SOME EXTENSIONS OF THE WALD-WOLFOWITZ-NOETHER THEOREM

By Jaroslav Hájek

Mathematical Institute of the Czechoslovak Academy of Sciences

1. Summary and introduction. Let $(R_{\nu 1}, \cdots, R_{\nu N_\nu})$ be a random vector which takes on the $N_\nu!$ permutations of $(1, \cdots, N_\nu)$ with equal probabilities. Let $\{b_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ and $\{a_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ be double sequences of real numbers. Put

$$S_\nu = \sum_{i=1}^{N_\nu} b_{\nu i} a_{\nu R_{\nu i}}. \tag{1.1}$$

We shall prove that the sufficient and necessary condition for asymptotic $(N_\nu \to \infty)$ normality of S_ν is of Lindeberg type. This result generalizes previous results by Wald-Wolfowitz [1], Noether [3], Hoeffding [4], Dwass [6], [7] and Motoo [8]. In respect to Motoo [8] we show, in fact, that his condition, applied to our case, is not only sufficient but also necessary.

Cases encountered in rank-test theory are studied in more detail in Section 6 by means of the theory of martingales. The method of this paper consists in proving asymptotic equivalency in the mean of (1.1) to a sum of infinitesimal independent components.

2. Three lemmas. Consider a sequence $U_1, \cdots, U_N$ of independent random variables each having uniform (rectangular) distribution over the interval $(0, 1]$. Let R_i be the rank of U_i, i.e.,

$$U_i = Z_{R_i}, \tag{2.1}$$

where $Z_1 < \cdots < Z_N$ is the sequence $U_1, \cdots, U_N$, reordered in ascending magnitude.

Take a nondecreasing sequence $a_1 \leqq \cdots \leqq a_N$ of real numbers and put

$$a(\lambda) = a_i \quad \text{for} \quad (i-1)/N < \lambda \leqq i/N \qquad (1 \leqq i \leqq N). \tag{2.2}$$

The function $a(\lambda)$ will be called a quantile function of $a_1 \leqq \cdots \leqq a_N$. As $(i-1)/N < i/(N+1) < i/N$, we have

$$a_i = a(i/N) = a[i/(N+1)]. \tag{2.3}$$

Furthermore,

$$\bar{a} = \frac{1}{N}\sum_{i=1}^{N} a_i = \int_0^1 a(\lambda)\, d\lambda \quad \text{and} \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(a_i - \bar{a})^2 = \int_0^1 [a(\lambda) - \bar{a}]^2\, d\lambda. \tag{2.4}$$

Received January 12, 1960; revised August 13, 1960.

LEMMA 2.1.

$$E\left[a(U_1) - a\left(\frac{R_1}{N}\right)\right]^2 \leqq 2 \max_{1 \leqq i \leqq N} |a_i - \bar{a}| \frac{1}{N}\left[2 \sum_{i=1}^{N} (a_i - \bar{a})^2\right]^{\frac{1}{2}}, \tag{2.5}$$

where the function $a(\cdot)$ *is given by* (2.2).

PROOF. If $Z_1, \cdots, Z_N$ are fixed, then U_1 takes on each of the values Z_i with probability $1/N$, and $U_1 = Z_i$ is equivalent to $R_1 = i$. Therefore,

$$\begin{aligned} E[a(U_1) - a(R_1/N)]^2 &= EE\{[a(U_1) - a(R_1/N)]^2 \mid Z_1, \cdots, Z_N\} \\ &= (1/N)E\sum_{i=1}^{N} [a(Z_i) - a(i/N)]^2. \end{aligned} \tag{2.6}$$

Now, first, consider a special quantile function

$$\begin{aligned} \epsilon[\lambda - (k/N)] &= 0 \quad \text{if} \quad \lambda \leqq k/N \\ &= 1 \quad \text{if} \quad \lambda > k/N. \end{aligned} \tag{2.7}$$

The quantile function $\epsilon[\lambda - (k/N)]$ corresponds, obviously, to the sequence $a_1 = \cdots = a_k = 0$, $a_{k+1} = \cdots = a_N = 1$. Let K denote the number of the U_i's smaller than k/N. Clearly, $Z_K < k/N < Z_{K+1}$. If $K \leqq k$, we have

$$\begin{aligned} \epsilon[Z_i - (k/N)] - \epsilon[(i-k)/N] &= 0 \quad \text{if } i = 1, \cdots, K, k+1, \cdots, N \\ &= 1 \quad \text{otherwise} \end{aligned} \tag{2.8}$$

so that

$$\sum_{i=1}^{N}\left[\epsilon\left(Z_i - \frac{k}{N}\right) - \epsilon\left(\frac{i-k}{N}\right)\right]^2 = |K - k|. \tag{2.9}$$

We can easily see that (2.9) also holds for $K \geqq k$. The result (2.9) together with (2.6) gives

$$E\left[\epsilon\left(U_1 - \frac{k}{N}\right) - \epsilon\left(\frac{R_1 - k}{N}\right)\right]^2 = \frac{1}{N} E\,|K - k|. \tag{2.10}$$

The distribution of K is, obviously, binomial with mean value k and variance $k[1 - (k/N)]$, so that

$$E|K - k| \leqq [E(K-k)^2]^{\frac{1}{2}} = [k\{1 - (k/N)\}]^{\frac{1}{2}},$$

and, therefore,

$$E\{\epsilon(U_1 - k/N) - \epsilon[(R_1 - k)/N]\}^2 \leqq N^{-1}\{k[1 - (k/N)]\}^{\frac{1}{2}}. \tag{2.11}$$

Now let us only suppose that $a_1 = 0$, and otherwise the sequence $a_1 \leqq \cdots \leqq a_N$ can be arbitrary. The quantile function of any such sequence may be expressed in the form

$$a(\lambda) = \sum_{k=1}^{N-1} (a_{k+1} - a_k)\, \epsilon[\lambda - (k/N)] \qquad (a_1 = 0, 0 < \lambda \leqq 1). \tag{2.12}$$

Actually, e.g., for $\lambda = i/N$, we have

$$a(i/N) = \sum_{k=1}^{N-1} (a_{k+1} - a_k)\, \epsilon\, [(i-k)/N] = \sum_{k=1}^{i-1} (a_{k+1} - a_k) = a_i \qquad (1 \leqq i \leqq N).$$

Now from (2.12) it follows that, first,

$$\begin{aligned} \sum_{i=1}^{N} a_i^2 &= \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j) \sum_{i=1}^{N} \epsilon\, [(i-k)/N]\, \epsilon\, [(i-j)/N] \\ &= \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j)(N - \max(k,j)), \end{aligned} \tag{2.13}$$

and, second,

$$\begin{aligned} \left[a(Z_i) - a\left(\frac{i}{N}\right)\right]^2 &= \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j) \\ &\cdot \left[\epsilon\left(Z_i - \frac{k}{N}\right) - \epsilon\left(\frac{i-k}{N}\right)\right]\left[\epsilon\left(Z_i - \frac{j}{N}\right) - \epsilon\left(\frac{i-j}{N}\right)\right]. \end{aligned} \tag{2.14}$$

Because $\epsilon[Z_i - (k/N)] - \epsilon[(i-k)/N]$ and $\epsilon[Z_i - (j/N)] - \epsilon[(i-j)/N]$ take on only the values 0 and ± 1, we have

$$\begin{aligned} &\left[\epsilon\left(Z_i - \frac{k}{N}\right) - \epsilon\left(\frac{i-k}{N}\right)\right]\left[\epsilon\left(Z_i - \frac{j}{N}\right) - \epsilon\left(\frac{i-j}{N}\right)\right] \\ &\qquad \leqq \left[\epsilon\left(Z_i - \frac{\max(k,j)}{N}\right) - \epsilon\left(\frac{i - \max(k,j)}{N}\right)\right]^2. \end{aligned} \tag{2.15}$$

On combining (2.14) (2.15) and (2.11), we get

$$\begin{aligned} E\left[a(U_1) - a\left(\frac{R_1}{N}\right)\right]^2 &= \frac{1}{N} E \sum_{i=1}^{N} \left[a(Z_i) - a\left(\frac{i}{N}\right)\right]^2 \\ &\leqq \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j) \\ &\qquad \cdot \frac{1}{N} E \sum_{i=1}^{N} \left[\epsilon\left(Z_i - \frac{\max(k,j)}{N}\right) - \epsilon\left(\frac{i - \max(k,j)}{N}\right)\right]^2 \\ &= \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j) E\left[\epsilon\left(U_1 - \frac{\max(k,j)}{N}\right) - \epsilon\left(\frac{R_1 - \max(k,j)}{N}\right)\right]^2 \\ &\leqq \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j) \frac{1}{N}\left[\max(k,j)\, \frac{N - \max(k,j)}{N}\right]^{\frac{1}{2}} \\ &\leqq \frac{1}{N} \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} (a_{k+1} - a_k)(a_{j+1} - a_j)[N - \max(k,j)]^{\frac{1}{2}} \end{aligned}$$

$$\leqq \frac{1}{N}\left\{\sum_{k=1}^{N-1}\sum_{j=1}^{N-1}(a_{k+1}-a_k)(a_{j+1}-a_j)\right.$$

$$\left.\cdot\sum_{k=1}^{N-1}\sum_{j=1}^{N-1}(a_{k+1}-a_k)(a_{j+1}-a_j)(N-\max(k,j))\right\}^{\frac{1}{2}}.$$

According to (2.13) and to

$$\sum_{k=1}^{N-1}\sum_{j=1}^{N-1}(a_{k+1}-a_k)(a_{j+1}-a_j) = a_N^2$$

the last expression equals $N^{-1}a_N(\sum_{i=1}^N a_i^2)^{\frac{1}{2}}$. This means that, for $a_1 = 0$,

$$E\left[a(U_1) - a\left(\frac{R_1}{N}\right)\right]^2 \leqq \frac{1}{N}a_N\left[\sum_{i=1}^N a_i^2\right]^{\frac{1}{2}}. \tag{2.16}$$

Generally, for an arbitrary a_1, we have

$$E\left[a(U_1) - a\left(\frac{R_1}{N}\right)\right]^2 \leqq \frac{1}{N}(a_N - a_1)\left[\sum_{i=1}^N (a_i - a_1)^2\right]^{\frac{1}{2}}. \tag{2.17}$$

On making use of the values $-a_N \leqq \cdots \leqq -a_1$ instead of $a_1 \leqq \cdots \leqq a_N$, we also get from (2.17) that

$$E\left[a(U_1) - a\left(\frac{R_1}{N}\right)\right]^2 \leqq \frac{1}{N}(a_N - a_1)\left[\sum_{i=1}^N (a_i - a_N)^2\right]^{\frac{1}{2}}. \tag{2.18}$$

It is now easy to derive (2.4). Let us put

$$\begin{aligned} a^+(\lambda) &= \bar{a} && \text{if } a(\lambda) \leqq \bar{a} \\ &= a(\lambda) && \text{if } a(\lambda) \geqq \bar{a} \end{aligned}$$

and

$$\begin{aligned} a^-(\lambda) &= a(\lambda) - \bar{a} && \text{if } a(\lambda) \leqq \bar{a} \\ &= 0 && \text{if } a(\lambda) \geqq \bar{a}. \end{aligned}$$

We then have

$$a(\lambda) = a^+(\lambda) + a^-(\lambda).$$

and, in view of (2.17) and (2.18), the following inequalities are clear:

$$\begin{aligned} E\left[a(U_1) - a\left(\frac{R_1}{N}\right)\right]^2 &\leqq 2E\left[a^+(U_1) - a^+\left(\frac{R_1}{N}\right)\right]^2 + 2E\left[a^-(U_1) - a^-\left(\frac{R_1}{N}\right)\right]^2 \\ &\leqq 2(a_N - \bar{a})[\sum_{a_i \geqq \bar{a}}(a_i - \bar{a})^2]^{\frac{1}{2}} + 2(\bar{a} - a_1)[\sum_{a_i \leqq \bar{a}}(a_i - \bar{a})^2]^{\frac{1}{2}} \\ &\leqq 2\max_{1 \leqq i \leqq N}|a_i - \bar{a}|\,\{[\sum_{a_i \geqq \bar{a}}(a_i - \bar{a})^2]^{\frac{1}{2}} + [\sum_{a_i \leqq \bar{a}}(a_i - \bar{a})^2]^{\frac{1}{2}}\} \\ &\leqq 2\max_{1 \leqq i \leqq N}|a_i - \bar{a}|\left[2\sum_{i=1}^N(a_i - \bar{a})^2\right]^{\frac{1}{2}}. \end{aligned}$$

This completes the proof.

Let us have a nondecreasing quadratically integrable function $\varphi(\lambda)$, $0 < \lambda < 1$, and put

$$(2.19) \qquad \varphi_N(\lambda) = \varphi[i/(N+1)] \quad \text{if} \quad (i-1)/N < \lambda \leqq i/N.$$

LEMMA 2.2. *The functions* $\varphi_N^2(\lambda)$, $0 < \lambda < 1$ *are uniformly* $(N \geqq 1)$ *integrable and*

$$(2.20) \qquad \lim_{N\to\infty} \int_0^1 [\varphi_N(\lambda) - \varphi(\lambda)]^2 \, d\lambda = 0.$$

PROOF. Suppose that $\varphi(0) \geqq 0$ so that both $\varphi(\lambda)$ and $\varphi^2(\lambda)$ are non-decreasing. Let A be a subset of $[0, 1]$ and $\mu(A)$ its Lebesgue measure. Put

$$I_k = ((k-1)/N, 1/N)$$

and $J_k = (k/N - \mu(A \cap I_k), k/N)$ and note that $\varphi_N(\lambda) = \varphi(k/(N+1))$ for $\lambda \,\varepsilon\, I_k$ and $\varphi_N(\lambda) \leqq \varphi(\lambda)$ for $k/(N+1) \leqq \lambda < k/N$. It holds, obviously,

$$(2.21) \quad \int_{A\cap I_k} \varphi_N^2(\lambda)\, d\lambda = \mu(A \cap I_k)\varphi^2(k/(N+1)) \leqq (N+1)k^{-1} \int_{J_k} \varphi^2(\lambda)\, d\lambda$$

so that, on making use of the first right hand expression for $k \leqq 3N/4$ and of the second one for $k \geqq N/4$, we get

$$(2.22) \quad \int_A \varphi_N^2(\lambda)\, d\lambda = \sum_{k=1}^N \int_{A\cap I_k} \varphi_N^2(\lambda)\, d\lambda \leqq \varphi^2(3/4)\mu(A) + 4\int_B \varphi^2(\lambda)\, d\lambda$$

where $\mu(B) = \mu(A)$. Inequality (2.22) clearly proves uniform integrability of the functions $\{\varphi_N^2(\lambda)\}$. A general $\varphi(\lambda)$ may be written in the form $\varphi(\lambda) = \varphi_1(\lambda) - \varphi_2(1-\lambda)$, where $\varphi_1(0) \geqq 0$, $\varphi_2(0) \geqq 0$ and both $\varphi_1(\lambda)$ and $\varphi_2(\lambda)$ are non-decreasing. This completes the proof of uniform integrability.

In order to prove (2.20), let us observe that $\varphi_N(\lambda) \to \varphi(\lambda)$ on the set of continuity points of $\varphi(\lambda)$, which, however, has Lebesgue measure 1. Convergence of $\varphi_N(\lambda)$ to $\varphi(\lambda)$ almost everywhere, together with uniform integrability of the function $\varphi_N^2(\lambda)$, implies (2.20). The proof is completed.

LEMMA 2.3. *Let* $c_1, \cdots, c_N, d_1, \cdots, d_N$ *be arbitrary real numbers and put* $\bar{c} = N^{-1}\sum_{i=1}^N c_i$, $\bar{d} = N^{-1}\sum_{i=1}^N d_i$. *Then*

$$(2.23) \qquad \begin{aligned} \operatorname{var}\left(\sum_{i=1}^N c_i d_{R_i}\right) &= \frac{1}{N-1}\sum_{i=1}^N (c_i - \bar{c})^2 \sum_{i=1}^N (d_i - \bar{d})^2 \\ &\leqq \frac{1}{N-1}\sum_{i=1}^N (c_i - \bar{c})^2 \sum_{i=1}^N d_i^2 . \end{aligned}$$

The proof is immediate.

3. Asymptotic equivalency in the mean. Random variables S_ν and T_ν will be called asymptotically equivalent in the mean (symbolically, $S_\nu \sim T_\nu$)

$$(3.1) \qquad \lim_{\nu\to\infty} E(S_\nu - T_\nu)^2/(\operatorname{var} T_\nu) = 0.$$

The relation is symmetric, and, if $S_\nu \sim T_\nu$ and $T_\nu \sim V_\nu$, then, clearly, $S_\nu \sim V_\nu$.

Let us take a sequence of independent random variables $U_1, U_2, \cdots$ each uniformly distributed over (0, 1] and denote the rank of U_i in the partial sequence $U_1, \cdots, U_{N_\nu}$ by $R_{\nu i}$, $1 \leqq i \leqq N_\nu$, $\nu \geqq 1$. The partial sequence $U_1, \cdots, U_{N_\nu}$, reordered in ascending magnitude, will be denoted by $Z_{\nu 1}, \cdots, Z_{\nu N_\nu}$.

The distribution of S_ν given by (1.1) does not depend on the ordering of the a's. So we may suppose that

$$a_{\nu 1} \leqq \cdots \leqq a_{\nu N_\nu} \qquad (\nu \geqq 1). \tag{3.2}$$

We shall assume that the a's fulfill the condition

$$\lim_{\nu \to \infty} \frac{\max_{1 \leqq i \leqq N_\nu} (a_{\nu i} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} = 0. \tag{3.3}$$

THEOREM 3.1. *Under the assumptions* (3.2) *and* (3.3) *the statistic S_ν given by* (1.1) *is asymptotically equivalent in the mean to the statistic*

$$T_\nu = \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu) a_\nu(U_i) + \bar{b}_\nu \sum_{i=1}^{N_\nu} a_{\nu i}, \tag{3.4}$$

where $a_\nu(\cdot)$ denotes the quantile functions of $a_{\nu 1} \leqq \cdots \leqq a_{\nu N_\nu}$ given by (2.2), *and $\bar{b}_\nu = N_\nu^{-1} \sum_{i=1}^{N_\nu} b_{\nu i}$.*

PROOF. On making use of (2.1) and (2.3), we may write

$$S_\nu - T_\nu = \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu) \left[a_\nu(Z_{\nu R_{\nu i}}) - a_\nu\left(\frac{R_{\nu i}}{N}\right)\right]. \tag{3.5}$$

As is well-known, the distribution of the ranks $(R_{\nu 1}, \cdots, R_{\nu N_\nu})$ is independent of the vector $(Z_{\nu 1}, \cdots, Z_{\nu N_\nu})$. In view of Lemma 2.3, where we put

$$c_i = b_{\nu i} - \bar{b}_\nu$$

and $d_i = a_\nu(Z_{\nu i}) - a_\nu(i/N)$, we can write

$$\begin{aligned} E\{(S_\nu - T_\nu)^2 \mid Z_{\nu 1}, \cdots, Z_{\nu N_\nu}\} &= \operatorname{var}\{S_\nu - T_\nu \mid Z_{\nu 1}, \cdots, Z_{\nu N_\nu}\} \\ &\leqq \frac{1}{N_\nu - 1} \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} \left[a_\nu(Z_{\nu R_{\nu i}}) - a_\nu\left(\frac{R_{\nu i}}{N_\nu}\right)\right]^2 \\ &= \frac{1}{N_\nu - 1} \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} \left[a_\nu(U_i) - a_\nu\left(\frac{R_{\nu i}}{N_\nu}\right)\right]^2. \end{aligned} \tag{3.6}$$

The first equality in (3.6) is ensured by $E(S_\nu - T_\nu \mid Z_1, \cdots, Z_{\nu N_\nu}) = 0$ which follows from $\sum (b_{\nu i} - \bar{b}_\nu) = 0$. Taking the mean value over $Z_{\nu 1}, \cdots, Z_{\nu N_\nu}$ on both sides of the inequality (3.6), we obtain

$$E(S_\nu - T_\nu)^2 \leqq \frac{1}{N_\nu - 1} \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} E\left[a_\nu(U_i) - a_\nu\left(\frac{R_{\nu i}}{N_\nu}\right)\right]^2. \tag{3.7}$$

Clearly,

$$E\left[a_\nu(U_i) - a_\nu\left(\frac{R_{\nu i}}{N_\nu}\right)\right]^2 = E\left[a_\nu(U_1) - a_\nu\left(\frac{R_{\nu 1}}{N_\nu}\right)\right]^2 \quad (1 \leqq i \leqq N_\nu, \nu \geqq 1), \tag{3.8}$$

so that (3.7) may be put in the form

$$(3.9) \qquad E\,(S_\nu - T_\nu)^2 \leqq \frac{N_\nu}{N_\nu - 1} E\left[a_\nu(U_1) - a_\nu\left(\frac{R_{\nu 1}}{N_\nu}\right)\right]^2 \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2.$$

On the other hand, it follows from (3.4) that

$$(3.10) \qquad \operatorname{var} T_\nu = 1/N_\nu \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2.$$

Making use of Lemma 2.1, (3.9) and (3.10) yield

$$(3.11) \qquad \frac{E(S_\nu - T_\nu)^2}{\operatorname{var} T_\nu} \leqq \frac{2\sqrt{2}\, N_\nu}{N_\nu - 1} \frac{\max |\, a_{\nu i} - \bar{a}_\nu \,|}{\left[\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2\right]^{\frac{1}{2}}}.$$

Consequently, the relation (3.1) follows from (3.11) and (3.3). The proof is completed.

THEOREM 3.2. *Let the function $\varphi(\lambda)$ be non-decreasing, non-constant and quadratically integrable. Suppose that*

$$(3.12) \qquad \lim_{\nu \to \infty} N_\nu = \infty.$$

Then the statistics

$$(3.13) \qquad S_\nu = \sum_{i=1}^{N_\nu} b_{\nu i} \varphi[R_{\nu i}/(N_\nu + 1)]$$

and

$$(3.14) \qquad T_\nu = \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)\varphi(U_i) + \bar{b}_\nu \sum_{i=1}^{N_\nu} \varphi[i/(N_\nu + 1)]$$

are asymptotically equivalent in the mean.

PROOF. If we put

$$(3.15) \qquad a_{\nu i} = \varphi[i/(N_\nu + 1)]\,,$$

the quantile function of $a_{\nu 1} \leqq \cdots \leqq a_{\nu N_\nu}$ will equal $\varphi_{N_\nu}(\lambda)$ expressed by (2.19). According to Lemma 2.2, the functions $\varphi^2_{N\nu}(\lambda)$ are uniformly integrable and hence

$$(3.16) \qquad \lim_{\nu \to \infty} N_\nu^{-1} \max_{1 \leqq i \leqq N_\nu} |a_{\nu i}|^2 = \lim_{\nu \to \infty} \max_{1 \leqq i \leqq N_\nu} \int_{(i-1)/N_\nu}^{i/N_\nu} \varphi^2_{N_\nu}(\lambda)\, d\lambda = 0.$$

On the other hand, from (2.20) and from non-constancy of $\varphi(\lambda)$, it follows that

$$(3.17) \qquad \lim_{\nu \to \infty} \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 = \int_0^1 \left[\varphi(\lambda) - \int_0^1 \varphi(x)\, dx\right]^2 d\lambda > 0.$$

Relations (3.16) and (3.17) imply that the sequences $a_{\nu 1} \leqq \cdots \leqq a_{\nu N_\nu}$ fulfill condition (3.3). Therefore, in view of Theorem 3.1, the statistic (3.13) is asymptotically equivalent to the statistic

$$T'_\nu = \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)\, \varphi_{N_\nu}(U_i) + \bar{b}_\nu \sum_{i=1}^{N_\nu} \varphi[i/(N_\nu + 1)]. \tag{3.18}$$

It remains to show that (3.18) is equivalent to (3.14). However, as is easy to see,

$$\frac{E(T_\nu - T'_\nu)^2}{\operatorname{var} T_\nu} = \frac{\int_0^1 [\varphi(\lambda) - \varphi_{N_\nu}(\lambda)]^2\, d\lambda}{\int_0^1 \left[\varphi(\lambda) - \int_0^1 \varphi(x)\, dx\right]^2 d\lambda}, \tag{3.19}$$

so that $T_\nu \sim T'_\nu$ is a consequence of the assumption (3.12) and of Lemma 2.2. Theorem 3.2 is proved.

4. Necessary and sufficient condition for asymptotic normality of S_ν. If $S_\nu \sim T_\nu$, then obviously, the asymptotic variance and the asymptotic distribution of S_ν and T_ν exist under the same conditions, and, if they exist, are the same. Thus the problem of the asymptotical distribution of S_ν is reduced to the problem of the asymptotical distribution of T_ν. The statistic T_ν, however, is a sum of independent addends, so that, if these addends are infinitesimal, it suffices to use well-known theory [11].

THEOREM 4.1. *Let us suppose that*

$$\lim_{\nu\to\infty} \frac{\max_{1 \leq i \leq N_\nu} (a_{\nu i} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} = 0 \tag{4.1}$$

and

$$\lim_{\nu\to\infty} \frac{\max_{1 \leq i \leq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2} = 0. \tag{4.2}$$

Then the statistic (1.1) *has an asymptotically normal distribution with mean value* ES_ν *and variance* var S_ν *if, and only if, for any* $\tau > 0$

$$\lim_{\nu\to\infty} 1/N_\nu \sum_{|\delta_{\nu ij}|>\tau}\sum \delta_{\nu ij}^2 = 0, \tag{4.3}$$

where

$$\delta_{\nu ij} = \frac{(b_{\nu i} - b_\nu)\,(a_{\nu j} - \bar{a}_\nu)}{\left[\frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2\right]^{\frac{1}{2}}} \quad (1 \leq i, j \leq N_\nu, \nu \geq 1). \tag{4.4}$$

PROOF. Assuming that $a_{\nu 1} \leq \cdots \leq a_{\nu N_\nu}$, then from (4.1) and Theorem 3.1 it follows that (1.1) is asymptotically equivalent in the mean to (3.4). Therefore, it suffices to show that, under the additional assumption (4.2), the condition (4.3) is necessary and sufficient for the asymptotic normality of T_ν with mean value ET_ν and variance var T_ν. The assumption (4.2), however, implies that the addends in (3.4) are infinitesimal, because

$$\frac{\max_{1 \leq i \leq N_\nu} \operatorname{var}[(b_{\nu i} - \bar{b}_\nu) a_\nu(U_i)]}{\operatorname{var} T_\nu} \tag{4.5}$$
$$= \frac{\max_{1 \leq i \leq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - a_\nu)^2} = \frac{\max_{1 \leq i \leq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}.$$

Consequently, we have to prove that (4.3) coincides with the Lindeberg condition for T_ν, namely with

$$\lim_{\nu \to \infty} \sum_{i=1}^{N_\nu} \frac{1}{\operatorname{var} T_\nu} \int_{|x| > \tau (\operatorname{var} T_\nu)^{\frac{1}{2}}} x^2 \, dP\{(b_{\nu i} - \bar{b}_\nu) a_\nu(U_i) < x\} = 0. \tag{4.6}$$

Clearly, we have

$$\frac{1}{\operatorname{var} T_\nu} \int_{|x| > (\operatorname{var} T_\nu)^{\frac{1}{2}}} x^2 \, dP\{(b_{\nu i} - \bar{b}_\nu) a_\nu(U_i) < x\} \tag{4.7}$$
$$= \frac{\frac{1}{N_\nu} \sum_{i=1}^{N_\nu} [(b_{\nu i} - \bar{b}_\nu)^2 \sum_{j \in E_{\nu i}} (a_{\nu j} - \bar{a}_\nu)^2]}{\frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2},$$

where

$$E_{\nu i} = \left\{ j : |(b_{\nu i} - b_\nu)(a_{\nu j} - \bar{a}_\nu)| > \tau \left[\frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \right]^{\frac{1}{2}} \right\}. \tag{4.8}$$

Now, observe that

$$\frac{1}{N_\nu} \sum\sum_{|\delta_{\nu ij}| > \tau} \delta_{\nu ij}^2 = \frac{\sum_{i=1}^{N_\nu} [(b_{\nu i} - b_\nu)^2 \sum_{j \in E_{\nu i}} (a_{\nu j} - \bar{b}_\nu)^2]}{\sum_{i=1}^{N_\nu} (b_{\nu i} - b_\nu)^2 \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2}, \tag{4.9}$$

so that (4.3) is actually equivalent to (4.6). The proof is accomplished.

The condition (4.3) is symmetrical in the a's and the b's. In applications, however, the a's and the b's often play a somewhat different role. In such cases the following theorem is useful:

THEOREM 4.2. *A double sequence* $\{a_{\nu i}, 1 \leq i \leq N_\nu, \nu \geq 1\}$ *satisfying the condition* (4.1) *fulfills the condition* (4.3) *for any double sequence* $\{b_{\nu i}, 1 \leq i \leq N_\nu, \nu \geq 1\}$ *satisfying the condition* (4.2) *if, and only if,*

$$\left[\lim_{\nu \to \infty} \frac{k_\nu}{N_\nu} = 0\right] \Rightarrow \lim_{\nu \to \infty} \frac{\max_{1 \leq i_1 < \cdots < i_{k_\nu} \leq N_\nu} \sum_{\alpha=1}^{k_\nu} (a_{\nu i_\alpha} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} = 0. \tag{4.10}$$

PROOF. First, let us prove that (4.2) and (4.10) implies (4.3). Put, for a fixed $\tau > 0$,

$$E_\nu = \left\{ j : (a_{\nu j} - \bar{a}_\nu)^2 > \frac{\tau^2 \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}{\max_{1 \leqq i \leqq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2} \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \right\} \tag{4.11}$$

and denote the number of elements in E_ν by k_ν. Clearly

$$\frac{\tau^2 \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}{\max_{1 \leqq i \leqq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2} \frac{k_\nu}{N_\nu} \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2 < \sum_{j \varepsilon E_\nu} (a_{\nu j} - \bar{a}_\nu)^2 \leqq \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2,$$

i.e.,

$$\frac{k_\nu}{N_\nu} \leqq \tau^{-2} \frac{\max_{1 \leqq i \leqq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2},$$

from which, in view of (4.2), it follows that

$$\lim_{\nu \to \infty} (k_\nu / N_\nu) = 0. \tag{4.12}$$

Relation (4.12), according to (4.10), implies that

$$\lim_{\nu \to \infty} \frac{\sum_{j \varepsilon E_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2} = 0. \tag{4.13}$$

Now, from (4.8) and (4.11) it follows that $E_{\nu i} \subset E_\nu$ and, consequently,

$$\frac{\sum_{i=1}^{N_\nu} \left[(b_{\nu i} - \bar{b}_\nu)^2 \sum_{j \varepsilon E_{\nu i}} (a_{\nu j} - \bar{a}_\nu)^2 \right]}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2} \leqq \frac{\sum_{j \varepsilon E_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2}. \tag{4.14}$$

Finally, (4.3) is an obvious consequence of (4.9), (4.14) and (4.13).

Second, let us assume that (2.10) does not hold. Then there exists a sequence B_ν of sets of integers such that, first, $B_\nu \subset \{1, \cdots, N_\nu\}$, second, the numbers of elements in B_ν, say l_ν, satisfy the relation

$$\lim_{\nu \to \infty} (l_\nu / N_\nu) = 0, \tag{4.15}$$

and, third,

$$\limsup_{\nu \to \infty} \frac{\sum_{j \varepsilon B_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2} > 0. \tag{4.16}$$

If

$$C_\nu = \left\{ j : (a_{\nu j} - \bar{a}_\nu)^2 > \frac{(N_\nu/l_\nu)^{\frac{1}{2}}}{N_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \right\}, \tag{4.17}$$

and $C_\nu^* = \{1, 2, \cdots, N_\nu\} - C_\nu$ then, clearly,

$$\frac{\sum_{j \varepsilon B_\nu \cap C_\nu^*} (a_{\nu j} - a_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} \leqq \left(\frac{N_\nu}{l_\nu}\right)^{\frac{1}{2}} \frac{l_\nu}{N_\nu} = \left(\frac{l_\nu}{N_\nu}\right)^{\frac{1}{2}}$$

and, therefore, according to (4.15),

$$\lim_{\nu \to \infty} \frac{\sum_{j \varepsilon B_\nu \cap C_\nu^*} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} = 0.$$

Consequently, in view of (4.16),

$$\limsup_{\nu \to \infty} \frac{\sum_{j \varepsilon C_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} \geqq \limsup \frac{\sum_{j \varepsilon B_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} > 0. \tag{4.18}$$

Now, put

$$b_{\nu 1} = \cdots = b_{\nu n_\nu} = 1, \qquad b_{\nu n_\nu + 1} = \cdots = b_{\nu N_\nu} = 0, \tag{4.19}$$

where n_ν is determined by

$$n_\nu \leqq (N_\nu/l_\nu)^{\frac{1}{2}} < n_\nu + 1 \qquad (\nu \geqq 1). \tag{4.20}$$

We have, obviously,

$$\lim_{\nu \to \infty} (n_\nu/N_\nu) = 0 \tag{4.21}$$

and, in view of (4.15),

$$\lim_{\nu \to \infty} n_\nu = \infty \tag{4.22}$$

Furthermore,

$$\bar{b}_\nu = n_\nu/N_\nu, \tag{4.23}$$

$$\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 = [n_\nu(N_\nu - n_\nu)]/N_\nu, \tag{4.24}$$

and

$$\frac{\max_{1 \leqq i \leqq N_\nu} (b_{\nu i} - \bar{b}_\nu)^2}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2} = \frac{\max\left\{\left(1 - \frac{n_\nu}{N}\right)^2, \left(\frac{n_\nu}{N_\nu}\right)^2\right\}}{\frac{n_\nu(N_\nu - n_\nu)}{N_\nu}} \leqq \max\left\{\frac{1}{n_\nu}, \frac{1}{N_\nu - n_\nu}\right\}. \tag{4.25}$$

Consequently, the relations (4.21) and (4.22) ensure that the condition (4.2) is satisfied.

The sets $E_{\nu i}$, given generally by (4.8), will be now for $1 \leqq i \leqq n_\nu$ defined by

$$(4.26)\quad E_{\nu i} = E_{\nu 1} = \left\{ j:(a_{\nu j} - \bar{a}_\nu)^2 > \tau^2 \frac{n_\nu}{N_\nu - n_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \right\} (1 \leqq i \leqq n_\nu).$$

From (4.20) it follows that, for ν sufficiently large, the set C_ν given by (4.17) will be included in the set $E_{\nu 1}$ given by (4.26). Therefore

$$\frac{1}{N_\nu} \sum_{|\delta_{\nu ij}| > \tau} \sum \delta^2_{\nu ij} = \frac{\sum_{i=1}^{N_\nu} [(b_{\nu i} - \bar{b}_\nu)^2 \sum_{j \varepsilon E_{\nu i}} (a_{\nu j} - \bar{a}_\nu)^2]}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2}$$

$$(4.27)\quad \geqq \frac{\sum_{i=1}^{n_\nu} [(b_{\nu i} - \bar{b}_\nu)^2 \sum_{j \varepsilon E_{\nu i}} (a_{\nu j} - \bar{a}_\nu)^2]}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2} = \frac{\left(1 - \frac{n_\nu}{N_\nu}\right)^2 n_\nu \sum_{j \varepsilon E_{\nu 1}} (a_{\nu j} - \bar{a}_\nu)^2}{n_\nu \left(1 - \frac{n_\nu}{N_\nu}\right) \sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2}$$

$$\geqq \left(1 - \frac{n_\nu}{N_\nu}\right) \frac{\sum_{j \varepsilon C_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{j=1}^{N_\nu} (a_{\nu j} - \bar{a}_\nu)^2} \qquad (\nu \geqq \nu_0).$$

The relations (4.27), (4.18) and (4.21) imply that (4.3) cannot hold, and the theorem is thereby proved.

Let $G_\nu(x)$ denote the distribution function of the numbers $a_{\nu 1} \leqq \cdots \leqq a_{\nu N_\nu}$, i.e.,

$$(4.28)\qquad G_\nu(x) = \frac{\text{number of the } a\text{'s smaller or equal to } x}{N_\nu}.$$

LEMMA 4.1. *Condition* (4.10) *may be expressed in either of the following three forms:*

(i) *The functions* $[a_\nu(\lambda) - \bar{a}_\nu]^2 \{\int_0^1 [a_\nu(\lambda) - \bar{a}_\nu]^2 \, d\lambda\}^{-1}$ *are uniformly integrable.*

(ii)

$$(4.29)\qquad [\lim_{\nu \to \infty} \epsilon_\nu = 0] \Rightarrow \left[\lim_{\nu \to \infty} \frac{\left(\int_0^{\epsilon_\nu} + \int_{1-\epsilon_\nu}^1\right) [a_\nu(\lambda) - \bar{a}_\nu]^2 \, d\lambda}{\int_0^1 [a_\nu(\lambda) - \bar{a}_\nu]^2 \, d\lambda} = 0 \right].$$

(iii)

$$(4.30)\qquad [\lim_{\nu \to \infty} K_\nu = \infty] \Rightarrow \left[\lim_{\nu \to \infty} \frac{1}{\sigma_\nu^2} \int_{|x - \bar{a}_\nu| > K_\nu \sigma_\nu} (x - \bar{a}_\nu)^2 \, dG_\nu(x) = 0 \right],$$

where

$$(4.31)\quad \sigma_\nu^2 = \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 = \int_0^1 [a_\nu(\lambda) - \bar{a}_\nu]^2 \, d\lambda = \int_{-\infty}^{\infty} (x - \bar{a}_\nu)^2 \, dG_\nu(x).$$

PROOF. If $a_{\nu 1} \leqq \cdots \leqq a_{\nu N_\nu}$ then surely

$$(4.32)\quad \max_{1 \leqq i_1 < \cdots < i_{k_\nu} \leqq N_\nu} \sum_{\alpha=1}^{k_\nu} (a_{\nu i_\alpha} - \bar{a}_\nu)^2 = \sum_{i=1}^{k_{\nu 1}} (a_{\nu i} - \bar{a}_\nu)^2 + \sum_{i=N_\nu - k_{\nu 2}}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2,$$

where $k_{\nu 1} + k_{\nu 2} = k_\nu$. On the other hand, we have

$$(4.33)\quad \begin{aligned} &\sum_{i=1}^{k_{\nu 1}} (a_{\nu i} - \bar{a}_\nu)^2 + \sum_{i=N_\nu - k_{\nu 2}}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \\ &\qquad = N_\nu \left(\int_0^{k_{\nu 1}/N_\nu} + \int_{1-(k_{\nu 2}/N_\nu)}^{1} \right) [a_\nu(\lambda) - \bar{a}_\nu]^2 \, d\lambda, \end{aligned}$$

which proves the equivalency of (4.10) to (ii) and, of course, to (i) as well.

Now we shall prove the equivalency of (4.10) to (4.30). Clearly,

$$(4.34)\quad \frac{1}{\sigma_\nu^2} \int_{|x-\bar{a}_\nu| > K_\nu \sigma_\nu} (x - \bar{a}_\nu)^2 \, dG_\nu(x) = \frac{1}{\sigma_\nu^2} \frac{1}{N_\nu} \sum_{|a_{\nu i} - \bar{a}_\nu| > K_\nu \sigma_\nu} (a_{\nu i} - \bar{a})^2.$$

Denoting the number of a's such that $|a_{\nu i} - \bar{a}_\nu| > K_\nu \sigma_\nu$ by k_ν, we have the following form of Tchebychev's inequality:

$$(4.35)\qquad k_\nu K_\nu^2 \leqq 1/\sigma_\nu^2 \sum_{|a_{\nu i} - \bar{a}_\nu| > K_\nu \sigma_\nu} (a_{\nu i} - \bar{a}_\nu)^2 \leqq N_\nu .$$

Assume, first, that (4.10) holds. Then $K_\nu \to \infty$ implies, in view of (4.35), $k_\nu/N_\nu \to 0$, so that, according to (4.10), the right side of (4.34) tends to 0, and, consequently, also the left side of (4.34) tends to 0. Thus (4.10) implies (4.30).

Assume, second, that (4.10) does not hold. On repeating the respective part of the proof of Theorem 4.2, we get again the relation (4.18), which is equivalent to

$$(4.36)\qquad \limsup_{\nu \to \infty} \frac{1}{\sigma_\nu^2} \int_{|x-\bar{a}_\nu| > (N_\nu/l_\nu)^{\frac{1}{2}} \sigma_\nu} (x - \bar{a}_\nu)^2 \, dG_\nu(x) > 0.$$

This means that (4.30) is not satisfied for $K_\nu = (N_\nu/l_\nu)^{\frac{1}{2}}$. So the negation of (4.10) implies the negation of (4.30), i.e., (4.30) implies (4.10). Lemma 4.1 is proved.

COROLLARY. *The statistic* (3.13) *is asymptotically normal with mean value* ES_ν *and variance* var S_ν *for any double sequence* $\{b_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ *satisfying* (4.2).

PROOF. By Lemma 2.2 and Lemma 4.1 the numbers $a_{\nu i} = \varphi[i/(N_\nu + 1)]$ fulfill the condition (4.10). It suffices to apply Theorem 4.2.

EXAMPLE 4.1. If the b's are given by (4.19), then the statistic

$$(4.37)\qquad S_\nu = \sum_{i=1}^{N_\nu} b_{\nu i} a_{\nu R_{\nu i}} = \sum_{i=1}^{n_\nu} a_{\nu R_{\nu i}}$$

represents a sum of n_ν elements selected by simple random sampling from the population $\{a_{\nu 1}, \cdots, a_{\nu N_\nu}\}$. The condition (4.2) is fulfilled, according to (4.25), if, and only if,

$$\lim_{\nu\to\infty} n_\nu = \lim_{\nu\to\infty} (N_\nu - n_\nu) = \infty. \tag{4.38}$$

Hence, provided that (4.1) holds, the distribution of S_ν is asymptotically normal with mean value ES_ν and variance var S_ν for all n_ν satisfying (4.38) if, and only if, the a's satisfy the condition (4.10).

As we shall see in Section 5, (4.10) is fulfilled, for example, if the populations $\{a_{\nu 1}, \cdots, a_{\nu N_\nu}\}$ have uniformly bounded excesses. See also [9] and [10] for further results concerning sampling from a finite population.

5. Comparison of various conditions. First let us introduce the following

Notation. The condition (4.1), introduced by Noether [3] and simplified by Hoeffding [4], will be denoted by N.

The Lindeberg condition (4.3) will be denoted by L.

The condition (4.10) will be denoted by Q.

The Wald-Wolfowitz [1] condition

$$\frac{\dfrac{1}{N_\nu}\displaystyle\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^r}{\left[\dfrac{1}{N_\nu}\displaystyle\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2\right]^{r/2}} = O(1) \qquad (r = 3, 4, \cdots) \tag{5.1}$$

where $O(1)$ denotes uniform boundedness, will be denoted by W.

The Hoeffding [4] condition

$$\lim_{\nu\to\infty} N_\nu^{\frac{1}{2}r-1} \frac{\displaystyle\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^r \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^r}{\left[\displaystyle\sum_{i=1}^{N_\nu} (a_{\nu i} - a_\nu)^2 \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2\right]^{r/2}} = 0 \tag{5.2}$$

will be denoted by H.

Observe that the conditions L and H concern $\{a_{\nu i}, b_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ whereas the conditions N, Q, W are applied to each double sequence $\{a_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ and $\{b_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ separately. The fact that $\{b_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ satisfies N and $\{a_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ satisfies Q will be denoted by NQ, and the symbols NN, NW and WW will have similar interpretations.

Theorem 5.1. $WW \Rightarrow NW \Rightarrow H$, $NQ \Rightarrow L \Rightarrow NN$.

Proof. For $WW \Rightarrow NW \Rightarrow H$ see Hoeffding [4], and for $H \Rightarrow L$ Motoo [8]. $NQ \Rightarrow L$ follows from Theorem 4.2. Thus it remains to prove $NW \Rightarrow NQ$, i.e., $W \Rightarrow Q$, and $L \Rightarrow NN$.

$W \Rightarrow Q$. If we take $r = 4$ in (5.1) and use the quantile function form, we get

$$\frac{\displaystyle\int_0^1 [a_\nu(\lambda) - \bar{a}_\nu]^4 \, d\lambda}{\left(\displaystyle\int_0^1 [a_\nu(\lambda) - \bar{a}_\nu]^2 \, d\lambda\right)^2} = O(1) \tag{5.3}$$

As is well-known from a theorem due to Vallé-Poussin, (5.3) implies that the functions $[a_\nu(\lambda) - \bar{a}_\nu]^2/\int_0^1[a_\nu(\lambda) - \bar{a}_\nu]\, d\lambda$ are uniformly integrable, which is equivalent to Q (see Lemma 4.1).

$L \Rightarrow NN$. This fact follows from the inequality

$$\frac{\max_{1\le j\le N_\nu} (a_{\nu j} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} = \sum_{i=1}^{N_\nu} \frac{\max_{1\le j\le N_\nu} (b_{\nu i} - \bar{b}_\nu)^2(a_{\nu j} - \bar{a}_\nu)^2}{\sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2 \sum_{i=1}^{N_\nu} (a_{\nu i} - \bar{a}_\nu)^2} \tag{5.4}$$
$$= \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} \max_{1\le j\le N_\nu} \delta_{\nu ij}^2 \le \epsilon + \frac{1}{N_\nu} \sum\sum_{|\delta_{\nu ij}|>\epsilon} \delta_{\nu ij}^2 \qquad (\epsilon > 0),$$

where the a's can be replaced be the b's.

REMARK 5.1. The condition (5.3) means that the excesses are uniformly bounded. Denoting this condition by W_4, we have $W \Rightarrow W_4 \Rightarrow Q$.

In [7] Dwass considered the empirical distribution function $G_\nu(x)$ of the values $a_{\nu 1}, \cdots, a_{\nu N_\nu}$ and supposed that, first, $\lim_{\nu\to\infty} G_\nu(x) = G(x)$ at every continuity point of a distribution function $G(x)$ and, second,

$$\int x\, dG_\nu(x) = \int x\, dG(x) = 0 \tag{5.5}$$

$$\int x^2\, dG_\nu(x) = \int x^2\, dG(x) = 1. \tag{5.6}$$

These assumptions imply that

$$[\lim_{\nu\to\infty} K_\nu = \infty] \Rightarrow \left[\lim_{\nu\to\infty} \int_{|x|>K_\nu} x^2\, dG_\nu = 0\right].$$

Consequently, by Lemma 4.1, the a's fulfill the condition Q, so that the respective part of Dwass theorem [7] is contained in Theorem 4.2.

6. A special case encountered in rank order test theory. In rank order tests theory there are used locally most powerful tests based on statistics of the form

$$S_\nu = \sum_{i=1}^{N_\nu} b_{\nu i} E\{\varphi(U_i) \mid R_{\nu i}\} \tag{6.1}$$

where $E\{\cdot \mid R_{\nu i}\}$ denotes the conditional mean value under the condition that the rank of U_i among the observations $U_1, \cdots, U_{N_\nu}$ equals $R_{\nu i}$. Let us observe that (6.1) is a special case of (1.1) for

$$a_{\nu i} = E\{\varphi(U_1) \mid R_{\nu 1} = i\} = E\{\varphi(Z_{\nu i})\}, \tag{6.2}$$

where, in the middle expression, the index 1 might be replaced by any index $j = 1, \cdots, N_\nu$.

For simplicity we shall suppose that, as in previous sections, the U's have a uniform distribution. This causes no loss of generality, since arbitrarily distributed observations may be expressed as (non-decreasing) functions of uniformly distributed observations. We shall also assume that

$$\int_0^1 \varphi(\lambda)\, d\lambda = 0 \tag{6.3}$$

and that

$$\int_0^1 \varphi^2(\lambda)\, d\lambda < \infty. \tag{6.4}$$

The assumption (6.3), however, is not essential. The integers N_ν will be assumed to tend to ∞ monotonically, i.e., $N_\nu \leqq N_{\nu+1}$, $\nu \geqq 1$.

LEMMA 6.1. *Put*

$$Y_\nu = Y_\nu(R_{\nu 1}) = E\{\varphi(U_1) \mid R_{\nu 1}\} \tag{6.5}$$

and assume that (6.3) *and* (6.4) *hold. Then*

(i)

$$P\{\lim_{\nu\to\infty} Y_\nu = \varphi(U_1)\} = 1, \tag{6.6}$$

(ii) $\{Y_1, Y_2, \cdots, \varphi(U_1)\}$ is a martingale and $\{Y_1^2, Y_2^2, \cdots, \varphi^2(U_1)\}$ is a semimartingale,

(iii) the random variables $Y_\nu^2(\nu \geqq 1)$ are uniformly integrable,

(iv)

$$\lim_{\nu\to\infty} E \mid Y_\nu^2 - \varphi^2(U_1)| = 0 \tag{6.7}$$

and

$$\lim_{\nu\to\infty} EY_\nu^2 = E\varphi^2(U_1). \tag{6.8}$$

PROOF. The Borel fields $\mathfrak{F}_\nu$ generated by the random vectors $(R_{\nu 1}, \cdots, R_{\nu N_\nu})$ form an increasing sequence of Borel fields. Denote by $\mathfrak{F}_\infty$ the smallest Borel field containing $\bigcup_1^\infty \mathfrak{F}_\nu$. As is well-known, the conditional distribution of U_1 for given $R_{\nu 1} = j$ is the Beta distribution with $p = j$ and $q = N - j + 1$. Hence

$$E\left(U_1 - \frac{R_{\nu 1}}{N_{\nu+1}}\right)^2 = \frac{1}{N_\nu} \sum_{j=1}^{N_\nu} \frac{j(N_\nu - j + 1)}{(N_\nu + 1)^2(N_\nu + 2)} < \frac{1}{N_\nu}, \tag{6.9}$$

from which follows that U_1 is equivalent (with probability 1) to a random variable measurable with respect to $\mathfrak{F}_\infty$. Consequently, $\varphi(U_1)$ is also equivalent to a random variable measurable with respect to $\mathfrak{F}_\infty$.

Since the conditional distribution of U_1 for fixed $R_{\nu 1}$ is independent of $R_{\nu 2}, \cdots, R_{\nu N_\nu}$, we may also write

$$Y_\nu = E\{\varphi(U_1)|R_{\nu 1}, \cdots, R_{\nu N_\nu}\} = E\{\varphi(U_1)|\mathfrak{F}_\nu\}. \tag{6.10}$$

Now we can apply the theory of martingales. The assertions (i) through (iv) are consequences of Doob [13], Chap. VII, Theorem 4.4, §1 Example 1, Theorem

1.1, Theorem 3.3 (III), Theorem 4.1 s, respectively. (6.8) is a simple consequence of (6.7).

LEMMA 6.2. *The numbers $a_{\nu i}$ given by* (6.2) *satisfy the condition Q given by* (4.10).

PROOF. Observe that Y_ν takes on the values $a_{\nu i}$ with equal probabilities so that condition (4.10) coincides, in view of (6.3) and (6.8), with uniform integrability of the random variables Y_ν, $\nu \geqq 1$, which has been proved by Lemma 6.1.

THEOREM 6.1. *If $\{b_{\nu i}, 1 \leqq i \leqq N_\nu, \nu \geqq 1\}$ fulfill the condition N given by* (4.2) *and the function $\varphi(\lambda)$ is non-vanishing and quadratically integrable, then the statistic* (6.1) *has an asymptotically normal distribution with mean value ES_ν and variance* var S_ν.

PROOF. It suffices to apply Theorem 4.2 and Lemma 6.2.

THEOREM 6.2. *Under the assumptions* (6.3) *and* (6.4), *the statistic* (6.1) *is asymptotically equivalent in the mean to the statistic*

$$T_\nu = \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)\varphi(U_i). \tag{6.11}$$

PROOF. By the method used in proving (3.9), we can show that

$$E(S_\nu - T_\nu)^2 \leqq [N_\nu/(N_\nu - 1)]E[\varphi(U_1) - Y_\nu]^2 \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2. \tag{6.12}$$

In view of (6.5), (6.12) is equivalent to

$$E(S_\nu - T_\nu)^2 \leqq [N_\nu/(N_\nu - 1)][E\varphi^2(U_1) - EY_\nu^2] \sum_{i=1}^{N_\nu} (b_{\nu i} - \bar{b}_\nu)^2. \tag{6.13}$$

Now it only remains to divide both sides of (6.13) by var T_ν and to apply (6.8).

REMARK 6.1. Theorem 6.1 generalizes the Dwass theorem [6], and the equality (6.8) generalizes the respective part of Hoeffding's Theorem 2 [5]. The condition of convexity of $\varphi(\lambda)$ is removed, which proves the conjecture made by Dwass in [14], p. 358.

7. Vector considerations. We shall briefly touch the question of asymptotical m-dimensional normality of a vector $(S_\nu^1, \cdots, S_\nu^m)$ where

$$S_\nu^g = \sum_{i=1}^{N_\nu} b_{\nu i} a^g_{\nu R_{\nu i}} \qquad (g = 1, \cdots, m), \tag{7.1}$$

i.e., the b's are fixed and the a's depend on $g = 1, \cdots, m$.

THEOREM 7.1. *Suppose that for any constants $\lambda_1, \cdots, \lambda_m$*

$$\sum_{i=1}^{N_\nu} \left[\sum_{g=1}^{m} \lambda_g (a_{\nu i}^g - \bar{a}_\nu^g)\right]^2 \geqq \epsilon \max_{1 \leqq g \leqq m} \left[\lambda_g^2 \sum_{i=1}^{N_\nu} (a_{\nu i}^g - \bar{a}_\nu^g)^2\right] \tag{7.2}$$

where ϵ is positive and independent of $\nu \geqq 1$. Assume that the b's fulfill the condition N given by (4.2) *and the a^g's* $(g = 1, \cdots, m)$ *the condition Q given by* (4.10). *Then the vector $(S_\nu^1, \cdots, S_\nu^m)$ has an asymptotically normal distribution with mean values ES_ν^g and covariances* cov (S_ν^g, S_ν^h), $1 \leqq g,h \leqq N_\nu$, $\nu \geqq 1$.

PROOF. According to a theorem by Cramér [12], p. 105, it suffices to prove asymptotic normality for any linear combination

$$\sum_{g=1}^{m} \lambda_g S_\nu^g = \sum_{i=1}^{N_\nu} b_{\nu i} \left[\sum_{g=1}^{m} \lambda_g a_{\nu R_{\nu i}}^g \right].$$

This will be done if we show that the numbers $c_{\nu i} = \sum_{g=1}^{m} \lambda_g a_{\nu i}^g$ fulfill the condition (4.10) for any $\lambda_1, \cdots, \lambda_m$. However, in view of (7.2), we have

$$\frac{\sum_{\alpha=1}^{k_\nu} (c_{\nu i_\alpha} - \bar{c}_\nu)^2}{\sum_{i=1}^{N_\nu} (c_{\nu i} - \bar{c}_\nu)^2} = \frac{\sum_{\alpha=1}^{k_\nu} \left[\sum_{g=1}^{m} \lambda_g (a_{\nu i_\alpha}^g - \bar{a}_\nu^g) \right]^2}{\sum_{i=1}^{N_\nu} \left[\sum_{g=1}^{m} \lambda_g (a_{\nu i}^g - \bar{a}_\nu^g) \right]^2}$$

$$\leqq \frac{\sum_{\alpha=1}^{k_\nu} m \sum_{g=1}^{m} \lambda_g^2 (a_{\nu i_\alpha}^g - \bar{a}_\nu^g)^2}{\epsilon \max_{1 \leqq g \leqq m} \left[\lambda_g^2 \sum_{i=1}^{N_\nu} (a_{\nu i}^g - \bar{a}_\nu^g)^2 \right]} \leqq \frac{m}{\epsilon} \sum_{g=1}^{m} \frac{\sum_{\alpha=1}^{k_\nu} (a_{\nu i_\alpha}^g - \bar{a}_\nu^g)^2}{\sum_{i=1}^{N_\nu} (a_{\nu i}^g - \bar{a}_\nu^g)^2}.$$

Now it suffices to note that, according to our suppositions the values $a_{\nu i}^g$ fulfill the condition (4.10).

REMARK 7.1. The condition (7.2.) simply means that all multiple correlation coefficients are uniformly bounded from 1.

REFERENCES

[1] A. WALD AND J. WOLFOWITZ, "Statistical tests based on permutations of the observations," *Ann. Math. Stat.*, Vol. 15 (1944), pp. 358–372.

[2] W. G. MADOW, "On the limiting distributions of estimates based on samples from finite universe," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 535–545.

[3] G. E. NOETHER, "On a theorem by Wald and Wolfowitz," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 455–558.

[4] W. HOEFFDING, "A combinatorial central limit theorem," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 169–566.

[5] W. HOEFFDING, "On the distribution of the expected values of the order statistics," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 93–100.

[6] M. DWASS, "On the asymptotic normality of certain rank order statistics," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 303–306.

[7] M. DWASS, "On the asymptotic normality of some statistics used in non-parametric tests," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 334–339.

[8] MINORU MOTOO, "On the Hoeffding's combinatorial central limit theorem," *Ann. Inst. Stat. Math.*, Vol. 8 (1957), pp. 145-154.

[9] PAUL ERDÖS AND ALFRÉD RÉNYI, "On a central limit theorem for samples from a finite population,"*Publ. Math. Inst. Hung. Acad. Sci.*, Vol. 4 (1959), pp. 49–61.

[10] J. HÁJEK, "Limiting distributions in simple random sampling from a finite population," *Publ. Math. Inst. Hung. Acad. Sci.*, Vol. 5 (1960), pp. 361–374.

[11] B. V. GNEDENKO AND A. N. KOLMOGOROV, *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Cambridge, 1954.

[12] HARALD CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.

[13] J. L. DOOB, *Stochastic Processes*, John Wiley & Sons, New York, 1953.

[14] M. DWASS, "The large sample power of rank order tests in the two-sample problem," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 352–374.