Z. Wahrscheinlichkeitstheorie verw. Gebiete
66, 379–386 (1984)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

An Estimate of the Remainder in a Combinatorial Central Limit Theorem

E. Bolthausen

Technische Universität Berlin, Fachbereich Mathematik, Straße des 17. Juni 135, D-1000 Berlin 12

§ 1. Introduction

Let $A=(a_{ij})$ be a $n \times n$ matrix of real numbers. Let

$$\mu_A = n a_{..}, \qquad \sigma_A^2 = \sum_{i,j} (a_{ij} - a_{i.} - a_{.j} + a_{..})^2/(n-1),$$

where

$$a_{i.} = \sum_j a_{ij}/n, \qquad a_{.j} = \sum_i a_{ij}/n, \qquad a_{..} = \sum_{i,j} a_{ij}/n^2.$$

Let further $\hat{a}_{ij} = (a_{ij} - a_{i.} - a_{.j} + a_{..})/\sigma_A$. A theorem which has been proved under various conditions by Hoeffding [3], Motoo [5] and others states that if π is uniformly distributed on the set of permutations of $\{1, 2, \ldots, n\}$ then $T_A = (\sum_i a_{i\pi(i)} - \mu)/\sigma = \sum_i \hat{a}_{i\pi(i)}$ is approximately standard normally distributed. We shall investigate the rate of convergence.

Estimates have been obtained by von Bahr [7] and Ho and Chen [2], but they yield the rate $O(n^{-1/2})$ only under some boundedness conditions, like $\sup_{i,j} |\hat{a}_{ij}| = O(n^{-1/2})$.

The special case where $a_{ij} = e_i d_j$ is of particular interest in non-parametric statistics and has been discussed by many authors, e.g. by Husková [4] in the case where the d_j satisfy boundedness and smoothness conditions and most recently and successfully by Does [1], whose results may cover most cases of statistical interest. Usually, d_j is assumed to be given by so-called score generating functions, e.g. $d_j = J(j/(n+1))$, where J is a function on (0, 1), satisfying some smoothness assumption. The main advantages of the approach given here are that no smoothness is needed and the a_{ij} are completely general. Von Bahr and Ho and Chen allow the a_{ij} to be random independent of π. The extensions to cover this case are straightforward and therefore omitted.

Theorem. *There is an absolute constant $K>0$, such that for all A with $\sigma_A^2>0$*

$$\sup_t |P(T_A \leqq t) - \Phi(t)| \leqq K \sum_{i,j} |\hat{a}_{ij}|^3/n$$

where Φ is the standard normal distribution function.

If one takes a sequence $(a_{ij}^{(n)})$ of $n \times n$ matrices, the theorem gives the convergence rate $n^{-1/2}$ if $\sum_{i,j} |\hat{a}_{ij}^{(n)}|^3/\sqrt{n}$ remains bounded and $\sigma^2=1$.

The proof given in § 3 is simpler than the Fourier theoretic approaches used e.g. by Does [1]. It is based on an improvement of the Stein method. Stein's method has also been used by Ho and Chen.

In §2 a proof of the classical Berry-Esseen Theorem is given using a version of the Stein method. This has also been done by Ho and Chen [2], but their proof depends on a concentration inequality and seems not to work for non-identically distributed variables. The approach given here is more flexible and the extension to the non-identically distributed case is straightforward. The proof in the simple situation of §2 gives the motivation for the proof of the theorem stated above.

§ 2. A Proof of the Classical Berry-Esseen Theorem

If $n \in \mathbb{N}$, $\gamma \geqq 1$, let $\mathscr{L}(n, \gamma)$ be the set of sequences $\underline{X} = \{X_1, \ldots, X_n\}$ of random variables, such that $X_1, \ldots, X_n$ are i.i.d. and $EX_i=0$, $EX_i^2=1$, $E|X_i|^3=\gamma$. If $\gamma<1$ $\mathscr{L}(n, \gamma)=\emptyset$. Let $S_k = \sum_{i=1}^{k} X_i/\sqrt{n}$, $1 \leqq k \leqq n$. If $z, x \in \mathbb{R}$, $\lambda>0$, let

$$h_{z,\lambda}(x) = ((1+(z-x)/\lambda) \wedge 1) \vee 0, \qquad h_{z,0}(x) = 1_{(-\infty, z]}(x).$$

Let

$$\delta(\lambda, \gamma, n) = \sup\{|E(h_{z,\lambda}(S_n)) - \Phi(h_{z,\lambda})| : z \in \mathbb{R}, \underline{X} \in \mathscr{L}(n, \gamma)\}.$$

Here $\Phi(g)$ is the standard normal expectation of g.

We write $\delta(\gamma, n) = \delta(0, \gamma, n)$. The Berry-Esseen theorem states that

$$\sup\{\sqrt{n}\,\delta(\gamma, n)/\gamma : \gamma \geqq 1, n \in \mathbb{N}\} < \infty. \tag{2.1}$$

By using $h_{z,0} \leqq h_{z,\lambda} \leqq h_{z+\lambda,0}$ one obtains

$$\delta(\gamma, n) \leqq \delta(\lambda, \gamma, n) + \lambda/\sqrt{2\pi}. \tag{2.2}$$

It obviously suffices to bound $\sqrt{n}\,\delta(\gamma, n)/\gamma$ for $n \geqq 2$ which is assumed from now on. We simply write h instead of $h_{z,\lambda}$ if there is no danger of confusion.

Let $f(x) = e^{x^2/2} \int_{-\infty}^{x} (h(z) - \Phi(h))\, e^{-z^2/2}\, dz$, which satisfies

$$f'(x) - x f(x) = h(x) - \Phi(h). \tag{2.3}$$

If $x \leqq 0$, then $|f(x)| \leqq \Phi(x)/\varphi(x)$, where φ is the standard normal density and if $x>0$: $|f(x)| \leqq (1-\Phi(x))/\varphi(x)$. Therefore

$$|f(x)| \leqq 1; \quad |x f(x)| \leqq 1; \quad |f'(x)| \leqq 2 \quad \text{for all } x. \tag{2.4}$$

(The last estimate by using (2.3)). From this one has

$$\begin{aligned}|f'(x+y)-f'(x)| &= |y f(x+y) + x(f(x+y)-f(x)) + h(x+y) - h(x)| \\ &\leqq |y| \left(1 + 2|x| + \frac{1}{\lambda} \int_0^1 1_{[z,\, z+\lambda]}(x+sy)\, ds\right),\end{aligned} \tag{2.5}$$

$$\begin{aligned} E(f'(S_n) - S_n f(S_n)) &= E(f'(S_n) - \sqrt{n}\, X_n f(S_n)) \\ &= E\left\{ f'(S_n) - f'(S_{n-1}) - X_n^2 \int_0^1 \left(f'\left(S_{n-1} + t \frac{X_n}{\sqrt{n}}\right) - f'(S_{n-1})\right) dt \right\}\end{aligned} \tag{2.6}$$

if $\underline{x} \in \mathscr{L}(n, \gamma)$. Using (2.5) one obtains

$$\begin{aligned} E|f'(S_n) - f'(S_{n-1})| &\leqq E\left\{ \frac{|X_n|}{\sqrt{n}} \left(1 + 2|S_{n-1}| + \frac{1}{\lambda} \int_0^1 1_{[z,\, z+\lambda]} \left(S_{n-1} + t \frac{X_n}{\sqrt{n}}\right) dt\right)\right\} \\ &\leqq \frac{c}{\sqrt{n}} (1 + \delta(\gamma, n-1)/\lambda)\end{aligned}$$

where we used the independence of S_{n-1} and X_n. Here and in the future c is used as a positive constant which depends only on the formula where it appears.

Similarly

$$E\left| X_n^2 \int_0^1 \left(f'\left(S_{n-1} + t \frac{X_n}{\sqrt{n}}\right) - f'(S_{n-1})\right) dt \right| \leqq \frac{c\gamma}{\sqrt{n}} (1 + \delta(\gamma, n-1)/\lambda).$$

Implementing these estimates into (2.6) and using (2.3) and (2.2) one obtains

$$\delta(\gamma, n) \leqq \frac{c\gamma}{\sqrt{n}} (1 + \delta(\gamma, n-1)/\lambda) + \lambda/\sqrt{2\pi}.$$

Choosing now $\lambda = 2c\gamma/\sqrt{n}$ (c here the same as above), then $\delta(\gamma, n) \leqq c\gamma/\sqrt{n} + \delta(\gamma, n-1)/2$. Using $\delta(\gamma, 1) \leqq 1$, this proves (2.1).

§ 3. Proof of the Theorem

c again denotes a constant which depends only on the formula where it appears. In contrast, $c_1, c_2, \ldots$ are positive constants which depend on nothing.

Let $\beta_A = \sum_{i,j} |\hat{a}_{ij}|^3$. If $n_0 \in \mathbb{N}$ and $\varepsilon_0 > 0$ are arbitrary but fixed then the statement of the theorem is true if $n \leqq n_0$ or $\beta_A > \varepsilon_0 n$ (β_A is bounded from below). Therefore, we assume that $n > n_0$ and $\beta_A \leqq \varepsilon_0 n$ where ε_0, n_0 will be specified later on but $n_0 \geqq 4$.

We first need a truncation: Let

$$a'_{ij}=\begin{cases}\hat{a}_{ij} & \text{if } |\hat{a}_{ij}|\leqq 1/2\\ 0 & \text{if } |\hat{a}_{ij}|>1/2\end{cases}$$

and $\Gamma=\{(i,j):|\hat{a}_{ij}|>1/2\}$. Clearly, $|\Gamma|$, the number of elements in Γ, is at most $8\beta_A$. Therefore,

$$\begin{aligned}P(\sum_i a'_{i\pi(i)}\neq T_A)&\leqq P(\sum_i 1_\Gamma(i,\pi(i))\geqq 1)\\ &\leqq E(\sum_i 1_\Gamma(i,\pi(i)))\\ &=|\Gamma|/n\leqq 8\beta_A/n. \qquad (3.1)\end{aligned}$$

If A' is the matrix $(a'_{i,j})$, then

$$\begin{aligned}|\mu_{A'}|&=\left|\frac{1}{n}\sum_{i,j}a'_{ij}\right|\leqq\frac{1}{n}\sum_{(i,j)\in\Gamma}|\hat{a}_{ij}|\\ &\leqq\frac{1}{n}|\Gamma|^{2/3}\beta_A^{1/3}\leqq c\beta_A/n. \qquad (3.2)\end{aligned}$$

We claim that

$$|\sigma_{A'}^2-1|\leqq c\beta_A/n. \qquad (3.3)$$

$$\begin{aligned}|\sigma_{A'}^2-1|&=|E(\sum_i a'_{i\pi(i)})^2-\mu_{A'}^2-E(\sum_i\hat{a}_{i\pi(i)})^2|\\ &\leqq|E(\sum_i a'^2_{i\pi(i)}+\sum_{i\neq j}a'_{i\pi(i)}a'_{j\pi(i)}-\sum_i\hat{a}^2_{i\pi(i)}-\sum_{i\neq j}\hat{a}_{i\pi(i)}\hat{a}_{j\pi(j)})|+\mu_A^2\\ &=\left|-\frac{1}{n}\sum_{(i,j)\in\Gamma}\hat{a}_{ij}^2+\frac{1}{n(n-1)}\sum_{i\neq j}\sum_{k\neq l}(a'_{ik}a'_{jl}-\hat{a}_{ik}\hat{a}_{jl})\right|+\mu_{A'}^2\\ &\leqq\frac{1}{n}\sum_{(i,j)\in\Gamma}\hat{a}_{ij}^2+\frac{2}{n(n-1)}\Big|\sum_{\substack{i\neq j,\,k\neq l\\(i,k)\in\Gamma}}\hat{a}_{ik}\hat{a}_{jl}\Big|+\mu_{A'}^2,\end{aligned}$$

$$\frac{1}{n}\sum_{(i,j)\in\Gamma}\hat{a}_{ij}^2\leqq c\beta_A/n \quad\text{and}\quad \frac{1}{n(n-1)}\Big|\sum_{(i,k)\in\Gamma}\sum_{j\neq i,\,l\neq k}\hat{a}_{ik}\hat{a}_{jl}\Big|$$

$$=\frac{1}{n(n-1)}\sum_{(i,k)\in\Gamma}\hat{a}_{ik}^2\leqq c\beta_A/n.$$

Therefore (3.3) is true.

If we choose ε'_0 small enough (depending on c in (3.3)), one has

$$|\sigma_{A'}^2-1|\leqq 1/3 \qquad \text{if } \beta_A\leqq\varepsilon'_0 n.$$

Therefore,

$$\begin{aligned}|\hat{a}'_{ij}|&=\left|\frac{1}{\sigma_{A'}}(a'_{ij}-a'_{i.}-a'_{.j}+a'_{..})\right|\\ &\leqq\tfrac{3}{4}+\tfrac{3}{2}(|a'_{i.}|+|a'_{.j}|+|a'_{..}|).\end{aligned}$$

By taking ε_0' small enough, this is easily seen to be $\leqq 1$. Therefore,

$$|\hat{a}_{ij}'| \leqq 1. \tag{3.4}$$

A simple calculation shows that

$$\beta_{A'} = \sum_{i,j} |\hat{a}_{ij}'|^3 \leqq c_1 \beta_A \tag{3.5}$$

if $\gamma > 0$, let $M_n(\gamma)$ be the set of $n \times n$ matrices B satisfying

$$\sigma_B^2 \neq 0, \quad |\hat{b}_{ij}| \leqq 1 \quad \text{and} \quad \sum_{i,j} |\hat{b}_{ij}|^3 \leqq \gamma.$$

Let

$$\delta(\lambda, \gamma, n) = \sup\{|E h_{z,\lambda}(T_B) - \Phi(h_{z,\lambda})| : z \in \mathbb{R}, B \in M_n(\gamma)\}$$

and $\delta(\gamma, n) = \delta(0, \gamma, n)$.

The considerations above show that if A is a $n \times n$ matrix with $\sigma_A^2 > 0$ and $\beta_A \leqq \varepsilon_0' n$, then

$$\begin{aligned}
\sup_t |P(T_A \leqq t) - \Phi(t)| &\leqq \sup_t |P(\sum_i a_{i\pi(i)}' \leqq t) - \Phi(t)| + \frac{8\beta_A}{n} \\
&\leqq \delta(c_1 \beta_{A'} n) + \sup_t |\Phi((t - \mu_{A'})/\sigma_{A'}) - \Phi(t)| + \frac{8\beta_A}{n} \\
&\leqq \delta(c_1 \beta_{A'} n) + c\beta_A/n.
\end{aligned} \tag{3.6}$$

We now want to show that $\delta(\gamma, n) \leqq c\gamma/n$, which together with (3.6) proves the theorem.

As $\beta_A \geqq c\sqrt{n}$, we may assume that $\gamma \geqq 1$. We fix $A \in M_n(\gamma)$ and estimate $|E h_{z,\lambda}(T_A) - \Phi(h_{z,\lambda})|$. Of course, we may assume $a_{ij} = \hat{a}_{ij}$ and therefore $a_{i.} = a_{.j} = 0$; $\frac{1}{n-1} \sum_{i,j} a_{ij}^2 = 1$. We denote the set of these matrices by $M_n^0(\gamma)$.

In order to apply the method of §2 we need some manipulations on the permutations which replace the independence of the summands in §2.

We define a random element (I_1, I_2, J_1, J_2) in N^4, where $N = \{1, \ldots, n\}$, in the following way: (I_1, I_2, J_1) is uniformly distributed on N^3, and given this, one has $J_2 = J_1$ on $\{I_1 = I_2\}$ and J_2 is uniformly distributed on $N - \{J_1\}$ on $\{I_1 \neq I_2\}$. Let π_1 be a random permutation, which is uniformly distributed on the permutations of N and independent of (I_1, I_2, J_1, J_2). Define

$$I_3 = \pi_1^{-1}(J_1), \quad I_4 = \pi_1^{-1}(J_2), \quad J_3 = \pi_1(I_1), \quad J_4 = \pi_1(I_2). \tag{3.7}$$
$$\underline{I} = (I_1, I_2, I_3, I_4), \quad \underline{J} = (J_1, J_2, J_3, J_4).$$

Of course, $I_1 = I_2$ holds if and only if $I_3 = I_4$. For each fixed $\underline{i} = (i_1, i_2, i_3, i_4) \in N^4$ which satisfies the condition $i_1 = i_2 \Leftrightarrow i_3 = i_4$, we fix once for all a permutation $t(\underline{i})$ of N, which maps i_1 to i_4 and i_2 to i_3 and which leaves the numbers outside $\{i_1, i_2, i_3, i_4\}$ fixed. Let further $s(i_1, i_2)$ be the transposition of i_1 and i_2. We put $\pi_2 = \pi_1 \circ t(\underline{I})$, $\pi_3 = \pi_2 \circ s(I_1, I_2)$.

Lemma. a) π_1, π_2, π_3 *have the same law.*
b) π_2 *is independent of* (I_1, J_1).

Proof. A simple calculation shows that $\underline{I}$ and π_1 are independent. Given $\underline{I}$, π_2 is a one to one function of π_1. Therefore, π_2 is also uniformly distributed and independent of $\underline{I}$.

Given π_2, (I_1, I_2) is a one to one function of (I_1, J_1). As (I_1, I_2) is uniformly distributed on N^2, it follows that π_2 is independent of (I_1, J_1). This proves b). Using this, an identical argument as that above shows that π_3 is uniformly distributed.

Let $T_i = \sum_j a_{j\pi_i(j)}$, $i=1,2,3$; $\Delta T_i = T_{i+1} - T_i$, $i=1,2$. ΔT_1 depends on $(\underline{I}, \underline{J})$, ΔT_2 on (I_1, I_2, J_1, J_2).

Therefore, if $f = f_{z,\lambda}$ from §2:

$$\begin{aligned} E(T_A f(T_A)) &= E(T_3 f(T_3)) = nE(a_{I_1 J_1} f(T_3)) \\ &= nE\left(a_{I_1 J_1} \Delta T_2 \int_0^1 (f'(T_1 + \Delta T_1 + t\Delta T_2) - f'(T_1))\, dt\right) \\ &\quad + nE(a_{I_1 J_1} \Delta T_2 f'(T_1)), \end{aligned}$$

where we used the independence of T_2 and $a_{I_1 J_1}$ (Lemma b). Using the independence of π_1 and (I_1, I_2, J_1, J_2) the second summand above equals $nE(a_{I_1 J_1} \Delta T_2) E(f'(T_1))$. Using this and (2.5), one obtains

$$\begin{aligned} &|E(f'(T_A) - T_A f(T_A))| \\ &\leqq 2|nE(a_{I_1 J_1} \Delta T_2) - 1| + n\int_0^1 E\Bigg\{|a_{I_1 J_1} \Delta T_2(\Delta T_1 + t\Delta T_2)| \Bigg(1 + 2|T_1| \\ &\quad + \frac{1}{\lambda}\int_0^1 1_{[z, z+\lambda]}(T_1 + s\Delta T_1 + st\Delta T_2)\, ds\Bigg)\Bigg\}\, dt \\ &\leqq 2|nE(a_{I_1 J_1} \Delta T_2) - 1| + nE(|a_{I_1 J_1} \Delta T_2|(|\Delta T_1| + |\Delta T_2|) \\ &\quad + 2nE(|a_{I_1 J_1} \Delta T_2|(|\Delta T_1| + |\Delta T_2|)|T_1|) \\ &\quad + \frac{n}{\lambda} E\Bigg(|a_{I_1 J_1} \Delta T_2|(|\Delta T_1 + \Delta T_2|) \\ &\quad \cdot \int_0^1 \int_0^1 1_{[z, z+\lambda]}(T_1 + s\Delta T_1 + st\Delta T_2)\, ds\, dt\Bigg) \\ &= A_1 + A_2 + A_3 + A_4, \quad \text{say.} \end{aligned} \tag{3.8}$$

A_1 is 0 and A_2 is easy to estimate:

$$E(|a_{I_1 J_1} \Delta T_2|(|\Delta T_1| + |\Delta T_2|))$$

contains summands of the form

$$E(|a_{I_1 J_1} a_{\alpha\beta} a_{\mu\nu}|)$$

where

$$\alpha, \mu \in \{I_1, \ldots, I_4\}, \qquad \beta, \nu \in \{J_1, \ldots, J_4\}.$$

An easy calculation shows that these summands are bounded by $c\gamma/n^2$. So

$$A_2 \leqq c\gamma/n. \tag{3.9}$$

In order to estimate A_3 and A_4 we look at the conditional distribution of T_1 given $\underline{I}=\underline{i}$, $\underline{J}=\underline{j}$. T_1 depends only on π_1 and the conditional distribution of π_1 is easy to describe: π_1 takes any permutation φ which satisfies $\varphi(i_k)=j_k$, $1 \leqq k \leqq 4$, with equal probability. If B is the matrix which is obtained from A by cancelling the rows i_1, i_2, i_3, i_4 and the columns j_1, j_2, j_3, j_4, the T_1, conditioned on $\underline{I}=\underline{i}$ and $\underline{J}=\underline{j}$ has the same law as

$$\sum_{i \in \{i_1, \ldots, i_4\}} a_{i\pi_1(i)} + \sum_{j=1}^{n-l} b_{j\sigma(j)},$$

where l is the number of (distinct) elements in $\{i_1, \ldots, i_4\}$ and σ is uniformly distributed on the permutations of $\{1, \ldots, n-l\}$. As $|a_{ij}| \leqq 1$ for all i, j, one has

$$\begin{aligned} E(|T_1| \,|\, \underline{I}=\underline{i}, \underline{J}=\underline{j}) &\leqq 4 + E(|\sum_i b_{i\sigma(i)}|) \\ &\leqq 4 + (E(\sum_i b_{i\sigma(i)})^2)^{1/2} \end{aligned}$$

which is bounded, uniformly in $\underline{i}$ and $\underline{j}$. Therefore, A_3 can be estimated in the same way as A_2, leading to

$$A_3 \leqq c\gamma/n. \tag{3.10}$$

If $1 \leqq l \leqq 4$, let $M_n^l(\gamma)$ be the set of $(n-l)\times(n-l)$-matrices, which can be obtained from matrices in $M_n^0(\gamma)$ by cancelling l rows and l columns. Introducing

$$\alpha(\lambda, \gamma, n) = \sup\{\|P(T_1 \in [z, z+\lambda] \,|\, \underline{I}, \underline{J})\|_\infty : z \in \mathbb{R},\ A \in M_n^0(\gamma)\}$$

we have:

$$\alpha(\lambda, \gamma, n) \leqq \sup\left\{P\left(\sum_{i=1}^{n-l} b_{i\sigma(i)} \in [z, z+\lambda]\right) : z \in \mathbb{R},\ B \in M_n^l(\gamma),\ 1 \leqq l \leqq 4\right\}. \tag{3.11}$$

If $B \in M_n^l(\gamma)$, then $|b_{i.}|$, $|b_{.j}|$, $|b_{..}|$ are $\leqq c/n$. Using this, one obtains

$$\left|\sigma_B^2 - \frac{1}{n-l-1}\sum_{i,j} a_{ij}^2\right| \leqq \frac{1}{n-l-1}\sum{}' a_{ij}^2 + o(1),$$

where the $\sum'$ is the sum over the cancelled matrix elements. Furthermore, $\sum' a_{ij}^2 \leqq cn\varepsilon_0^{2/3}$ if $\beta_A \leqq \varepsilon_0 n$ and if $\varepsilon_0 \leqq \varepsilon_0'$ is small enough and n_0 sufficiently large, we have $|\sigma_B^2 - 1| \leqq 1/2$ and therefore $\sigma_B^2 \geqq 1/2$. Using this, one gets

$$\beta_B = \sum_{i,j} |\hat{b}_{ij}|^3 \leqq c_2 \beta_A. \tag{3.12}$$

Therefore

$$\sup_z P(\sum_i b_{i\sigma(i)} \in [z, z+\lambda]) \leqq \sup_z P(\sum_i \hat{b}_{i\sigma(i)} \in [z, z+2\lambda]),$$

which, if $c_2 \varepsilon_0 \leqq \varepsilon_0'$ according to (3.6), is

$$\begin{aligned}&\leqq 2\delta(c_1 \beta_B, n-l) + c\beta_B/n + \lambda/\sqrt{2\pi}\\ &\leqq 2\delta(c_1 c_2 \beta_A, n-l) + c_2 c\beta_A/n + \lambda/\sqrt{2\pi}.\end{aligned}$$

Using (3.11), one has

$$\alpha(\lambda, \gamma, n) \leqq 2 \max_{1 \leqq l \leqq 4} \delta(c_1 c_2 \beta_A, n-l) + c\left(\lambda + \frac{\beta_A}{n}\right).$$

Using this with the estimate (3.9) of A_2, one obtains

$$A_4 \leqq \frac{c\gamma}{n}\left(1 + \frac{1}{\lambda}\frac{\gamma}{n} + \frac{1}{\lambda} \max_{1 \leqq l \leqq 4} \delta(c_1 c_2 \gamma, n-l)\right).$$

Combining this with (3.9) and (3.10) in (3.8) and using (2.2) and (2.3) one obtains

$$\delta(\gamma, n) \leqq c_3 \frac{\gamma}{n}\left(1 + \frac{1}{\lambda}\frac{\gamma}{n} + \frac{1}{\lambda} \max_{1 \leqq l \leqq 4} \delta(c_1 c_2 \gamma, n-l)\right) + \frac{\lambda}{\sqrt{2\pi}}.$$

Now one may choose λ at ones liking, so we take $\lambda = 2c_1 c_2 c_3 \gamma/n$ leading to

$$\delta(\gamma, n) \leqq \frac{c\gamma}{n} + \frac{1}{2c_1 c_2} \delta(c_1 c_2 \gamma, n-l),$$

and if $n \geqq 8$, this gives

$$\sup_\gamma \frac{n\delta(\gamma, n)}{\gamma} \leqq c + \frac{1}{2} \max_{1 \leqq l \leqq 4} \sup_\gamma \frac{(n-l)}{\gamma} \delta(\gamma, n-l).$$

This proves $\delta(\gamma, n) \leqq c\gamma/n$ and, using again (3.6), this proves the theorem.

References

1. Does, R.J.M.M.: Berry-Esseen theorems for simple linear rank statistics. Ann. Probability **10**, 982–991 (1982)
2. Ho, S.T., Chen, L.H.Y.: An L_p bound for the remainder in a combinatorial central limit theorem. Ann. Probability **6**, 231–249 (1978)
3. Hoeffding, W.: A combinatorial central limit theorem. Ann. Math. Statist. **22**, 558–566 (1951)
4. Husková, Marie: The Berry-Esseen theorem for rank statistics. Comment. Math. Univ. Carolina, **20**, 399–415 (1979)
5. Motoo, M.: On the Hoeffding's combinatorial central limit theorem. Ann. Inst. Statist. Math. **8**, 145–154 (1957)
6. Stein, Ch.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. Proc. Sixth Berkeley Sympos. Math. Statist. Probability **2**, 583–602 (1972)
7. von Bahr, B.: Remainder term estimate in a combinatorial limit theorem. Z. Wahrscheinlichkeitstheorie verw. Gebiete **35**, 131–139 (1976)

Received November 15, 1982; in revised form October 5, 1983