The Annals of Statistics
1975, Vol. 3, No. 2, 528-531

CONVERGENCE OF THE REDUCED EMPIRICAL PROCESS FOR NON-I.I.D. RANDOM VECTORS

BY GEORG NEUHAUS

University of Freiburg

Any triangular array of row independent random vectors with continuous df's has a standard reduction to random vectors with values in the unit cube. The reduced empirical process belonging to the transformed random vectors is always relatively compact. Weak convergence to a (necessarily Gaussian) process holds iff the corresponding covariance kernel converges pointwise.

1. Notation and results. For every $n \geqq 1$ let $X_n^j = (X_{n1}^j, \cdots, X_{nk}^j), j = 1, \cdots, n$, be independent random k-vectors with continuous df's F_n^j. The df $F_n = (1/n) \sum_{j=1}^n F_n^j$ has marginal df's $F_{ni}, i = 1, \cdots, k$ (say) which define random vectors

$$U_n^j = (U_{n1}^j, \cdots, U_{nk}^j) = (F_{n1}(X_{n1}^j), \cdots, F_{nk}(X_{nk}^j)), \qquad j = 1, \cdots, n, \tag{1.1}$$

with df's

$$G_n^j = F_n^j \circ (F_{n1}^{-1}, \cdots, F_{nk}^{-1}), \qquad j = 1, \cdots, n, \tag{1.2}$$

where F_{ni}^{-1} is the left continuous inverse of $F_{ni}, i = 1, \cdots, k$.
The df $G_n = (1/n) \sum_{j=1}^n G_n^j$ has uniform marginals G_{ni} and consequently

$$|G_n(t) - G_n(s)| \leqq k|t - s| \qquad \forall\, t, s \in E_k = [0, 1]^k \quad \forall\, n \geqq 1, \tag{1.3}$$

where $|t| = \max\{|t_i| : i = 1, \cdots, k\}$, $t = (t_1, \cdots, t_k) \in R^k$.
Let G_n^e be the resulting empirical df of the $U_n^j, j = 1, \cdots, n$. Then we get the reduced empirical process of $X_n^j, j = 1, \cdots, n$,

$$X_n(t) = n^{\frac{1}{2}}(G_n^e(t) - G_n(t)), \qquad t \in E_k, \tag{1.4}$$

with corresponding covariance kernel

$$EX_n(t) \cdot X_n(s) = K_n(t, s) = G_n(s \wedge t) - \frac{1}{n} \sum_{j=1}^n G_n^j(t) \cdot G_n^j(s), \qquad t, s \in E_k, \tag{1.5}$$

where $s \wedge t = (s_1 \wedge t_1, \cdots, s_k \wedge t_k)$.
Since $X_n(\ )$ takes values in D_k, see e.g. Neuhaus (1971), we can formulate a

THEOREM. (i) *For any triangular array of row independent random vectors* X_n^j, $j = 1, \cdots, n$, $n \geqq 1$, *having continuous* df*'s*

(1.6) *the sequence* $(X_n(\cdot))_{n \geqq 1}$ *is relatively compact on* D_k *(furnished with the Skorohod-topology).*

Received May 1973; revised April 1974.
AMS 1970 *subject classifications.* Primary 60B10; Secondary 60G15.
Key words and phrases. Non. i.i.d. random-vectors, convergence of the empirical process.

(ii) *If for some K*

$$K_n(s, t) \to K(s, t) \qquad \forall\, s, t \in E_k\,, \tag{1.7}$$

then there exists a Gaussian process X with sample paths in C_k and covariance kernel K for which

$$X_n \Rightarrow X \quad \text{on} \quad D_k \quad (\textit{weak convergence})\,. \tag{1.8}$$

On the other hand, if the finite dimensional distributions of X_n converge to those of some process X, then X is Gaussian and (1.7) *holds with K the covariance kernel of X.*

2. Proof. We shall point out that the proof of Neuhaus (1971), pages 1293–1294, where the i.i.d. case is treated extends almost literally to the non-i.i.d. case considered here. The proof *loc. cit.* is a k-dimensional extension of the "classical" 1-dimensional proof as presented in Parthasarathy (1967). Since this proof is quick and almost self-contained it seems an appropriate one in handling the reduced empirical process.

While proving the 1-dimensional version of the above theorem Shorack (1973) introduced the linearly interpolated process S_n of X_n having $S_n(i/n) = X_n(i/n)$ for $i = 0, \cdots, n$ and used only the supremum metric; but we feel it is much simpler to introduce a Skorohod type metric and work directly with X_n. The crucial point of his proof is the not quite immediate inequality

$$E|S_n(t) - S_n(s)|^4 \leqq 144|t - s|^2 \qquad \forall\, s, t \in [0, 1]\,. \tag{1.9}$$

For X_n itself such an inequality does not hold in general; nevertheless, it holds almost:

For every $k \geqq 1$, $r \geqq 1$ there exists a constant $Q = Q(k, r)$ independent of $n \geqq 1$ with

$$E(X_n(t) - X_n(s))^{2r} \leqq Q \cdot |t - s|^r \qquad t, s \in E_k,\ |t - s| \geqq \frac{1}{n}\,. \tag{1.10}$$

For the 1-dimensional case (1.10) (with $r = 2$) is elementary and was used by Shorack (1973), too. But even in the general case $k \geqq 1$, (1.10) (with $r = k + 1$) and (1.3) are all that one really needs, since the proof of the statement

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\{\sup\{|X_n(t) - X_n(s)| : |t - s| \leqq \delta\} \geqq \varepsilon\}) = 0\,, \qquad \forall\, \varepsilon > 0\,, \tag{1.11}$$

as given in Neuhaus (1971), pages 1293–1294, works without any change if (4.1) resp. Lemma 5.2 *loc. cit.* is replaced by (1.3) resp. (1.10) above. The rest of the Theorem then follows immediately; see e.g. Shorack (1973).

Therefore, all that remains to do is verify (1.10) for general $k \geqq 1$, $r \geqq 1$. By the way, let us mention that in Sen (1970) (1.10) was claimed for $k = 1$ and every $r \geqq 1$ without the restriction $|t - s| \geqq (1/n)$; but for example in the 1-dimensional i.i.d. case (1.10) (with $r = 2$) becomes incorrect for $|t - s| \to 0$.

PROOF OF (1.10). For every family $X_1, \cdots, X_n$ of independent integrable rv's the rth moment

$$E(\sum_{i=1}^n X_i)^r = \sum_{i_1,\cdots,i_r=1}^n E(X_{i_1} \cdots X_{i_r})$$

can be rewritten as a finite sum of terms like

$$(1.12) \qquad \pm \prod_{\mu=1}^r \sum_{j=1}^n \prod_{\nu=1}^r (EX_j^{a_{\nu\mu}})^{b_{\nu\mu}}, \qquad a_{\nu\mu}, b_{\nu\mu} \in \{0, 1, \cdots, r\}, b_{\nu\mu} \geqq 1,$$

with $\sum_{\mu=1}^r \sum_{\nu=1}^r a_{\nu\mu} b_{\nu\mu} = r$. The number of terms (1.12) in the sum as well as the $a_{\nu\mu}$'s and $b_{\nu\mu}$'s are independent of n. In the case $r = 3$ we have for example

$$(1.13) \qquad E(\sum_{i=1}^n X_i)^3 = (\sum_{i=1}^n EX_i)^3 + 2\sum_{i=1}^n (EX_i)^3 + \sum_{i=1}^n EX_i^3 - 3\sum_{i=1}^n EX_i^2 EX_i - 3\sum_{i=1}^n (EX_i)^2 \sum_{i=1}^n EX_i + 3\sum_{i=1}^n EX_i^2 \sum_{i=1}^n EX_i .$$

If the X_i's have centered $B(1, p_i)$ distribution we get from (1.12), utilizing $EX_i = 0$ and $|EX_i^s| \leqq p_i$, $s \geqq 2$:

$$(1.14) \qquad E(\sum_{j=1}^n X_j)^{2r} \leqq \sum_{i=1}^r A_i(\sum_{j=1}^n p_j)^i \quad \text{with } A_i \geqq 0, A_i \text{ independent of } n .$$

For $s, t \in E_k$ (without loss of generality $s \leqq t$ componentwise) we define $X_j = 1_t(U_n^j) - 1_s(U_n^j) - p_j$, $p_j = G_n^j(t) - G_n^j(s)$. Then (1.10) follows from (1.14) and (1.3). □

REMARK. Koul (1970) used still another approach in proving relative compactness of the 1-dimensional "weighted" reduced empirical process. He employed the well-known fluctuation inequalities of Billingsley (1968) (Section 12). It is clear how to get k-dimensional versions of Koul's (1970) results by using the multidimensional fluctuation inequalities of Bickel and Wichura (1971). Since this approach is rather involved one would look for a straightforward proof of relative compactness of the empirical process. This was our aim in the proof above. On the other hand, adapting our approach to the weighted empirical process seems not quite immediate and would make the proof involved, too.

Acknowledgment. I am thankful to the referee for his suggestions improving the readability of this note.

REFERENCES

[1] BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1671.

[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] KOUL, H. L. (1970). Some convergence theorems for ranks and weighted empirical cumulatives. *Ann. Math. Statist.* **41** 1768–1773.

[4] NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

[5] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

[6] SEN, P. K. (1970). On the distribution of the one-sample rank order statistics. *Nonparametric Techniques in Statistical Inference*, (M. Puri ed.). Cambridge Univ. Press.

[7] SHORACK, G. R. (1973). Convergence of reduced empirical and quantile processes with application to functions of order statistics in the non-i.i.d. case. *Ann. Statist.* **1** 146–152.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
DER ALBERT-LUDWIGS-UNIVERSITÄT
78 FREIBURG IM BREISGAU
HERMANN-HERDER STRASSE 10
GERMANY