The Annals of Probability
1976, Vol. 4, No. 3, 474-479

A NOTE ON INVARIANCE PRINCIPLES FOR INDUCED ORDER STATISTICS[1]

BY PRANAB KUMAR SEN

University of North Carolina

Weak convergence of a sequence of two-dimensional time parameter stochastic processes constructed from partial sums of induced order statistics to a standard Brownian sheet process is established.

1. Introduction. Let $\{(X_i, Y_i), i \geqq 1\}$ be a sequence of independent and identically distributed random vectors (i.i.d.rv) with a bivariate distribution function (df) H, and let F and G_x be respectively the marginal df of X_1 and the conditional df of Y_1 given $X_1 = x$; F is assumed to be continuous so that ties among the X_i can be neglected in probability. For every n ($\geqq 1$), let $X_{n,1} < \cdots < X_{n,n}$ be the order statistics corresponding to $X_1, \cdots, X_n$, and, as in Bhattacharyya (1974), the *induced order statistics* $Y_{n1}, \cdots, Y_{nn}$ are defined by

$$Y_{nk} = Y_j \quad \text{if} \quad X_{n,k} = X_j \qquad \text{for } j, k = 1, \cdots, n . \tag{1.1}$$

Let $m(x) = E(Y_1 \mid X_1 = x)$, $\sigma^2(x) = E(\{Y_1 - m(x)\}^2 \mid X_1 = x)$ and assume that

$$0 < \sigma^2 = \int_{-\infty}^{\infty} \sigma^2(x)\, dF(x) < \infty . \tag{1.2}$$

Let F_n be the empirical df of $X_1, \cdots, X_n$, $F_n^{-1}(t) = \inf\{x : F_n(x) \geqq t\}$, $t \in I = [0, 1]$,

$$\psi_n(t) = \int_0^{F_n^{-1}(t)} \sigma^2(x)\, dF_n(x) \quad \text{and} \quad \psi(t) = \int_0^{F^{-1}(t)} \sigma^2(x)\, dF(x) , \quad t \in I , \tag{1.3}$$

so that both ψ_n and ψ are nondecreasing (in t) and, in addition, ψ_n is stochastic in nature. For every n ($\geqq 1$), consider a stochastic process $W_n = \{W_n(t), t \in I\}$ by introducing a sequence of integer-valued, nondecreasing and right continuous functions $\{k_n(t), t \in I\}$ where $k_n(t) = \max\{k : \psi_n(k/n) \leqq t\psi_n(1)\}$, $t \in I$, and then letting $W_n(t) = \{n\psi_n(1)\}^{-\frac{1}{2}} S_{nk_n(t)}$, $t \in I$, where

$$S_{nk} = \sum_{j=1}^{k} \{Y_{nj} - m(X_{n,j})\} , \quad k = 1, \cdots, n ; \qquad S_{n0} = 0 . \tag{1.4}$$

By an interesting application of the Skorokhod embedding under a conditional setup, Bhattacharyya (1974) has shown that under suitable regularity conditions, W_n weakly converges (in the Skorokhod J_1-topology on $D[0, 1]$) to a standard Wiener process. We shall show that for a sequence of two-dimensional time

Received June 4, 1975; revised October 20, 1975.

[1] Work partially supported by the Air Force Office of Scientific Research, A.F.S.C., U.S.A.F., Grant No. AFOSR 74-2736 and partially by a grant of the Deutsche Forschungsgemeinschaft during the residence of the author at Freiburg i. Br., W. Germany (visiting Albert-Ludwig Universität).

AMS 1970 *classification.* 60F05.

Key words and phrases. Brownian sheets, induced order statistics, martingales, uniform integrability, weak convergence and Wiener processes.

parameter stochastic processes constructed from the S_{nk} in (1.4), weak convergence to a Brownian sheet process holds under less stringent conditions and the conclusion applies to W_n as well. Since for a multiparameter process, the classical embedding technique runs into difficulties, the task is completed here by using certain convergence properties of $\psi_n(t)$, defined in (1.3). The main results are presented in Section 2 and the proofs in the concluding section.

2. The main results. We assume that the following *uniform integrability* condition (less restrictive than Condition 1 of [1]) holds:

$$\sup_{x\in R} E(\{Y_1 - m(x)\}^2 I(|Y_1 - m(x)| > s) \mid X_1 = x) \to 0 \quad \text{as} \quad s \to \infty, \tag{2.1}$$

where $I(A)$ stands for the indicator function of the set A and $R = (-\infty, \infty)$.

Let us now consider a two-dimensional time parameter stochastic process $W_n^* = \{W_n^*(\mathbf{t}), \mathbf{t} \in I^2\}$, $I^2 = [0, 1]^2$, $\mathbf{t} = (t_1, t_2)$, where we set

$$W_n^*(\mathbf{t}) = \{n\psi_n(1)\}^{-\frac{1}{2}} S_{[nt_1]k_n(\mathbf{t})}, \qquad \mathbf{t} \in I^2, \tag{2.2}$$

$[q]$ being the largest integer $\leqq q\ (> 0)$ and

$$k_n(\mathbf{t}) = \max\{k : \psi_{[nt_1]}(k/[nt_1]) \leqq t_2\psi_{[nt_1]}(1)\}, \qquad \mathbf{t} \in I^2. \tag{2.3}$$

Note that W_n^* belongs to the space $D^2[0, 1]$. Also, let $W^* = \{W^*(\mathbf{t}), \mathbf{t} \in I^2\}$ be a standard Brownian sheet on I^2. Then, our main theorem may be presented as follows.

THEOREM 1. *Under* (1.2) *and* (2.1), *as* $n \to \infty$,

$$W_n^* \to_{\mathscr{D}} W^*, \quad \textit{in the } J_1\textit{-topology on} \quad D^2[0, 1]. \tag{2.4}$$

The proof of the theorem is outlined in Section 3. In the rest of this section, we consider the following two results which are required in the sequel. Let $\mathscr{B}_{n,k}$ be the sigma-field generated by $\{(X_{n,j}, Y_{nj}), 1 \leqq j \leqq k\}$, for $k = 1, \cdots, n$ and $\mathscr{B}_{n,0}$ be the trivial sigma-field. Also, let $\mathscr{A}_n$ be the sigma-field generated by $(X_1, \cdots, X_n)$, $n \geqq 1$. Finally, let $\{c_{ni}, 1 \leqq i \leqq n; n \geqq 1\}$ be a double sequence of arbitrary constants and we define

$$S_{nk}^* = \sum_{j=1}^k c_{nj}\{Y_{nj} - m(X_{n,j})\}, \quad k = 1, \cdots, n; \qquad S_{n0}^* = 0. \tag{2.5}$$

LEMMA 2. *For every* $n\ (\geqq 1)$, $\{S_{nk}^*, \mathscr{B}_{n,k}; 1 \leqq k \leqq n\}$ *is a martingale closed on the right by* S_{nn}^*.

PROOF. Note that by Lemma 1 of Bhattacharyya (1974), given $\mathscr{A}_n$, the Y_{nj} are all conditionally independent with Y_{nj} having the conditional df $G_{X_{n,j}}$ and conditional mean $m(X_{n,j})$, $j = 1, \cdots, n$. Hence, on writing $E(S_{nn}^* \mid \mathscr{B}_{n,k}) = E(E\{S_{nn}^* \mid \mathscr{A}_n, \mathscr{B}_{n,k}\})$ it follows by standard arguments that by (2.5), $E(S_{nn}^* \mid \mathscr{B}_{n,k}) = S_{nk}^*$, $k \leqq n$. □

Note that the $\sigma^2(X_i)$ are i.i.d.rv's with mean σ^2, so that by the Khintchine strong law of large numbers, as $n \to \infty$,

$$\psi_n(1) = \int_{-\infty}^{\infty} \sigma^2(x)\, dF_n(x) = n^{-1} \sum_{i=1}^n \sigma^2(X_i) \to \sigma^2, \quad \text{almost surely (a.s.).} \tag{2.6}$$

Also, by (2.1),

$$\sup_{x \in R} \sigma^2(x) < \infty . \tag{2.7}$$

Finally, by the Glivenko-Cantelli theorem, as $n \to \infty$,

$$\max_{1 \leq k \leq n} |F(X_{n,k}) - k/n| \to 0 \quad \text{a.s.}, \tag{2.8}$$

and hence, by (1.3), (2.7) and (2.8), we arrive at the following.

LEMMA 3. *Under* (1.2) *and* (2.1), $\sup\{\psi_n(t) : t \in I\} \leqq \sup\{\sigma^2(x) : x \in R\}$ *for all* n, *and*

$$|\psi_n(t) - \psi(t)| \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty , \quad \text{for every} \quad t \in I . \tag{2.9}$$

3. Proof of Theorem 1. We need to show that (i) the finite dimensional distributions (f.d.d.) of $\{W_n^*\}$ converge to the correspondig ones of W^* and (ii) W_n^* is tight. Unlike the case of partial sums of independent rv's, here for $k_j \leqq n_j$, $j = 1, 2$, with $n_1 \leqq n_2$, $(X_{n_1,1}, \cdots, X_{n_1,k_1}) \cap (X_{n_2,1}, \cdots, X_{n_2,k_2})$ need not be equal to $(X_{n_1,1}, \cdots, X_{n_1,k})$ with $k = k_1 \wedge k_2 = \min(k_1, k_2)$, and this introduces additional complications in the proof.

First, consider the convergence of the f.d.d.'s. Note that

$$EW^*(\mathbf{s})W^*(\mathbf{t}) = \mathbf{s} \wedge \mathbf{t} = (s_1 \wedge t_1)(s_2 \wedge t_2) \qquad \text{for all } \mathbf{s}, \mathbf{t} \in I^2 . \tag{3.1}$$

We shall show that W_n^* has asymptotically the same covariance structure. For this, first, consider nonstochastic integers

$$n_j = [n\alpha_j] , \qquad k_j = [n_j \gamma_j] , \quad (\alpha_j, \gamma_j) \in I^2 \qquad \text{for} \quad j = 1, 2 . \tag{3.2}$$

Note that for α_j or γ_j equal to 0, $S_{n_j k_j} = 0$, and hence, we need to confine ourselves only to the range $0 < \alpha_j, \gamma_j \leqq 1$, $j = 1, 2$. Also, note that

$$\{n\psi_n(1)\}^{-1} E(S_{n_1k_1} S_{n_2k_2}) = \{n\psi_n(1)\}^{-1} E\{E(S_{n_1k_1} S_{n_2k_2} \mid \mathscr{A}_n)\} , \tag{3.3}$$

where by Lemma 3 and the Schwarz inequality, $|E(S_{n_1k_1}S_{n_2k_2} \mid \mathscr{A}_n)|/n\psi_n(1) \leqq \{n\psi_n(1)\}^{-1} \cdot \{n_1\psi_{n_1}(k_1/n_1) n_2 \psi_{n_2}(k_2/n_2)\}^{\frac{1}{2}}$ is bounded for all n, and thus, $E(S_{n_1k_1}S_{n_2k_2} \mid \mathscr{A}_n)/n\psi_n(1) \to_p c$, a constant, implies that $E(S_{n_1k_1}S_{n_2k_2})/n\psi_n(1) \to c$, as $n \to \infty$. For this reason, first, we show that under (2.1) and (3.2),

$$\{n\psi_n(1)\}^{-1} E(S_{n_1k_1} S_{n_2k_2} \mid \mathscr{A}_n) \to_p (\alpha_1 \wedge \alpha_2)(\gamma_1 \wedge \gamma_2) , \qquad \text{as} \quad n \to \infty . \tag{3.4}$$

If $\alpha_1 = \alpha_2$, then, by Lemma 1 of [1] and our Lemma 2 (with all the $c_{ni} = 1$), we have $E(S_{n_1k_1}S_{n_1k_2} \mid \mathscr{A}_n)/n\psi_n(1) = E(S_{n_1k}^2 \mid \mathscr{A}_n)/n\psi_n(1) = n_1\psi_{n_1}(k/n_1)/n\psi_n(1) \to \alpha_1(\gamma_1 \wedge \gamma_2)$ a.s., as $n \to \infty$, by (2.3) and (2.6), where $k = k_1 \wedge k_2$. Hence, (3.4) holds. Next, consider the case of $\alpha_1 < \alpha_2$. It may be noted that for $n_1 \leqq n_2$, $(X_{n_1,1}, \cdots, X_{n_1,k_1}) \cap (X_{n_2,1}, \cdots, X_{n_2,k_2}) = (X_{n_1,1}, \cdots, X_{n_1,k_1-q})$ where $q\,(\leqq k_1 \wedge (n_2 - n_1))$ is a nonnegative integer valued random variable. Thus, in this case, the lhs (left hand side) of (3.4) reduces to

$$\{n_1\psi_{n_1}((k_1 - q)/n_1)/n\psi_n(1)\} = n^{-1}n_1\{\psi_{n_1}(1)/\psi_n(1)\}\{\psi_{n_1}((k_1 - q)/n_1)/\psi_{n_1}(1)\} . \tag{3.5}$$

Hence, by virtue of (2.6) and (3.2), it remains to show that for $\alpha_1 < \alpha_2$,

$$(k_1 - q)/n_1 \to_p \gamma_1 \wedge \gamma_2 , \qquad \text{as} \quad n \to \infty . \tag{3.6}$$

First, consider the case of $\gamma_1 < \gamma_2$. Let $u(t) = 1$ or 0 according as t is $\geqq$ or < 0 and let $M = \sum_{i=n_1+1}^{n_2} u(X_{n_1,k_1} - X_i)$. Then, a little examination reveals that $q = 0$ if $M \leqq k_2 - k_1$. Also,

$$(3.7)\qquad P\{M = m\} = n_1\binom{n_1-1}{k_1-1}\binom{n_2-n_1}{m} \int_{-\infty}^{\infty} \{F(x)\}^{k_1+m-1}\{1 - F(x)\}^{n_2-k_1-m}\, dF(x) .$$

Further, $(k_2 - k_1)/n \to (\alpha_2\gamma_2 - \alpha_1\gamma_1) = (\alpha_2 - \alpha_1)\gamma_1 + \alpha_2(\gamma_2 - \gamma_1)$, $\gamma_2 > \gamma_1$, so that from (3.7), it readily follows that as $n \to \infty$,

$$(3.8)\qquad P\{M \leqq k_2 - k_1\} \to 1 , \qquad \text{i.e.,}\quad P\{q = 0\} \to 1 .$$

Thus, (3.6) holds. Let us next consider the case of $\gamma_1 \geqq \gamma_2$ but $\alpha_2\gamma_2 \geqq \alpha_1\gamma_1$. Note that, by definition,

$$(3.9)\qquad X_{n_1,k_1-q} < X_{n_2,k_2} < X_{n_1,k_1-q+1} ,$$

so that $F(X_{n_1,k_1-q}) < F(X_{n_2,k_2}) < F(X_{n_1,k_1-q+1})$. Also, by (2.8), $|F(X_{n_2,k_2}) - k_2/n_2| \to 0$ a.s. and $\max\{|F(X_{n_1,j}) - j/n_1| : 1 \leqq j \leqq n_1\} \to 0$ a.s., as $n \to \infty$. Hence, using (3.2), we obtain immediately that $(k_1 - q)/n_1 \to_p \gamma_2$, which proves (3.6). Finally, the case of $\alpha_1 < \alpha_2$ and $\gamma_1 \geqq \gamma_2$ but $\alpha_2\gamma_2 < \alpha_1\gamma_1$ can be dealt with in a similar manner. Hence, (3.4) holds in general. To obtain (3.4) for $k_j = k_n(\mathbf{t}_j)$, $j = 1, 2$, defined by (2.3) [instead of (3.2)], we note that, by definition, $0 \leqq t_2\psi_{[nt_1]}(1) - \psi_{[nt_1]}(k_n(\mathbf{t})/[nt_1]) \leqq [nt_1]^{-1} \max\{\sigma^2(X_{[nt_1],j}) : 1 \leqq j \leqq [nt_1]\} \to 0$; $\mathbf{t} = (t_1, t_2)$, and hence, $\psi_{[nt_1]}(k_n(\mathbf{t})/[nt_1])/\psi_{[nt_1]}(1) \to t_2$, in probability, as $n \to \infty$. As a result, by (2.2), (2.3) and (3.4), we conclude that

$$(3.10)\qquad \begin{aligned} &E\{W_n^*(\mathbf{s})W_n^*(\mathbf{t}) \mid \mathscr{A}_n\} \to_p (\mathbf{s} \wedge \mathbf{t}) &&\text{as } n \to \infty ,\\ &E\{W_n^*(\mathbf{s})W_n^*(\mathbf{t})\} \to \mathbf{s} \wedge \mathbf{t} , &&\text{as } n \to \infty, \text{ for all } \mathbf{s}, \mathbf{t} \in I^2 . \end{aligned}$$

Now, for every fixed m ($\geqq 1$) and $\mathbf{t}_1, \cdots, \mathbf{t}_m \in I^2$, consider an arbitrary linear compound

$$(3.11)\qquad T_n = \sum_{j=1}^{m} \lambda_j W_n^*(\mathbf{t}_j) \qquad \text{where}\quad \boldsymbol{\lambda} \neq \mathbf{0} \quad\text{and}\quad ||\boldsymbol{\lambda}|| < \infty .$$

By virtue of (2.2) and (2.3), (3.11) may be rewritten as

$$(3.12)\qquad T_n = \{n\psi_n(1)\}^{-\frac{1}{2}} \sum_{i=1}^{n} c_{ni}\{Y_{ni} - m(X_{n,i})\} , \qquad \text{where}\quad \max_{1\leqq i\leqq n} |c_{ni}| < c < \infty ,$$

and the c_{ni} depend on (i) $\boldsymbol{\lambda}$, (ii) $\mathbf{t}_j$, $j = 1, \cdots, m$ and (iii) the triangular array of order statistics $\{X_{k,j}, 1 \leqq j \leqq k; 1 \leqq k \leqq n\}$. Now, given $\mathscr{A}_n$, the $Y_{n,j} - m(X_{n,j})$ are all conditionally independent with means 0 and conditional variances $\sigma^2(X_{n,j})$, the c_{ni} are all held fixed and (2.1) insures that under this conditional setup, the Lindeberg condition holds for the sequence $\{c_{nj}(Y_{nj} - m(X_{n,j})), 1 \leqq j \leqq n\}$. So that, conditionally, given $\mathscr{A}_n$, T_n is asymptotically normal with mean 0 and variance

$$(3.13)\qquad \{n\psi_n(1)\}^{-1} \sum_{i=1}^{n} c_{ni}^2\sigma^2(X_{n,i}) .$$

On the other hand, if $\mathbf{V}_{n,m}$ be the conditional (given $\mathscr{A}_n$) covariance matrix of

$\{W_n^*(\mathbf{t}_1), \cdots, W_n^*(\mathbf{t}_m)\}$, then (3.13) is equal to $\boldsymbol{\lambda}'\mathbf{V}_{n,m}\boldsymbol{\lambda}$, and by (3.10), it converges in probability to $\boldsymbol{\lambda}'\mathbf{V}_m\boldsymbol{\lambda}$, where $\mathbf{V}_m = ((\mathbf{t}_j \wedge \mathbf{t}_k))_{j,k=1,\cdots,m}$ is positive definite. Hence, unconditionally too, T_n, is asymptotically normal with mean 0 and variance $\boldsymbol{\lambda}'\mathbf{V}_m\boldsymbol{\lambda}$. Thus, for every $\mathbf{t}_1, \cdots, \mathbf{t}_m \in I^2$, the joint df of $\{W_n^*(\mathbf{t}_1), \cdots, W_n^*(\mathbf{t}_m)\}$ is asymptotically the same as that of $\{W^*(\mathbf{t}_1), \cdots, W^*(\mathbf{t}_m)\}$, and the proof of the convergence of the f.d.d.'s is complete.

Let us now consider the proof of tightness of $\{W_n^*\}$. Note that, for every $(s_1, t_1) < (s_2, t_2)$, the increment over the block is

$$
\begin{aligned}
&W_n^*((s_1, t_1), [s_2, t_2]) \\
(3.14) \qquad &= W_n^*(s_2, t_2) - W_n^*(s_2, t_1) - W_n^*(s_1, t_2) + W_n^*(s_1, t_1) \\
&= \{n\psi_n(1)\}^{-\frac{1}{2}}(S_{n_2k_2} - S_{n_2k_1} - S_{n_1q_2} + S_{n_1q_1}) \\
&= \{n\psi_n(1)\}^{-\frac{1}{2}}(\textstyle\sum_{j=k_1+1}^{k_2}\{Y_{n_2j} - m(X_{n_2,j})\} - \sum_{j=q_1+1}^{q_2}\{Y_{n_1j} - m(X_{n_1,j})\}),
\end{aligned}
$$

where $n_j = [ns_j]$, $\psi_{n_2}(k_j/n_2)/\psi_{n_2}(1) \to t_j$ and $\psi_{n_1}(q_j/n_1)/\psi_{n_1}(1) \to s_j$ for $j = 1, 2$. Note that the sums on the rhs of (3.14) may contain a common subset. However, this drops out with the result that for some h ($\geqq 0$), there are $k_2 - k_1 - h$ and $q_2 - q_1 - h$ terms for which the corresponding $X_{n_j,r}$ are all distinct. A similar representation holds for any other neighbouring block. Thus, if we set $s_1 < s_2 < s_3$ and $t_1 < t_2 < t_3$ such that the Lebesgue measure of the blocks are equal, i.e., $(s_3 - s_2)(t_3 - t_2) = (s_2 - s_1)(t_3 - t_2) = \cdots = (s_2 - s_1)(t_2 - t_1) = \lambda$, say, then, by virtue of (3.14) and Lemma 1 of [1], we can again show by steps similar to those employed in the first part of the proof of the theorem that under the conditional model (given $\mathscr{A}_n$),

$$
(3.15) \qquad E(\{W_n^*((s_1, t_1), [s_2, t_2])W_n^*((s_2, t_1), [s_3, t_2])\}^2 \mid \mathscr{A}_n) \leqq c_n\lambda^2,
$$

almost everywhere,

where c_n is bounded for every n and $\lim_{n\to\infty} c_n = 1$; a similar inequality holds for any other neighbouring blocks. Hence, using the multiparameter extension (viz., [2]) of the Billingsley inequality ([3], page 128), the tightness of $\{W_n^*\}$ follows readily from (3.15). □

REMARKS. Bhattacharyya (1974) considered the convergence of $\{n\psi(1)\}^{-\frac{1}{2}}S_{[nt]}$, $t \in I$ and in his Skorokhod representation, he needed an additional condition that $\sigma^2(x)$ is of bounded variation on R. By changing $S_{[nt]}$ to $S_{nk_n(t)}$, $t \in I$, we are able to eliminate the above condition in so far as the weak convergence result is concerned. Also, if we consider the weak convergence of $\{W_n^*(1, t), t \in I\}$ to a standard Wiener process, the proof simplifies a lot. Here, the martingale result of Lemma 2 (for $c_{ni} = 1$, $i \geqq 1$) and (2.1) provide the access to the first theorem of Section 3 of McLeish (1974) and the proof follows quite simply. The condition that F is continuous can also be dropped as in [1].

Whereas the weak convergence of $\{W_n^*(1, t), t \in I\}$ has been used in [1] to provide some asymptotic tests for regression functions, our Theorem 1 may be used to provide some sequential analogues of these tests.

Acknowledgment. Thanks are due to the referee for his critical reading of the manuscript and the set of useful comments on it.

REFERENCES

[1] BHATTACHARYYA, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics. *Ann. Statist.* **2** 1034–1039.

[2] BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multi-parameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1670.

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

[4] MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probability* **2** 620–628.

DEPARTMENT OF BIOSTATISTICS
THE UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27514