.34). Hence the lemma.

With the above lemmas, we are now in the position to bound $\|F - \Phi\|_p$. We choose f in the identity (4.12) to be f_z in (1.7) which is the unique bounded solution of the differential equation (1.5). Then for all real z we have

$$|F(z) - \Phi(z)| = |Eh_z(W) - \Phi(z)| = |Ef_z'(W) - EWf_z(W)|$$
$$= |Ef_z'(W) - E \int f_z'(V^{**} + t)K(t)\,dt - EZf_z(W)|$$

which by (4.20)

$$= \left|[Ef_z'(W)]\left[E \int K(t)\,dt + \frac{\sigma^2}{n}\right] - E \int f_z'(V^{**} + t)K(t)\,dt - EZf_z(w)\right|$$
$$\leqq |E \int [f_z'(V) - f_z'(V^{**} + t)]K(t)\,dt| + |Ef_z'(V)E \int K(t)\,dt - E \int f_z'(V)K(t)\,dt|$$
$$+ \frac{\sigma^2}{n} E|f_z'(W)| + |EZf_z(W)|$$

where we used $Ef'(W) = Ef'(V)$.

Thus, using Lemmas 1.1 and 2.1 together with Lemmas 4.8, 4.10, 4.12 and the inequality $\sigma^2 \leqq 1$, we obtain

$$\|F - \Phi\|_\infty \leqq 24[\varepsilon + 3L_n(\varepsilon)] + \left[\frac{18}{n} + 8\varepsilon(1 + 4\varepsilon)\right]\sigma$$

and

$$\|F - \Phi\|_1 \leqq 24[\varepsilon + 3L_n(\varepsilon)] + \left[\frac{18}{n} + 8\varepsilon(1 + 4\varepsilon)\right]\sigma .$$

These, together with $\|\cdot\|_p^p \leqq \|\cdot\|_1 \cdot \|\cdot\|_\infty^{p-1}$, prove (3.1). The proof of the theorem is completed.

Acknowledgment. We wish to thank the referee for suggesting Corollary 3.1 and his helpful suggestions on an earlier version of the paper.

REFERENCES

[1] CHEN, LOUIS H. Y. (1975). An approximation theorem for sums of certain randomly selected indicators. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **33** 69–74.

[2] CHEN, LOUIS H. Y. (1978). Two central limit problems for dependent random variables. To appear.

[3] ERICKSON, R. V. (1973). On an L_p version of the Berry-Esseen theorem for independent and m-dependent variables. *Ann. Probability* **1** 497–503.

[4] ERICKSON, R. V. (1974). L_1 bounds for asymptotic normality of m-dependent sums using Stein's technique. *Ann. Probability* **2** 522–529.

[5] FELLER, W. (1968). On the Berry-Esseen theorem. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **10** 261–268.

[6] FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics*. Wiley, New York.

[7] HÁJEK, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Ann. Math. Statist.* **32** 506–523.

[8] HO, SOO-THONG (1975). The remainders in the central limit theorem and in a generalization of Hoeffding's combinatorial limit theorem. M. Sc. thesis, Univ. of Singapore.

[9] HOEFFDING, W. (1951). A combinatorial limit theorem. *Ann. Math. Statist.* **22** 558–566.

[10] KOLCHIN, V. F. and CHISTYAKOV, V. P. (1973). On a combinatorial limit theorem. *Theor. Probability Appl.* **18** 728–739.

[11] KOLCHIN, V. F. and CHISTYAKOV, V. P. (1974). On limit distributions of a statistic. *Theor. Probability Appl.* **19** 359–365.

[12] LOH, WEI-YIN (1975). On the normal approximation for sums of mixing random variables. M. Sc. thesis, Univ. of Singapore.

[13] MOTOO, M. (1957). On the Hoeffding's combinatorial central limit theorem. *Ann. Inst. Statist. Math.* **8** 145–154.

[14] NOETHER, G. E. (1949). On a theorem by Wald and Wolfowitz. *Ann. Math. Statist.* **20** 455–458.

[15] OSIPOV, L. V. (1966). Refinement of Lindeberg's theorem. *Theor. Probability Appl.* **11** 299–302.

[16] PURI, M. L. and SEN, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*. Wiley, New York.

[17] ROBINSON, J. (1972). A converse to a combinatorial limit theorem. *Ann. Math. Statist.* **43** 2053–2057.

[18] STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 583–602, Univ. of California Press.

[19] WALD, A. and WOLFOWITZ, J. (1944). Statistical tests on permutations of observations. *Ann. Math. Statist.* **15** 358–372.

TEMASEK JUNIOR COLLEGE
BEDOK SOUTH AVENUE 2
SINGAPORE 16
REPUBLIC OF SINGAPORE

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SINGAPORE
BUKIT TIMAH ROAD
SINGAPORE 10
REPUBLIC OF SINGAPORE