A CENTRAL LIMIT THEOREM FOR PARTLY DEPENDENT VARIABLES

By H. J. Godwin and S. K. Zaremba

University College of Swansea, Wales

1. Introduction and definitions. It is well-known that the Central Limit Theorem can be extended to cases in which the random variables under consideration are not entirely independent. In particular, various theorems have been produced with the purpose of dealing with variables which are dependent only when they are in some sense near to each other. The case of m-dependent variables (see, for instance, [13] and [16]) belongs to this category. Another case of this kind arises when the variables have several indices and are regarded as near to each other when they have at least one index value in common (for a somewhat special instance of this case see [9]). The importance of the latter case is due to the fact that it covers a large class of statistics for which W. Hoeffding [5] suggested the name of U-statistics; however, these statistics are only a special instance of it, as can be seen from the reduced number of degrees of freedom.

The purpose of the present paper is to prove a general form of the Central Limit Theorem for partly dependent variables. Its statement is believed to include, as special cases, all the hitherto published propositions on these lines, to cover most, if not all, the situations which have been treated ad hoc, and to go, in some directions, beyond the previously obtained results. As remarked by Feller [2], limiting distributions of normalized sums of random variables should not depend on the existence of moments; accordingly, no moments are postulated, and indeed the most general form of the Central Limit Theorem for independent random variables [2] is contained in the theorem which follows. The statement of the latter may appear slightly cumbersome but it implies, as corollaries, a variety of simpler propositions which are given in Section 3; on the other hand, its proof, which is a generalization of the argument in [16], and does not reduce the general case to that of independent variables, remains conceptually as simple as it would be if the argument were confined to some of the special cases of partly dependent variables. In order to simplify the language, the whole argument is stated for one-dimensional variables, but there is no difficulty in extending it to multi-dimensional variables; a general expression for the mixed moments given, for instance, in [7] is useful in applying the multivariate form of the Second Limit Theorem (discussed, for instance, in [10], Section 7).

In order to avoid misunderstandings, it should be remembered that pairwise disjoint sets of random variables are called (mutually) *independent* if the joint probability distribution function of their union is the product of the joint probability distribution functions of the various sets. A set of random variables will be called *irreducible* if it cannot be decomposed into two (mutually) independent proper subsets. But the factorization of a joint probability distribution function

Received April 26, 1960; revised December 12, 1960.

applies also to all the corresponding marginal distributions. Hence, if the pairwise disjoint sets $S_1, \cdots, S_n$ of random variables are independent and if $S_1', \cdots, S_n'$ are subsets of $S_1, \cdots, S_n$ respectively, these subsets are also independent. It follows that the partition of any denumerable set of ranom variables into irreducible sets is unique.

The analysis of relations of dependence between random variables is complicated by the well-known fact (see, for instance, [1], Section 14.4) that pairwise independence does not imply independence in general.

In order to overcome this difficulty, it is proposed to describe as *linkedness* any symmetrical and reflexive relation between random variables of a given set which satisfies the condition that any two subsets are (mutually) independent whenever no variable of one subset is linked with any variable of the other.

If two variables are not independent then they must be regarded as linked and, at the other extreme, we can construct the relation trivially by making any two variables linked. It is usually most convenient to restrict the linkedness as far as possible; in some of the applications listed below (Section 3) it is, in fact, necessary to regard variables as linked only when they are correlated, but in the case of the method of paired comparisons a wider grouping has to be linked. In the case of m-dependent variables two variables can be regarded as linked when their indices differ by not more than m. One can also think of a family of random variables with several indices and with a relation of linkedness equivalent to the presence of a given number of common index values.

2. The Main Theorem.

THEOREM. *Let $\{x_k\}$, with k belonging to K, be a denumerable family of random variables with a well-defined relation of linkedness, and K_t, with t belonging to T, a family of finite subsets of K, S_t being the set of all the variables x_k for which k belongs to K_t. The precise nature of K and T does not need to be prescribed; we can take T to be a topological space with the point ∞ adjoined so that $t \to \infty$ is meaningful.*[1] *Assume the existence of a number d, a function $\gamma(m)$ defined for all integral values of m greater than 2, and a function $\theta(t)$ defined for all t in T, with the property that, for all m and t in T, $\gamma(m)\theta(t)^{m-d}$ is an upper bound for the number of sequences of elements of S_t having m terms, beginning with any two arbitrarily given linked terms and forming irreducible sets. Moreover, assume that, for a suitable family of positive numbers a_t corresponding to any t in T and for any positive η, the following four conditions are satisfied:*

$$\text{(i)} \qquad \sum_{k \text{ in } K_t} P[|x_k| > a_t] \to 0 \quad \text{as} \quad t \to \infty;$$

---

[1] In each of the examples given in Section 3, K can be regarded as a vector space, but imposing on K and T any restrictions beyond the conditions of the theorem would have no bearing whatsoever on its proof, and indeed could only help to obscure the gist of the argument.

(ii) $$a_t^{-1} \sum_{k \text{ in } K_t} \int_{\eta\theta(t)^{-1}a_t<|x|\leq a_t} |x|\, dF_k(x) \to 0 \quad \text{as} \quad t \to \infty,$$

where $F_k(x)$ *is the probability distribution function of* x_k ;

(iii) $$\lim_{t\to\infty} a_t^{-2} \sum_{(k,l)}^{(t)} \int_{|x|,|y|\leq\eta\theta(t)^{-1}a_t} [x - b_{t,k}^{(\eta)}][y - b_{t,l}^{(\eta)}]\, dF_{k,l}(x, y) = 1,$$

where $F_{k,l}(x, y)$ *is the joint probability distribution function of* x_k *and* x_l, $\sum_{(k,l)}^{(t)}$ *denotes a sum extended to all pairs of values of* k *and* l *belonging to* K_t *and corresponding to linked variables* x_k *and* x_l, *and*

$$b_{t,k}^{(\eta)} = \int_{|x|\leq\eta\theta(t)^{-1}a_t} x\, dF_k(x);$$

(iv) $$\theta(t)^{2-d}a_t^{-2} \sum_{(k,l)}^{(t)} \int_{|x|,|y|\leq\eta\theta(t)^{-1}a_t} |x - b_{t,k}^{(\eta)}||y - b_{t,l}^{(\eta)}|\, dF_{k,l}(x, y)$$

is bounded. Then, as $t \to \infty$, *the distribution of the random variable*

$$X_t = a_t^{-1} \sum_{k \text{ in } K_t} (x_k - b_{t,k}^{(\theta(t))})$$

tends to be normal with zero mean and unit variance.

PROOF. It is easy to see that, if the conditions of the theorem are satisfied, there exists a family of positive numbers ϵ_t corresponding to every t in T in such a way that $\epsilon_t \to 0$ as $t \to \infty$, and that the conditions (ii′), (iii′) and (iv′), obtained from (ii), (iii) and (iv) by substituting ϵ_t for η, are satisfied. Put

$$\begin{aligned} u_{t,k} &= x_k, & z_{t,k} &= y_{t,k} = 0 \quad \text{if} \quad |x_k| > a_t ;\\ z_{t,k} &= x_k, & u_{t,k} &= y_{t,k} = 0 \quad \text{if} \quad \epsilon_t\theta(t)^{-1}a_t < |x_k| \leq a_t ;\\ y_{t,k} &= x_k, & u_{t,k} &= z_{t,k} = 0 \quad \text{if} \quad |x_k| \leq \epsilon_t\theta(t)^{-1}a_t ;\\ Y_t &= a_t^{-1} \sum_{k \text{ in } K_t} (y_{t,k} - b_{t,k}^{(\theta(t))}). \end{aligned}$$

Thus

(1) $$X_t = Y_t + a_t^{-1} \sum_{k \text{ in } K_t} z_{t,k} + a_t^{-1} \sum_{k \text{ in } K_t} u_{t,k}.$$

But (i) entails

$$P\,[\sum_{k \text{ in } K_t} u_{t,k} \neq 0] \to 0 \quad \text{as} \quad t \to \infty,$$

and *a fortiori*

(2) $$\operatorname{p\,lim}_{t\to\infty} a_t^{-1} \sum_{k \text{ in } K_t} u_{t,k} = 0.$$

On the other hand, in the new notation, (ii′) becomes

$$a_t^{-1} \sum_{k \text{ in } K_t} E(|z_{t,k}|) \to 0 \quad \text{as} \quad t \to \infty.$$

A fortiori,

(3) $$a_t^{-1}E(|\sum_{k \text{ in } K_t} z_{t,k}|) \to 0 \quad \text{as} \quad t \to \infty.$$

Hence, by an application of the Bienaymé-Chebyshev inequality,

(4) $$\text{p} \lim_{t\to\infty} a_t^{-1} \sum_{k \text{ in } K_t} z_{t,k} = 0.$$

The boundedness of each of the variables y_t ensures the existence of all its moments. In particular, according to the definitions of $b_{t,k}^{(\eta)}$, $y_{t,k}$, $z_{t,k}$ and Y_t,

$$E(Y_t) = -a_t^{-1}E(\sum_{k \text{ in } K_t} z_{t,k}),$$

so that (3) implies

(5) $$E(Y_t) \to 0 \quad \text{as} \quad t \to \infty.$$

Remembering that, according to the definition of $y_{t,k}$, $E(y_{t,k}) = b_{t,k}^{(\epsilon_t)}$, put

(6) $$\xi_{t,k} = a_t^{-1}[y_{t,k} - E(y_{t,k})] = a_t^{-1}[y_{t,k} - b_{t,k}^{(\epsilon_t)}].$$

There is no difficulty in verifying that the mutual independence of subsets of S_t entails that of the corresponding subsets of $\{\xi_{t,k}\}$; hence, in particular, it is admissible to regard, for any k and l in K_t, $\xi_{t,k}$ and $\xi_{t,l}$ as linked if, and only if, x_k and x_l are linked. On the other hand, now

$$Y_t - E(Y_t) = \sum_{k \text{ in } K_t} \xi_{t,k},$$

and, consequently, for any positive integer m,

$$E(Y_t - E(Y_t))^m = \sum_{k(1) \text{ in } K_t} \cdots \sum_{k(m) \text{ in } K_t} E(\xi_{t,k(1)} \cdots \xi_{t,k(m)}).$$

However, the last summand can be decomposed into a product of expectations according to the unique decomposition of $\xi_{t,k(1)} \cdots \xi_{t,k(m)}$ into irreducible sets, and terms corresponding to the same factorization can be grouped together. If, in the corresponding partial sum, we neglect to omit the products of moments in which the sets of variables under the various expectation signs are not mutually independent, the summation variables under each expectation sign run independently of those under the other expectations and, in consequence, the partial sum can be factorized into a product of sums of moments.

In any of these products, sums of first moments are equal to 0, while owing to (iii′) sums of second moments tend to 1. A sum of moments of any order q exceeding 2 will be shown to tend to 0 as $t \to \infty$. Indeed, this will be seen to remain true even if all the moments are replaced by the corresponding absolute moments. In the first place it should be noted that under each expectation the first variable is necessarily linked with at least one of the others. Accordingly the terms of the sum can be grouped into $q - 1$ overlapping classes, and since the moments are absolute, it suffices to show that the sum of all the terms in any of these classes tends to 0. Without loss of generality, it can be assumed that the first two variables are linked. The absolute values of the moments can only increase if the last $q - 2$ variables are each replaced by $2\epsilon_t\theta(t)^{-1}$, which is an obvious upper bound for their values. According to the assumptions of the theorem, the number of terms corresponding to any combination of values of the first two summation indices does not exceed $\gamma(q)\theta(t)^{q-d}$ and, therefore, the sum is bounded by

$$(7) \qquad (2\epsilon_t)^{q-2}\gamma(q)\theta(t)^{2-d} \sum_{(k,l)} E(|\xi_{t,k}\xi_{t,l}|),$$

which tends to 0 according to (iv′) and to the definition of $\xi_{t,k}$.

There remains to show that the sum of all the terms we failed to omit tends to 0 as $t \to \infty$. It is easy to see that this sum is a linear combination, with coefficients independent of t, of expressions obtained from products of sums of the type considered above by a process which could be described as one of amalgamating factors. Two, or more, factors are thus amalgamated if they are replaced by one sum of products of former summands, the new sum being restricted to sets of index values giving rise to irreducible sets of random variables (in addition to the original restriction requiring the variables under each expectation to form an irreducible set). But such amalgamated factors involve at least four summation indices, and, clearly, their absolute values are bounded in the same way as the previously considered sums of absolute moments. Consequently, the amalgamated factors, and, therefore, also the corresponding products, as well as their linear combinations mentioned above, tend to 0 as $t \to \infty$.

Hence, apart from terms the sum of which tends to 0, $E(Y_t - E(Y_t))^m$ is a sum of products of sums of second moments. Since no such products can arise when m is odd,

$$(8) \qquad \lim_{t\to\infty} E(Y_t - E(Y_t))^m = 0 \quad \text{when } m \text{ is odd.}$$

If m is even, we are left with $m!/[(\frac{1}{2}m)!2^{\frac{1}{2}m}]$ products of sums of second moments, arising out of the same number of possible partitions of m variables into $\frac{1}{2}m$ irreducible pairs. Since each of these products tends to 1,

$$(9) \qquad \lim_{t\to\infty} E(Y_t - E(Y_t))^m = m!/[(\tfrac{1}{2}m)!\,2^{\frac{1}{2}m}] \quad \text{when } m \text{ is even.}$$

The limits, given by (8) and (9), of the central moments of Y_t are those of a normal distribution with zero mean and unit variance. Consequently, according to the Second Limit Theorem (see, for instance, [3]), this is the limiting distribution of $Y_t - E(Y_t)$. Finally, owing to a proposition given by Cramér ([1], Section 20.6) and as a consequence of (1), (5), (4) and (2), this is also the limiting distribution of X_t. Hence the proof is complete.

Remark I. If the moments of $a_t^{-1}\sum_{k \text{ in } K_t}(z_{t,k} + u_{t,k})$ of some even order m tend to 0 as $t \to \infty$ (which can easily be expressed in terms of the joint probability distribution functions of m variables $\{x_t\}$), then, owing to the Hölder inequality and to the results obtained while proving the main theorem, the moments of X_t up to the order m tend to the corresponding moments of the limiting distribution.

REMARK II. The argument above also yields a form of the weak law of large numbers for partly dependent variables. More precisely, if, in the statement of the theorem, the second condition is relaxed by requiring only that (ii) should hold for *some* positive η and if the last two conditions are replaced by the conditions that, for the same η,

$$\text{(v)} \qquad \lim_{t\to\infty} a_t^{-2} \sum_{(k,l)}^{(t)} \int_{|x|,|y|\leq\eta\theta(t)^{-1}a_t} [x - b_{t,k}^{(\eta)}][y - b_{t,l}^{(\eta)}]\, dF_{k,l}(x, y) = 0$$

then the conclusion is that

$$\text{p}\lim_{t\to\infty} X_t = 0.$$

Indeed, if in the definition of $y_{t,k}$ and $z_{t,k}$, ϵ_t is replaced by η, the arguments proving (2), (4) and (5) still apply, while (v) is equivalent to

$$\text{l.i.m.}_{t\to\infty} Y_t - E(Y_t) = 0;$$

the proposition follows immediately from the last relation in conjunction with (1), (2), (4) and (5).

3. Special cases and applications. An important particular case of the main theorem arises when $d = 2$ and $\theta(t)$ is an upper bound for the number of elements of S_t which are linked with any random variable belonging to this set; then we can take $\gamma(m) = (m - 1)!$ and the last condition of the theorem is simplified by the omission of the factor $\theta(t)^{2-d}$.

In one of the main applications, K is the set of all the sets of, say, r positive integers, and K_t is the subset of K determined by the requirement that all these integers should be less than or equal to t, two variables being linked when the two index sets have at least one element in common. More particularly, if the joint probability distribution of any (finite) number of variables depends only on which indices have the same value and not on that particular value or on the values of the other indices, if the variables have finite second moments, and if linked variables are correlated, then the conditions of the theorem are satisfied, provided that we put

$$a_t^2 = \text{var} \sum_{k \text{ in } K_t} x_k .$$

Indeed, then, $a_t = O(t^{r-\frac{1}{2}})$ and

$$P[|x_k| > a_t] < a_t^{-2} \int_{|x_k|>a_t} x^2\, dF_k(x) = o(t^{1-2r}).$$

Hence

$$\sum_{k \text{ in } K_t} P[|x_k| > a_t] = o(t^{1-r}),$$

which proves (i). Furthermore, $\theta(t) = t^r - (t - 1)^r$ so that $\theta(t)^{-1}a_t = O(t^{\frac{1}{2}})$. Thus

$$(10)\quad \begin{cases} \int_{\eta\theta(t)^{-1}a_t<|x|\leqq a_t} |x|\,dF_k(x) \leqq \int_{|x|>\eta\theta(t)^{-1}a_t} |x|\,dF_k(x) \\ \leqq \eta^{-1}\theta(t)a_t^{-1}\int_{|x|>\eta\theta(t)^{-1}a_t} x^2\,dF_k(x) = O(t^{-\frac{1}{2}}), \end{cases}$$

and consequently

$$(11)\qquad a_t^{-1}\sum_{k \text{ in } k_t}\int_{|x|>\eta\theta(t)^{-1}a_t} |x|\,dF_k(x) \to 0 \quad \text{as} \quad t\to\infty,$$

which shows that (ii) is also satisfied. As a consequence of (ii),

$$E(x_t) - b_{t,k}^{(\eta)} \to 0 \quad \text{as} \quad t\to\infty$$

and this, combined with (10), implies (iii). Finally, it is easy to see that the sum in (iv) is $O(t^{2r-1})$, ensuring that this condition is also satisfied.

Consequently, given a family of independent and identically distributed random variables $\{Xi\}$ $(i = 1, 2, \cdots)$ as well as a symmetrical function f of r arguments, such that $f(X_1, \cdots, X_r)$ should have a second moment and $f(X_1, X_2, \cdots, X_r)$ and $f(X_1, X_{r+1}, \cdots, X_{2r-1})$ should be correlated, the conditions of the theorem will be satisfied by making $f(X_{(1)}, \cdots, X_{i(r)}) = x_k$, with k denoting the set $\{i(1), \cdots, i(r)\}$. Apart from the special form given to the covariances in [5] and from the fact that the theorem is stated, there, in its multidimensional version, this is Hoeffding's limit theorem for independently distributed variables. It should be noted that the case when $f(X_1, X_2, \cdots, X_r)$ and $f(X_1, X_{r+1}, \cdots, X_{2r-1})$ are uncorrelated is a trivial case of Hoeffding's theorem since, owing to his choice of the normalizing factor, the variances tend to 0. On the other hand, in the statement above, the random variables $\{X_i\}$ can have any number of dimensions, and if several functions are simultaneously considered the statement applies to any linear combination of these functions, implying an asymptotically normal joint distribution of the functions themselves (see, for instance, Section 7 in [10]).

As pointed out by Hoeffding, the statements above can be applied to a whole range of statistics, such as Gini's mean difference, Gini's coefficient of concentration, functions of rank and of the signs of differences of random variables, difference-sign and rank correlations in samples, tests of independence, etc.

In general, when the random variables in question have a common upper bound, i.e., when $P[|x_k| > A] = 0$ for some A and all k in K, the first two conditions of the main theorem are automatically satisfied, provided that $\theta(t)^{-1}a_t \to \infty$ as $t\to\infty$. This applies in particular to the test function in Wilcoxon's test (see, for instance, [15]). Given two samples $x_1, \cdots, x_m$ and $y_1, \cdots, y_n$ of two random variables, the test function is

$$(12)\qquad U = \sum_{i=1}^{m}\sum_{k=1}^{n} z_{ik}$$

where

$$z_{ik} = \begin{cases} 1 & \text{if } x_i > y_k \\ 0 & \text{if } x_i \leqq y_k . \end{cases}$$

The scope of the test will not be discussed here (see, however, [4], [11], [14], [18]). The theorem proved by Hoeffding in [5] cannot be applied here because the parts played by the two different sets of variables are not symmetrical. However, under the assumption that the two samples arise from two identically distributed populations, the asymptotic normality of U was proved in [11] (where the distribution for small samples was obtained as well), and could be deduced from a theorem in [6]; without this assumption, but with the restriction that m/n be constant, it was proved in [8]. On the other hand, apart from the trivial case when the distributions do not overlap and so U is constant, the asymptotic normality of U follows from the main theorem of the present paper under any hypothesis about the distribution of the two random variables in question and without any requirement on m/n; t becomes a two-dimensional vector (t_1, t_2) tending to infinity when both $t_1 \to \infty$ and $t_2 \to \infty$, and K_t becomes the set of all the pairs (i, k) of positive integers such that $i \leqq t_1$, $k \leqq t_2$.

In some applications, the patterns of dependence of the random variables are fairly complicated, and it was in order to cover such cases that, in the statement of the main theorem, the constant d was allowed to take values other than 2. One such case arises in the treatment of the test function in the method of "paired comparisons" [12] for the investigation of the transitivity of preferences. The subject of the experiment is asked to choose between each pair formed with the entities $v_1, \cdots, v_t$. Write $P_{\{i,j,k\}} = 1$ if the preferences confined to v_i, v_j and v_k are not transitive, and $P_{\{i,j,k\}} = 0$ if they are. Under the null hypothesis that all the choices between pairs are independent and equally probable, Moran [12] proved the asymptotic normality of $\sum P_{\{i,j,k\}}$ by an argument *ad hoc*.

The same result can be obtained by a direct application of the main theorem of the present paper. It is easily seen that $P_{\{i,j,k\}}$ and $P_{\{i',j',k'\}}$ have to be regarded as linked whenever the two sets of index values have at least two elements in common, so that $\theta(t) = 3t - 8$. Nevertheless, two sets of variables are independent if there is only one link between them. Hence we can take $d = 3$. Then $\gamma(m)$ is the number of possible patterns of links ensuring the irreducibility of a set of m of the variables $P_{\{i,j,k\}}$ (allowing for repetitions). Furthermore, putting

$$a_t^2 = \operatorname{var} \sum_{i,j,k \leqq t} P_{\{i,j,k\}} = \sum_{i,j,k \leqq t} \operatorname{var} P_{\{i,j,k\}},$$

we obtain $a_t = O(t^{\frac{3}{2}})$, and consequently, $a_t \theta(t)^{-1} \to \infty$ as $t \to \infty$, which, in view of the boundedness of the distribution of the variables, causes the conditions (i), (ii) and (iii) to be satisfied automatically. The last condition of the theorem is easily seen to be also satisfied.

In the same way, the theorem could be applied to the distribution of the test function under alternative hypotheses respecting the pattern of dependence of the variables set by the null hypothesis.

If $\theta(t)$ can be taken independent of t, as, for instance, in the case of m-dependent variables, this simplifies the whole situation in a way which is quite different from the simplification due to the boundedness of the distributions in question. Conditions (i) and (ii) can, then, be combined into one condition: $\lim_{t\to\infty} \sum_{k \text{ in } K_t} P[|x_k| > \eta a_t] = 0$ for every positive η, $\theta(t)$ can be omitted in the statement of condition (iii), while, owing to the Schwarz inequality, condition (iv) can be replaced by the simpler condition that

$$a_t^{-2} \sum_{k \text{ in } K_t} \int_{|x| \leq a t} [x - b_{t,k}^{(\eta)}]^2 \, dF_k(x)$$

should be bounded. Thus, as a special case of the theorem proved in the preceding section, we find the most general statement of the Central Limit Theorem for m-dependent variables published so far (see [13] and [16]), and, therefore, as a still more special case, also the Central Limit Theorem for independent random variables under conditions which are not only sufficient, but necessary as well, at least when the normalized random variables are infinitesimal (see [2] and [17]).

REFERENCES

[1] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.

[2] Willy Feller, "Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung," *Math. Z.*, Vol. 40 (1935), pp. 521–559.

[3] M. Fréchet and J. Shohat, "A proof of the generalized Second Limit Theorem," *Trans. Amer. Math. Soc.*, Vol. 33 (1931), pp. 533–543.

[4] B. V. Gnedenko, "On the Wilcoxon test of comparing of two samples," (Russian, English summary), *Bull. Acad. Polon. Sci., Sér. Sci. Math. Astr. Phys.*, Vol. 6 (1958), 615–618.

[5] Wassily Hoeffding, "A class of statistics with asymptotically normal distributions," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 293–325.

[6] Wassily Hoeffding, "A combinatorial central limit theorem," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 558–566.

[7] J. Halcombe Laning, jr. and Richard H. Battin, *Random Processes in Automatic Control*, McGraw-Hill Book Co., New York, 1956.

[8] E. L. Lehmann, "Consistency and unbiasedness of certain nonparametric tests," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 165–179.

[9] Z. A. Lomnicki and S. K. Zaremba, "A further instance of the central limit theorem for dependent random variables," *Math. Z.*, Vol. 66 (1957), pp. 490–494.

[10] Z. A. Lomnicki and S. K. Zaremba, "On some moments and distributions occurring in the theory of linear stochastic processes, I," *Monatsh. Math.*, Vol. 61 (1957), pp. 318–358.

[11] H. B. Mann and D. R. Whitney, "On a test whether one of two random variables is stochastically larger than the other," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 50–60.

[12] P. A. P. Moran, "On the method of paired comparisons," *Biometrika*, Vol. 34 (1947), pp. 363–365.

[13] Steven Orey, "A central limit theorem for m-dependent random variables," *Duke Math. J.*, Vol. 25 (1948), pp. 543–546.

[14] D. VAN DANTZIG, "On the consistency and power of Wilcoxon's two sample test," *Nederl. Akad. Wetensch. Proc.*, Ser. A., Vol. 54 *Indag. Math.*. Vol. 13 (1951), pp. 1–8.

[15] B. L. VAN DER WAERDEN, *Mathematische Statistik*, Springer, Berlin, 1957.

[16] S. K. ZAREMBA, "Note on the central limit theorem," *Math. Z.*, Vol. 69 (1958). pp. 295–298.

[17] S. K. ZAREMBA, "On necessary conditions for the central limit theorem." *Math. Z.*, Vol. 70 (1958), pp. 281–282.

[18] S. K. ZAREMBA, "A generalization of Wilcoxon's test," *Monatsh. Math.*, Vol. 65 (1961).