Journal of Multivariate Analysis **79**, 191–218 (2001)
doi:10.1006/jmva.2000.1967, available online at http://www.idealibrary.com on IDEAL®

A Nonparametric Test of Serial Independence for Time Series and Residuals

Kilani Ghoudi

Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada
E-mail: ghoudi@uqtr.uquebec.ca

Reg J. Kulperger

University of Western Ontario, London, Ontario, Canada
E-mail: rjk@fisher.stats.uwo.ca

and

Bruno Rémillard

Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada
E-mail: remillar@uqtr.uquebec.ca

Received November 17, 1998; published online June 19, 2001

This paper presents nonparametric tests of independence that can be used to test the independence of p random variables, serial independence for time series, or residuals data. These tests are shown to generalize the classical portmanteau statistics. Applications to both time series and regression residuals are discussed.
© 2001 Academic Press
AMS 1990 subject classifications: 62G10; 60F05; 62E20.
Key words and phrases: independence; serial independence; empirical processes; pseudo-observations; residuals; weak convergence; Cramér–von Mises statistics.

1. INTRODUCTION

Testing for independence is very important in statistical applications. These tests arise in many different settings, in particular when checking the dependence of p random variables, one usually carries out an independence test. Such a test is also required when verifying that consecutive observations of a time series are independent. Finally when checking the hypotheses of most linear models, one often needs to test serial independence of the error terms.

0047-259X/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

The problem of testing the independence of p random variables is quite old. The first tests were based on correlation measures (Kendall, Spearman). More powerful tests are based on the empirical distribution function, and were considered by Hoeffding (1948), Blum *et al.* (1961), and Cotterill and Csörgő (1982, 1985).

In a time series setting, one is more interested in testing serial independence, that is one would like to verify if consecutive observations $U_i, ..., U_{i+p-1}$ are independent. This problem received considerable attention in the literature. It is usually tackled using Portmanteau statistics based on the autocorrelation functions (see Brockwell and Davis, 1991; Kulperger and Lockhart, 1998). Recently, Skaug and Tjøstheim (1993) proposed a test for serial pairwise independence based on the empirical distribution function. Their work generalizes Hoeffding (1948) to serial independence. Interesting extensions of this test can be found in Hong (1998) and Hong (2000). Delgado (1996) used a Blum, Kiefer, and Rosenblatt statistic in the serial independence context. He showed that the process converges weakly, but that the limiting process is not very useful when trying to tabulate critical values of test statistics.

Portmanteau type statistics are also used when checking serial independence of the errors of a linear model. The Durbin Watson test is the standard diagnostic for serial independence of the errors of linear regression and is also based on some measure of the correlation between the errors (see Jobson, 1991).

This paper develops nonparametric tests of independence and serial independence that can be applied in either of the above three cases. In other words, the tests proposed here apply when testing the independence of p random variables or the serial independence of time series data and residuals. These tests are Cramér–von Mises or Kolmogorov–Smirnov functionals of some empirical processes. This paper shows that under the independence (serial independence) hypothesis these empirical processes converge to Gaussian limits with quite convenient covariance functions.

It is also shown that if the U_i's have continuous distribution function then the limiting distributions of the test statistics do not depend on the underlying law of the U_i's. This holds when testing independence of p random variables, serial independence of time series data and serial independence of residuals of a classical linear regression. In other cases such as residuals of an autoregressive model, the limiting distribution depends, in general, on the law of the U_i's.

The test of the independence of p random variables shall be called Setting 1 while that of the serial independence in time series data is called Setting 2. Test of serial independence for residuals or residual-likes observations is referred to as Setting 3, and using the terminology of Ghoudi and Rémillard (1998a), this shall also be called the pseudo-observations situation.

The idea behind the construction in the time series setting is quite similar to that proposed by Skaug and Tjøstheim (1993) and Delgado (1996). It uses the famous method of time delay, which is well known in the chaotic time series literature. The construction together with few definitions are given in the next section. Section 3 presents the properties of the limiting processes. It is shown, in particular, that these processes have very convenient covariance functions and that they admit attractive representations in terms of Brownian drums. Section 4 defines test statistics used in this work. It is shown that for Cramér–von Mises type statistics, one obtains a closed form for the asymptotic distribution. The Cornish–Fisher asymptotic expansion (Abramowitz and Stegun, 1964) and Imhof's characteristic function inversion algorithm (Imhof (1961)) are used to tabulate the distribution and the quantiles of these statistics. Section 5 discusses applications of these statistics. A first subsection considers the case where U_i's are time series data. It uses a special alternative to provide a power study. The second subsection provides a simulation study comparing the power of the tests discussed in Section 4 with Delgado's (1996) test. Section 6 is devoted to the proofs of the results stated within this work.

2. DEFINITIONS AND RESULTS

This section defines and states the results for the asymptotic behavior of the empirical processes used to develop the test procedures. First a characterization of the independence of p random variables is provided. Then the section gets divided into three subsections. Each subsection presents one of the three particular cases described earlier.

Let $U_1, ..., U_p$ be $p \geqslant 2$ random variables. For $1 \leqslant j \leqslant p$, let $K^{(j)}$ denotes the marginal distribution function of U_j and for any $t = (t^{(1)}, ..., t^{(p)})$ in $\mathbb{R}^p$, let $K_p(t) = P\{U_1 \leqslant t^{(1)}, ..., U_p \leqslant t^{(p)}\}$ be the joint distribution function of $U_1, ..., U_p$. Now for any $A \subset I_p = \{1, ..., p\}$, and any $t \in \mathbb{R}^p$, set

$$\mu_A(t) = \sum_{B \subset A} (-1)^{|A \backslash B|} K_p(t^B) \prod_{j \in A \backslash B} K^{(j)}(t^{(j)}),$$

where $|A|$ denotes the number of elements in A, where by convention, $\prod_{\varnothing} = 1$ and where (t^B) is the vector with components

$$(t^B)^{(i)} = \begin{cases} t^{(i)}, & i \in B; \\ \infty, & i \in I_p \backslash B. \end{cases}$$

Then one can state the following characterization of the independence of $U_1, ..., U_p$.

PROPOSITION 2.1. *$U_1, ..., U_p$ are independent if and only if $\mu_A \equiv 0$, for all $A \subset \{1, ..., p\}$.*

2.1. *Testing the Independence of p Random Variables.* This section deals with the classical problem of testing the independence of p random variables. This problem received considerable attention in the literature. Here one is mainly interested in tests based on the empirical distribution function. Such tests were considered by Hoeffding (1948), Blum *et al.* (1961), and Cotterill and Csörgő (1982, 1985). In particular, if one lets $\varepsilon_i = (\varepsilon_i^{(1)}, ..., \varepsilon_i^{(p)})$; $i = 1, ..., n$ be a random sample of $\mathbb{R}^p$ valued random variables, it is desired to test the independence of the components $\varepsilon^{(1)}, ..., \varepsilon^{(p)}$. Blum *et al.* (1961) proposed the following empirical process

$$\beta_{n,p}(t) = \sqrt{n}\left\{K_{n,p}(t) - \prod_{i=1}^{p} K_n^{(i)}(t^{(i)})\right\},$$

where $K_{n,p}$ is the joint empirical distribution function and where $K_n^{(i)}$ is the ith empirical marginal distribution function. It is shown that, except for the case $p = 2$, the asymptotic covariance function of the process $\beta_{n,p}$ is not very convenient. In other words it does not provide a nice way to produce critical values of some useful test statistics. Here alternative processes are proposed. The idea behind the introduction of these processes comes from the characterization of independence given in Proposition 2.1.

To this end, for any $A \subset I_p = \{1, ..., p\}$ and any $t = (t^{(1)}, ..., t^{(p)}) \in \mathbb{R}^p$ let

$$R_{n,A}(t) = \sqrt{n} \sum_{B \subset A} (-1)^{|A \setminus B|} K_{n,p}(t^B) \prod_{i \in A \setminus B} K_n^{(i)}(t^{(i)}).$$

Using the multinomial formula (8) in Section 6, the above reduces to

$$\begin{aligned} R_{n,A}(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{B \subset A} (-1)^{|A \setminus B|} \prod_{k \in B} \mathbb{I}\{\varepsilon_i^{(k)} \leqslant t^{(k)}\} \prod_{j \in A \setminus B} K_n^{(j)}(t^{(j)}) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{k \in A} [\mathbb{I}\{\varepsilon_i^{(k)} \leqslant t^{(k)}\} - K_n^{(k)}(t^{(k)})], \end{aligned}$$

where $\mathbb{I}$ denotes the indicator function. The asymptotic behaviour of these processes is stated next.

THEOREM 2.1. *Let $\varepsilon_1, ..., \varepsilon_n$ be independent and identically distributed random vectors and suppose that $K^{(k)}$, the marginal distribution function of $\varepsilon_1^{(k)}$, is continuous for $k = 1, ..., p$. Then if $\varepsilon_1^{(1)}, ..., \varepsilon_1^{(p)}$ are independent, the* *processes* $(R_{n,A})_{A \subset I_p}$ *converge in* $D(\mathbb{R}^p)$ *to independent mean zero Gaussian processes* R_A *having covariance functions given by*

$$C_A(s,t) = \mathrm{Cov}\{R_A(s), R_A(t)\}$$

$$= \prod_{i \in A} [\min\{K^{(i)}(s^{(i)}), K^{(i)}(t^{(i)})\} - K^{(i)}(s^{(i)})\, K^{(i)}(t^{(i)})].$$

In fact one can easily verify that the processes $R_{n,A}$ are related to the process $\beta_{n,p}$ through the following representations

$$\beta_{n,p}(t) = \sum_{B \subset I_p;\, |B| > 1} R_{n,B}(t^B) \prod_{i \in I_p \backslash B} K_n^{(i)}(t^{(i)}),$$

and

$$R_{n,A}(t) = \sum_{B \subset A} (-1)^{|A \backslash B|}\, \beta_{n,p}(t^B) \prod_{i \in A \backslash B} K_n^{(i)}(t^{(i)}).$$

As a corollary one gets the asymptotic behaviour of the $\beta_{n,p}$ given in Blum *et al.* (1961). The inverse is also true, that is the above theorem can be obtained using the results of Blum *et al.* (1961). This does not shorten the proof, therefore a direct proof is presented in Section 6.

Test statistics based on these processes are introduced in Section 4. In particular, asymptotics of Cramér–von Mises and Kolmogorov–Smirnov type statistics shall be discussed. The presentation of these tests is delayed to Section 4, in order to treat the three settings together, that is p random variables, serial independence in time series and serial independence with pseudo-observations.

2.2. *Testing Serial Independence of Time Series Data.* Here the time series framework is presented. Let $\{U_i\}_{i \geqslant 1}$ be a stationary and ergodic time series. Let $p \geqslant 2$ be a fixed integer. For any integer $1 \leqslant i \leqslant n-p+1$, set $\varepsilon_i = (U_i, U_{i+1}, ..., U_{i+p-1})$ and for any $t = (t^{(1)}, ..., t^{(p)}) \in \mathbb{R}^p$, define

$$K_{n,p}(t) = \frac{1}{n} \sum_{i=1}^{n-p+1} \mathbb{I}\{\varepsilon_i^{(1)} \leqslant t^{(1)}, ..., \varepsilon_i^{(p)} \leqslant t^{(p)}\}.$$

Let K_p be the distribution function of $(\varepsilon_1^{(1)}, ..., \varepsilon_1^{(p)})$, that is,

$$K_p(t) = P(\varepsilon_1^{(1)} \leqslant t^{(1)}, ..., \varepsilon_1^{(p)} \leqslant t^{(p)}) = P(U_1 \leqslant t^{(1)}, ..., U_p \leqslant t^{(p)}).$$

As in the previous section, for any set $A \subset I_p = \{1, 2, ..., p\}$, let

$$R_{n,A}(t) = \sqrt{n} \sum_{B \subset A} (-1)^{|A \backslash B|}\, K_{n,p}(t^B) \prod_{i \in A \backslash B} K_n^{(i)}(t^{(i)}) \qquad (1)$$

It follows that $R_{n,A}=0$ if $|A|\leqslant 1$ and that the coefficient of $\prod_{i\in A} K_n^{(i)}(t^{(i)})$ in $R_{n,A}$ is $(-1)^{|A|-1}(|A|-1)$. One can also rewrite

$$R_{n,A}(t)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n-p+1}\sum_{B\subset A}(-1)^{|A\backslash B|}\prod_{k\in B}\mathbb{I}\{\varepsilon_i^{(k)}\leqslant t^{(k)}\}\prod_{j\in A\backslash B}K_n^{(j)}(t^{(j)})$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n-p+1}\prod_{k\in A}[\mathbb{I}\{\varepsilon_i^{(k)}\leqslant t^{(k)}\}-K_n^{(k)}(t^{(k)})].$$

In this setting, it is clear that a translate $B=A+k$ of a given set A generates basically the same process $R_{n,A}$. Therefore, without loss of generality, one can restrict the attention to the processes $R_{n,A}$ where $A\in\mathcal{A}_p=\{A\subset I_p;\ 1\in A, \text{ and } |A|>1\}$.

Let K denotes the common distribution function of the U_i's. The asymptotic behaviour of $R_{n,A}$ is stated in the following theorem.

THEOREM 2.2. *If K is continuous and if the U_i's are independent, then the processes $(R_{n,A})_{A\in\mathcal{A}_p}$ converge in $D(\mathbb{R}^p)$ to independent mean zero Gaussian processes R_A having covariance functions given by*

$$C_A(s,t)=\operatorname{Cov}\{R_A(s),R_A(t)\}=\prod_{i\in A}[\min\{K(s^{(i)}),K(t^{(i)})\}-K(s^{(i)})\,K(t^{(i)})].$$

Skaug and Tjøstheim (1993) based their test for serial independence on Cramer–von Mises functional of the above process $R_{n,A}$ where A is of the form $\{1,k\}$ for some $2\leqslant k\leqslant p$. This shows that their test is only for serial pairwise independence and not serial independence in general. But it was argued that for many alternatives, this test is more powerful then tests based on the joint distribution of p variables.

Delgado (1996) used the following process

$$\beta_{n,p}(t)=K_{n,p}(t)-\prod_{i=1}^{p}K_n^{(i)}(t^{(i)}).$$

Once again this process is related to our processes via the representations

$$\beta_{n,p}(t)=\sum_{B\subset I_p;\ |B|>1}R_{n,B}(t^B)\prod_{i\in I_p\backslash B}K_n^{(i)}(t^{(i)}),$$
$$R_{n,A}(t)=\sum_{B\subset A}(-1)^{|A\backslash B|}\beta_{n,p}(t^B)\prod_{i\in A\backslash B}K_n^{(i)}(t^{(i)}).$$

Next, set

$$\beta_p(t)=\sum_{B\subset I_p;\ |B|>1}R_B(t^B)\prod_{i\in I_p\backslash B}K(t^{(i)}),$$

As a corollary one gets the asymptotic behavior of the $\beta_{n,p}$ given in Lemma 1 of Delgado (1996). The inverse is also true, as illustrated in the remark following Theorem 2.1.

COROLLARY 2.1. *If K is continuous and if the U_i's are independent, then the process $\beta_{n,p}$ converges in $D(\mathbb{R}^p)$ to a continuous Gaussian process β_p with covariance function Γ_β given by*

$$\begin{aligned}\Gamma_\beta(x, y) = {} & K_p(x_p \wedge y_p) - K_p(x_p)\, K_p(y_p) \\ & + \sum_{k=1}^{p-1} K_k(x_k)\, K_k((\tau^{p-k} y)_k)\{K_{p-k}(y_{p-k} \wedge (\tau^k x)_{p-k}) \\ & - K_{p-k}(y_{p-k})\, K_{p-k}((\tau^k x)_{p-k})\} \\ & + \sum_{k=1}^{p-1} K_k(y_k)\, K_k((\tau^{p-k} x)_k)\{K_{p-k}(x_{p-k} \wedge (\tau^k y)_{p-k}) \\ & - K_{p-k}(x_{p-k})\, K_{p-k}((\tau^k y)_{p-k})\}, \qquad x, y \in \mathbb{R}^p,\end{aligned}$$

where for any $d \geqslant 1$, K_d is the distribution function of $(U_1, ..., U_d)$ and where $a \wedge b$ is the vector with components $\{\min(a^{(1)}, b^{(1)}), ..., \min(a^{(k)}, b^{(k)})\}$, $a, b \in \mathbb{R}^k$ and $(\tau^j x)_k = (x^{(j+1)}, ..., x^{(j+k)})$.

As noted by Delgado the asymptotic covariance function of $\beta_{n,p}$ is not convenient for the tabulation of critical values of Cramér–von Mises functionals. In his paper Delgado proposed the use of a permutation method to approximate these critical values. However, one should be very careful when using simulations to tabulate critical values of tests of independence, since any simulation procedure uses pseudo-random variables having some kind of serial dependence. The effect of using simulation could be negligible in relatively small samples, as pointed by Delgado (1996) and the results in Section 4. However, theoretically one should be able to detect this kind of serial dependence at least for very large samples. Thus using simulation will result in miscalculating the critical values. It is therefore essential to have an alternate method for evaluating these critical values. This is exactly the main aim of this paper. As shown in Sections 3 and 4, the covariance functions of the processes $R_{n,A}$ are quite easy to handle. In fact, explicit form of their eigenvalues and eigenfunctions will be given.

Remark 2.1. If the distribution function K is known, one could use the statistics $\tilde{R}_{n,A}$ obtained by replacing $K_n^{(i)}$ by K in (1). For example, this would be the case if one wishes to verify if sequence of observations is a sequence of independent and identically distributed random variables with uniform marginals. As noted in Section 6, the processes $\tilde{R}_{n,A}$ and $R_{n,A}$ are asymptotically equivalent.

2.3. *Testing Serial Independence with Pseudo-Observations.* This section deals with the pseudo-observations situation. To be precise, let $\{X_i\}_{i\geq 1}$ be an $\mathcal{X}$ valued time series. Let H be a function from $\mathcal{X}$ to an interval T of $\mathbb{R}$ and consider the series $\{U_i = H(X_i)\}_{i\geq 1}$. Suppose that the series of U_i's is stationary and ergodic and that the distribution function K of U_1 is continuous. The aim is to test if $U_i ..., U_{i+p-1}$ are independent. If H is known, this reduces to the time series setting discussed in the previous section. On the other hand, if H is unknown and is estimated by some function H_n and U_i is estimated by $\tilde{U}_i = H_n(X_i)$, then this is called the pseudo-observations case. Even though the $\tilde{U}_i$'s depend on n, no subscript is added for the sake of simplicity. Residuals are just a special case of pseudo-observations. Empirical processes based on pseudo-observations like the $\tilde{U}_i$'s are considered by Barbe *et al.* (1996) and Ghoudi and Rémillard (1998a, 1998b). This section redefines the processes $R_{n,A}$, $A \in \mathcal{A}_p$ for this setting and studies their asymptotic behavior.

For each $i = 1, ..., n-p+1$ set $\varepsilon_i = (U_i, ..., U_{i+p-1})$ and $e_i = (\tilde{U}_i, ..., \tilde{U}_{i+p-1})$. Let $\tilde{K}_{n,p}$ be the empirical distribution function of the e_i's and let $\tilde{K}_n^{(k)}$; $k = 1, ..., p$ be the kth empirical marginal distribution. Then, for this setting, the process $R_{n,A}(t)$ is given by

$$\begin{aligned} R_{n,A}(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n-p+1} \sum_{B \subset A} (-1)^{|A \setminus B|} \prod_{k \in B} \mathbb{I}\{e_i^{(k)} \leq t^{(k)}\} \prod_{j \in A \setminus B} \tilde{K}_n^{(j)}(t^{(j)}) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n-p+1} \prod_{k \in A} [\mathbb{I}\{e_i^{(k)} \leq t^{(k)}\} - \tilde{K}_n^{(k)}(t^{(k)})]. \end{aligned} \tag{2}$$

One also defines the processes

$$\tilde{\beta}_{n,p}(x) = \sqrt{n} \left\{ \tilde{K}_{n,p}(t) - \prod_{i=1}^{p} \tilde{K}_n^{(i)}(t^{(i)}) \right\},$$

and

$$\tilde{\beta}^*_{n,p}(x) = \sqrt{n} \left\{ \tilde{K}_{n,p}(t) - \prod_{i=1}^{p} K(t^{(i)}) \right\}.$$

The process $\tilde{\beta}^*_{n,p}$ is a special case of the empirical process based on pseudo-observations studied by Ghoudi and Remillard (1998b). In particular their Theorem 2.1 applied to this context yields the asymptotic behavior of $\tilde{\beta}^*_{n,p}$. Before the precise statement of this result, we introduce the following conditions.

(**R1**) There exists some positive continuous function $r: \mathcal{X} \to \mathbb{R}$ such that $\inf_{x \in \mathcal{X}} r(x) > 0$ and $E\{r(X)\}$ is finite. Further let $\mathcal{C}_r$ be a closed subset of the Banach space of all continuous functions f from $\mathscr{X}$ to $\mathbb{R}$ such that $\|f\|_r = \sup_{x \in \mathscr{X}} |f(x)/r(x)|$ is finite. Assume that there exists a continuous version $\tilde{H}_n$ of H_n such that $\sqrt{n}\, \|\tilde{H}_n(x) - H_n(x)\|_r$ converges in probability to zero.

(**R2**) Suppose also that for any $f \in \mathscr{C}_r$ with $g = f/r$, and any continuous ψ on $\mathbb{R}$, $0 \leqslant \psi \leqslant 1$, the processes

$$\alpha_{n,\, j,\, \psi \circ g}(s, t)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n-p+1} \Bigg[\psi(g(X_{i+j}))\, \mathbb{I}\{\varepsilon_i^{(j)} \leqslant t^{(j)} + s r(X_{i+j})\} \prod_{k \neq j} \mathbb{I}\{\varepsilon_i^{(k)} \leqslant t^{(k)}\}$$

$$- E\left\{ \psi(g(X_j))\, \mathbb{I}\{\varepsilon_1^{(j)} \leqslant t^{(j)} + s r(X_j)\} \prod_{k \neq j} \mathbb{I}\{\varepsilon_1^{(k)} \leqslant t^{(k)}\} \right\} \Bigg],$$

are such that for any compact subset C of $\mathbb{R}$ and for $s \in \mathbb{R}$,

$$\sup_{t \in C} |\alpha_{n,\, j,\, \psi \circ g}(s/\sqrt{n}, t) - \alpha_{n,\, j,\, \psi \circ g}(0, t)| \tag{3}$$

converge in probability to zero. Finally suppose that if $\alpha_n(t) = \alpha_{n,1,1}(0, t)$ and $\mathbb{H}_n = \sqrt{n}(\tilde{H}_n - H)$, then $(\alpha_n, \mathbb{H}_n)$ converges in $C(\mathbb{R}^d) \times \mathscr{C}_r$ to a process $(\alpha, \mathscr{H})$.

(**R3**) The support T of K is an interval of $\mathbb{R}$, K admits a density $k(\cdot)$ on T which is bounded on every compact subset of T and that there exists a version of the conditional distribution of $X \prod_{k \neq j} \mathbb{I}\{\varepsilon_1^{(k)} \leqslant t^{(k)}\}$ given $\varepsilon_1^{(j)} = H(X_j) = t^{(j)}$, denoted by $P_{j,t}$, such that for any $f = f_0 + \theta r$ with $f_0 \in \mathscr{C}_r$ and $\theta \in \mathbb{R}$, the mappings

$$t \mapsto \mu_j(t, f) = k_j(t^{(j)})\, E\left\{ f(X_j) \prod_{k \neq j} \mathbb{I}\{\varepsilon_1^{(k)} \leqslant t^{(k)}\} \mid \varepsilon_1^{(j)} = t^{(j)} \right\},$$

are continuous on T.

Finally suppose that for any compact subset C of T,

$$\lim_{M \to \infty} \int_M^\infty \sup_{s \in C} P_s(r(X) > u)\, du = 0.$$

With these notations Theorem 2.1 of Ghoudi and Remillard (1998b) applied to this context may be restated as follows

THEOREM 2.3. *If* $\{U_i\}_{i \geq 1}$ *is a stationary and ergodic time series and if* $U_1, ..., U_p$ *are independent, then if conditions* R1–R3 *are satisfied the process* $(\tilde{\beta}^*_{n,p})$ *converges in* $D(\mathbb{R}^p)$ *to*

$$\tilde{\beta}^*_p(t) = \beta_p(t) - \sum_{j=1}^{d} \mu_j(t, \mathbb{H}).$$

As a corollary one gets.

COROLLARY 2.2. *Under the conditions of Theorem* 2.3 *the process* $R_{n,A}(t)$ *converges in* $D(\mathbb{R}^p)$ *to*

$$\tilde{R}_A(t) = \sum_{B \subset A} (-1)^{|A \setminus B|} \tilde{\beta}^*_p(t^B) \prod_{i \in A \setminus B} K(t^{(i)}).$$

As an example it will be shown how these results apply to the linear regression residuals.

Linear Regression Residuals. Consider a classical linear regression model, $Y = a + b'Z + U$, where $Y \in \mathbb{R}$, $Z \in \mathbb{R}^d$ and where Z and U are independent. To apply the results of this paper to this case, note that in this context $X = (Y, Z)$ and $U = H(X) = H(Y, Z) = Y - a - b'Z$. One also has $H_n(x) = H_n(y, z) = y - a_n - b'_n z$ where a_n and b_n could be taken as the least square estimate of a and b respectively. $r(y, x) = r(z) = 1 + \|z\|$, $\mathscr{C}_r = \{a + b'z; a \in \mathbb{R} \text{ and } b \in \mathbb{R}^d\}$ and $\mathbb{H}(y, z) = A + B'z$ where (A, B, β_p) is the joint weak limit of $(\sqrt{n}(a_n - a), \sqrt{n}(b_n - b), \beta_{n,p})$. The $\mu_j(t, \mathbb{H})$'s reduce to

$$\mu_j(t, \mathbb{H}) = k(t^{(j)}) \prod_{k \neq j} K(t^{(k)}) \{A + B'E(Z)\}.$$

The application of Corollary 2.2 to this setting yields the following proposition.

PROPOSITION 2.2. *Suppose that the design matrix is not singular,* $E(\|Z\|^2)$ *is finite and* U *admits a support* T *that is an interval of* $\mathbb{R}$ *and a continuous bounded density on this support. Then if* a_n *and* b_n *are the least square estimates of* a *and* b, *the processes* $R_{n,A}$, *defined by* (2) *for* $A \in \mathscr{A}_p$, *converge to the independent centered Gaussian processes* $\{R_A\}_{A \in \mathscr{A}_p}$ *given in Theorem* 2.2.

Remark 2.2. Note that for general linear models the limiting process does not necessarily simplify to R_A, in particular, for the autoregressive model, $Y_i - \mu = \phi(Y_{i-1} - \mu) + U_i$, the process $\tilde{R}_A$ is not equal to R_A. In fact, even for the simple case of $A=\{1, j\}$, it follows from (3.7) in Ghoudi and Rémillard (1998b) that

$$\tilde{R}_{\{1,j\}}(t)=R_{\{1,j\}}(t)+k(t^{(j)})\,G(t^{(1)})\,\phi^{j-2}\Phi,$$

where $G(s)=E(\varepsilon_1 \mathbb{I}\{\varepsilon_1 \leqslant s\})$ and Φ is a random variable representing the limit of $\sqrt{n}(\hat{\phi}_n-\phi)$.
Note also that for $\tilde{\beta}_{n,p}$, even in the linear regression setting, the extra term in the limit does not simplify.

3. PROPERTIES OF THE LIMITING PROCESSES

This section shows that the limiting processes admit very convenient covariance functions and that they can be represented in term of the process β_p defined in Corollary 2.1 or more appropriately in terms of Brownian drums.

The covariance functions C_A, $A\in\mathcal{A}_p$ are very easy to use. In fact C_A is the product of $|A|$ covariance functions of Brownian bridges. That is, the eigenvalues and the eigenfunctions of C_A are quite easy to obtain and are summarized in the next proposition.

PROPOSITION 3.1. *Let $k=|A|$. Then the covariance function C_A admits eigenvalues and eigenfunctions given by*

$$\lambda_{j_1,\ldots,j_k}=\prod_{l=1}^{k}(j_l\pi)^{-2} \qquad \text{and} \qquad f_{j_1,\ldots,j_k}(t_1,\ldots,t_k)=\prod_{l=1}^{k}\sin(j_l\pi t_l),$$

respectively, for $(j_1,\ldots,j_k)\in\mathbb{N}^k$.

Next consider the representation of the process R_A. The first result is straightforward and is stated in the next proposition.

PROPOSITION 3.2. *If K is a continuous distribution function then*

$$R_A(t)=\sum_{B\subset A}(-1)^{|A\setminus B|}\,\beta_p(t^B)\prod_{i\in A\setminus B}K(t^{(i)}).$$

For the second representation, recall that by Theorem 2.2, the processes $(R_A)_{A\in\mathcal{A}_p}$ are all independent with covariance function C_A, $A\in\mathcal{A}_p$. This representation shows that these processes can also be written in terms of

Wiener sheets. For, let W be a Wiener sheet on $[0, 1]^p$, that is a mean zero continuous Gaussian process on $[0, 1]^p$ with covariance function given by

$$E\{W(s)\, W(t)\} = \prod_{i=1}^{p} s^{(i)} \wedge t^{(i)}, \qquad s, t \in [0, 1]^p.$$

Next, define the Wiener drums or Brownian drums D_A, $A \in \mathscr{A}_p$, in terms of W by

$$D_A(s) = \sum_{B \subset A} (-1)^{|A \setminus B|}\, W(s_B) \prod_{k \in A \setminus B} s^{(k)}, \qquad A \in \mathscr{A}_p,$$

where $s_B^{(i)} = s^{(i)}$ if $i \in B$ and 1 otherwise. To simplify the statement of the results, assume K is continuous and consider the following rescaled version of the processes R_A.

$$\tilde{D}_A(s) = R_A\{K^{-1}(s^{(1)}), ..., K^{-1}(s^{(p)})\}, \qquad A \in \mathscr{A}_P.$$

The representation of $\tilde{D}_A(s)$ is given in the following proposition.

PROPOSITION 3.3. *The joint law of* $(D_A)_{A \in \mathscr{A}_p}$ *is the same as that of* $(\tilde{D}_A)_{A \in \mathscr{A}_p}$. *Moreover*

$$D_A(s) = W(s_A) - E\{W(s_A) \mid \mathscr{K}_A\},$$

where $\mathscr{K}_A$ *is the sigma-algebra generated by the values of the Wiener sheet on the boundary of* $[0, 1]^A$, *that is,*

$$\mathscr{K}_A = \sigma\{W(t);\ t^{(i)} = 0 \text{ or } 1, \text{ for some } i \in A\}.$$

It is thus justified to call D_p a Wiener drum or a Brownian drum, since it vanishes on the boundary of $[0, 1]^p$. The proof of the proposition requires two steps and is given in Section 6.

4. TEST STATISTICS

This section studies test statistics based on the processes considered earlier. In fact for the three settings of Subsections 2.1, 2.2, and 2.3 one can introduce Cramér–von Mises or Kolmogorov–Smirnov type statistics using the processes $R_{n,A}$'s. It will be shown that the limiting distribution of the Cramér–von Mises statistics is in general easy to obtain.

First, define the Cramér–von Mises statistics as

$$T_{n,A} = \int R_{n,A}^2(t)\, dK_{n,p}(t), \tag{4}$$

for the first two settings and

$$\tilde{T}_{n,A} = \int R_{n,A}^2(t)\, d\tilde{K}_{n,p}(t),$$

for the pseudo-observations setting. Next the Kolmogorov–Smirnov statistic is given by

$$S_{n,A} = \sup_t |R_{n,A}(t)|. \tag{5}$$

To test the independence or the serial independence one can, in particular, use the statistics $V_{n,p} = \sum_A T_{n,A}$, $\bar{V}_{n,p} = \max_A T_{n,A}$ or $W_{n,p} = \max_A S_{n,A}$, where A ranges over all the subset of I_P for the test of independence of p random variables and over the class $\mathscr{A}_p$ for the test of serial independence in a time series. When dealing with pseudo-observations one replaces $T_{n,A}$ by $\tilde{T}_{n,A}$ in the above.

Next the asymptotic distribution of $T_{n,A}$ is established. But first, set

$$\xi_k = \sum_{(i_1, \dots, i_k) \in \mathbb{N}^k} \frac{1}{\pi^{2k}(i_1 \cdots i_k)^2} Z_{i_1, \dots, i_k}^2,$$

where the $Z_{i_1, \dots, i_k}$'s are independent $N(0, 1)$ random variables. The asymptotics of $T_{n,A}$ are given next

Lemma 4.1. *Under the conditions of Theorem* 2.1 *or Theorem* 2.2, *the statistic $T_{n,A}$ converges in law to $\xi_{|A|}$.*

The critical values of the asymptotic distribution of ξ_k are easy to compute. In fact this can be achieved by first computing the cumulants given by (6), and then applying the Cornish Fisher asymptotic expansion, or by inversion of the characteristic function. This inversion is obtained by the numerical integration method proposed by Imhof (1961), or the improved version of this algorithm introduced by Deheuvels and Martynov (1996). The following provides the cumulant of order m of the ξ_k

$$\kappa_m = \frac{2^{m-1}(m-1)!}{\pi^{2km}} \zeta(2m)^k, \tag{6}$$

where $\zeta(\cdot)$ denotes the Riemann zeta function.

TABLE I

Critical Values of the Distribution of ξ_k

p	$k=2$	$k=3$	$k=4$	$k=5$
0.900000	0.047071	0.006695	0.000992	0.000152
0.905000	0.047951	0.006783	0.001000	0.000153
0.910000	0.048882	0.006877	0.001010	0.000154
0.915000	0.049872	0.006976	0.001020	0.000155
0.920000	0.050928	0.007082	0.001030	0.000156
0.925000	0.052058	0.007195	0.001042	0.000157
0.930000	0.053273	0.007317	0.001054	0.000158
0.935000	0.054584	0.007448	0.001067	0.000159
0.940000	0.056009	0.007590	0.001081	0.000161
0.945000	0.057565	0.007745	0.001096	0.000162
0.950000	0.059279	0.007915	0.001113	0.000164
0.955000	0.061183	0.008104	0.001131	0.000166
0.960000	0.063323	0.008315	0.001152	0.000168
0.965000	0.065760	0.008555	0.001176	0.000170
0.970000	0.068585	0.008833	0.001203	0.000173
0.975000	0.071938	0.009161	0.001235	0.000176
0.980000	0.076052	0.009560	0.001274	0.000180
0.985000	0.081359	0.010071	0.001323	0.000185
0.990000	0.088810	0.010779	0.001391	0.000191
0.995000	0.101347	0.011940	0.001500	0.000202

Table I provides an approximation of the cut-off values obtained from the Cornish Fisher asymptotic expansion with the first six cumulants.

A careful examination of the asymptotic distributions of the $T_{n,A}$'s shows that their expectations and their variances diminish considerably as the cardinality of A increases. For example, the mean of ξ_k is equal to $1/6^k$ and the asymptotic variance is given by $\mathrm{Var}(\xi_k)=2/90^k$. So when using the statistics $V_{n,p}$ or $\bar{V}_{n,p}$, the biggest contribution tends to come from the sets A of small sizes. To avoid this problem it is more convenient to work with a standardized version of the statistics. To be specific let $T^*_{n,A}=(T_{n,A}-E(\xi_k))/\sqrt{\mathrm{Var}(\xi_k)}$ and define $V^*_{n,p}=\sum_A T^*_{n,A}$ and $\bar{V}^*_{n,p}=\max_A T^*_{n,A}$, where the range of the sets A is that given in the definition of $V_{n,p}$ and $\bar{V}_{n,p}$. Lemma 4.1 implies that $T^*_{n,A}$ converges in distribution to $\xi^*_k=(\xi_k-E(\xi_k))/\sqrt{\mathrm{Var}(\xi_k)}$, that $V^*_{n,p}$ converges in distribution to V^*_p and that $\bar{V}^*_{n,p}$ converges in distribution to $\bar{V}^*_p$, where $V^*_p=\sum_A \xi^*_{|A|}$ and $\bar{V}^*_p=\max_A \xi^*_{|A|}$ with

TABLE II

Distributions of ξ_k^*; $k = 2$, 3, and 4 and of $\bar{V}_3^*$ and $\bar{V}_4^*$ for Settings 1 and 2

x	$P\{\xi_2^* \le x\}$	$P\{\xi_3^* \le x\}$	$P\{\xi_4^* \le x\}$	Setting 1 $P\{\bar{V}_3^* \le x\}$	Setting 2 $P\{\bar{V}_3^* \le x\}$	Setting 1 $P\{\bar{V}_4^* \le x\}$	Setting 2 $P\{\bar{V}_4^* \le x\}$
0.0	0.64549	0.69481	0.86226	0.28950	0.18687	0.01454	0.07779
0.2	0.71697	0.75618	0.88928	0.38871	0.27869	0.03949	0.14171
0.4	0.77282	0.80462	0.91065	0.48055	0.37138	0.08131	0.21895
0.6	0.81656	0.84278	0.92762	0.56195	0.45887	0.1387	0.30234
0.8	0.85104	0.87292	0.94116	0.63223	0.53805	0.20762	0.38587
1.0	0.87840	0.89687	0.95200	0.69198	0.60783	0.28290	0.46542
1.2	0.90028	0.91588	0.96074	0.74232	0.66829	0.35993	0.53858
1.4	0.91789	0.93116	0.96780	0.78452	0.72010	0.43513	0.60426
1.6	0.93216	0.94347	0.97353	0.81980	0.76418	0.50606	0.66222
1.8	0.94378	0.95345	0.97820	0.84925	0.80150	0.57126	0.71273
2.0	0.95329	0.96156	0.98202	0.87383	0.83301	0.63005	0.75635
2.2	0.96110	0.96819	0.98514	0.89434	0.85955	0.68228	0.79377
2.4	0.96755	0.97363	0.98770	0.91146	0.88188	0.72817	0.82570
2.6	0.97288	0.97810	0.98981	0.92576	0.90066	0.76813	0.85286
2.8	0.97730	0.98178	0.99155	0.93772	0.91644	0.80269	0.87588
3.0	0.98098	0.98483	0.99299	0.94772	0.92969	0.83241	0.89536
3.2	0.98404	0.98735	0.99417	0.95608	0.94082	0.85786	0.91182
3.4	0.98659	0.98944	0.99515	0.96309	0.95018	0.87958	0.92570
3.6	0.98873	0.99117	0.99597	0.96896	0.95804	0.89807	0.93741
3.8	0.99052	0.99262	0.99664	0.97388	0.964644	0.91378	0.94727
4.0	0.99201	0.99382	0.99720	0.97801	0.97020	0.92710	0.95557

$\xi_{|A|}^*$ independent for different A's. Table II provides the distribution function of the statistics ξ_k^* and $\bar{V}_p^*$, as approximated by Imhof's technique.

Imhof's technique can also be used to tabulate the distribution of V_p^*. Unfortunately, no closed form can be given for the Kolmogorov–Smirnov statistic. However, for continuous U_i's and for both the independence of p random variables and the serial independence in a time series, $S_{n,A}$ and $T_{n,A}$ are distribution free, therefore one can approximate their limiting distributions by simulating the limiting Brownian drum process. Note that one can also simulate sequences of independent uniform (0, 1) random variables. Table III reports the result of 5000 simulation of pseudo-random sequences generated using the KISS algorithm, Marsaglia and Zaman (1995). It shows that the effect of the dependence contained in pseudo-random sequence is negligible.

This fact is also illustrated in Table IV, where the simulation procedure used Splus random number generator, for Cramér–von Mises (4) and

TABLE III

Sample 0.95 Quantiles for $\bar{V}_{n,2}^*$

Sample size	100	500	1000	Asymptotic value
95th quantile	1.921	1.898	1.940	1.937

TABLE IV

Sample 0.95 Quantiles for Lags $k = 1, 2, 3$

Sample Size	T_{n,A_k}			S_{n,A_k}		
	k			k		
	1	2	3	1	2	3
100	.0596	.0598	.0588	.6605	.6553	.6605
200	.0595	.0596	.0567	.6890	.6825	.6814
400	.0584	.0571	.0626	.7126	.7164	.7318

Kolmogorov–Smirnov (5) functionals in the time series setting, for sets A of the form $\{1, 1+k\}$ where $k = 1$, 2 and 3. k is often called the lag. These simulations were done for different sample sizes $n = 100$, $n = 200$ and $n = 400$ and with 2500 Monte Carlo replicates. Observe that the results for the Cramér–von Mises statistics compares very well with the asymptotic quantiles given in Table I. One also notices from the simulation results that the asymptotics take effect for reasonable sample sizes and does so more quickly for the Cramér–von Mises statistic than the Kolmogorov–Smirnov statistic. One also notices that the critical values are indeed consistent with $R_{n,A}$ being identically distributed for different $|A|$. It is also seen that the Cramér–von Mises statistic is quite consistent across different sample sizes n, but that the Kolmogorov–Smirnov statistic has critical values that change a small amount as the sample size n increases. This is quite consistent with the results for Kolmogorov–Smirnov statistic in the usual setting where a finite sample correction is often used; see, for example, Stephens (1986).

Moreover, a careful examination of these results show that, as expected, the processes for different lags are independent and identically distributed Gaussian processes. This pairwise independence of the processes occurs for moderate sample size n (in the range of 100).

Note that by Proposition 2.2, all the results stated above for the serial independence in the time series setting will apply to the test of serial independence for the residuals of linear regression models. In fact, when working with the residuals of a classical linear regression model, the limiting processes are exactly the same as those obtained for the time series setting. In particular Tables I, II and IV apply to these residuals.

5. POWER STUDIES

This section presents two power studies. The first is a comparison with the classical portmanteau statistics. The second discusses the performance of the tests presented earlier compared to that of Delgado (1996). The next subsection introduces Portmanteau processes as a special case of the processes introduced in Section 2. Then it presents a simulation study for the power of the statistics, introduced in Section 4, in detecting a product alternative.

5.1. *Portmanteau Processes.* The classical portmanteau statistic is based on sample autocorrelations. Its sampling distribution is based on the fact that if the data comes from an i.i.d. sequence, the normalized sample autocorrelations are asymptotically independent standard normal random variables. This section shows that by properly choosing the set A, the processes $R_{n,A}$ are in fact empirical processes based on lags and that the classical Portmanteau statistics are functionals of these processes. Theorem 2.2 shows that these processes are asymptotically independent Gaussian processes. Some functional of these processes, such as Cramér–von Mises and Kolmogorov–Smirnov statistics, are distribution free. In this sense it will be shown that these empirical processes play the role of a generalized Portmanteau process. This section presents some uses for this process. It also gives a power study, for tests based on these processes, against a product process alternative which consists of a 1-dependent sequence with zero lag 1 covariance.

Assume one disposes of a stationary and ergodic sequence of random variables $\{U_i\}_{i\geqslant 1}$ with common distribution K, the idea of the portmanteau statistics for $p=2$ is to consider pairs of random variables (U_i, U_{i+k}) at lag k. In fact, to simplify the presentation, only the case $p=2$ will be discussed here. Now, set $A_k=\{1, k+1\}$ and consider the process R_{n,A_k}. From the previous results one concludes that if the U_i's are independent then the processes R_{n,A_k} converge to independent Gaussian processes with common covariance function $C(s_1, s_2, t_1, t_2)=\{K(s_1)\wedge K(t_1)-K(s_1)\,K(t_1)\}\{K(s_2)\wedge K(t_2)-K(s_2)\,K(t_2)\}$.

In the classical setting of the portmanteau statistics, the lag k covariance is given by $c(k)=\mathrm{Cov}(U_i, U_{i+k})$, and its sample estimate is obtained via

$$c_n(k)=\frac{1}{n-k}\sum_{i=1}^{n-k}(U_i-\bar{U})(U_{i+k}-\bar{U}),$$

where $\bar{U}$ is the sample mean. First note that the normalizing factor $n-k$ can be replaced with n without affecting the asymptotic of $c_n(k)$. With this modification one easily obtains

$$\sqrt{n}\,c_n(k)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{n,A_k}(t_1, t_2)\,dt_1\,dt_2.$$

It is well known that under the independence hypothesis, the $c_n(k)$'s converge to independent $N(0, \tau^2)$ distributions, where

$$\tau^2 = \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{K(s_1) \wedge K(t_1) - K(s_1)\, K(t_1)\}\; ds_1\, dt_1 \right]^2 = \sigma^4,$$

where $\sigma^2 = \mathrm{Var}(U_1)$ and where the last equality follows from Hoeffding's Lemma (see Block and Fang, 1988). As discussed in the previous section one might consider other more powerful statistics such as those of the Cramér–von Mises or the Kolmogorov–Smirnov types discussed earlier.

5.1.1. *Detection of a Product Process Alternative.* To get an idea about the power of these tests, the following product process alternative is considered.

Assume that the data generating mechanism for this alternative consist of the following product process

$$U_i = X_{i-1} X_i, \tag{7}$$

where the X_i's are independent and identically distributed random variables with mean zero and finite variance. This process is a 1-dependent sequence with zero lag 1 covariance $c(1)$. The statistics described earlier are used to detect if the sequence of U_i's form an i.i.d. sequence.

The classical portmanteau statistic is usually defined as the sum of squares of sample correlations, (see Brockwell and Davis, 1991). Such a test will have poor power, in particular the power does not tend to 1 as the sample size tends to ∞. That is, the process (7) is a particularly difficult alternative to be detected by a portmanteau statistic.

Since the joint bivariate distributions at various lags for this alternative process are not products of the marginals, one should expect a test based on the process (1) to have some power against this type of product alternative. To this end, a simulation study is considered next. For the purpose of this study assume that X_i is a $N(0, 1)$ random variable. The simulation

TABLE V

Product Process Rejection Rates at 0.05 Level Test, at Various Lags $k = 1, 2, 3$

Sample Size	T_{n,A_k}			S_{n,A_k}		
	k			k		
	1	2	3	1	2	3
100	.218	.056	.050	.158	.052	.053
200	.455	.049	.050	.332	.052	.063
400	.986	.058	.058	.640	.049	.056

consists of generating sequences of observations from the alternative (12) and noting the rejection rates for each of the Cramér–von Mises and the Kolmogorov–Smirnov functionals discussed earlier. This was done for different sample sizes with 2000 Monte Carlo replicates. The estimated critical values from Table IV were used and the study was conducted for three different lags. Table V summarizes the results.

Since the product process is a 1-dependent sequence, the rejection rates using the lag 2 or 3, processes should be and are 0.05. The lag 1 process has good power rejecting the null hypothesis of independence. The Cramér–von Mises statistic does better than the Kolmogorov–Smirnov statistic. The Cramér–von Mises statistic has power 0.218 at sample size 100 and increases to power 0.986 at sample size 400. The Kolmogorov–Smirnov statistic has somewhat smaller power, but still has good power for the product process alternative. The rejection rates for the lag 2 and 3 processes are 0.05. Thus the tests based on these lags also recognizes that the data is consistent with a 1-dependent sequence. Again one should notice that a portmanteau

TABLE VI

Percentage of Rejection of $U_t = bU_{t-1} + \delta_t$

		Statistic														
		$V_{n,p}$			$\bar{V}_{n,p}$			$V^*_{n,p}$			$\bar{V}^*_{n,p}$			Delgado		
		p			p			p			p			p		
b	n	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4
0.0	20	5.1	5.1	5.1	5.1	5.1	4.7	5.1	4.7	5.5	5.1	4.7	4.9	5.1	5.3	4.9
	50	4.9	4.4	5.4	4.9	4.1	4.9	4.9	5.0	5.5	4.9	4.9	5.7	4.9	4.3	5.3
	100	5.4	5.6	4.7	5.4	5.7	4.6	5.4	4.8	5.1	5.4	5.1	5.3	5.4	4.9	4.5
0.1	20	7.6	7.2	5.8	7.6	6.9	5.7	7.6	6.7	6.4	7.6	5.8	6.1	7.6	9.2	8.5
	50	11.3	8.6	8.9	11.3	8.6	8.0	11.3	8.5	8.2	11.3	7.6	6.5	11.3	11.2	13.4
	100	17.4	13.6	10.6	17.4	13.5	10.8	17.4	11.4	8.8	17.4	11.7	7.5	17.4	16.9	16.5
0.2	20	12.9	9.9	8.8	12.9	9.4	7.6	12.9	8.6	9.6	12.9	6.7	8.1	12.9	14.3	14.0
	50	26.0	18.1	17.4	26.0	17.9	16.9	26.0	16.6	14.4	26.0	14.7	10.8	26.0	25.1	27.5
	100	46.7	38.2	31.6	46.7	38.5	32.2	46.7	32.6	24.5	46.7	32.0	22.2	46.7	45.7	42.2
0.3	20	20.5	16.4	12.4	20.5	15.7	11.3	20.5	14.7	14.0	20.5	11.3	12.4	20.5	23.4	21.3
	50	47.8	37.2	33.9	47.8	38.1	33.0	47.8	34.0	28.2	47.8	31.5	23.2	47.8	47.3	46.5
	100	77.8	70.4	60.9	77.8	71.3	62.9	77.8	64.2	50.4	77.8	64.5	49.9	77.8	77.0	71.0
0.4	20	30.7	23.7	18.6	30.7	22.6	17.1	30.7	20.3	21.1	30.7	15.6	17.8	30.7	33.3	31.5
	50	70.4	60.3	54.8	70.4	61.6	55.1	70.4	56.8	47.6	70.4	54.7	41.7	70.4	70.0	67.7
	100	95.2	92.1	87.2	95.2	93.1	88.8	95.2	88.7	79.7	95.2	89.9	80.6	95.2	94.8	91.1
0.5	20	43.8	33.5	29.6	43.8	32.4	26.4	43.8	29.7	31.2	43.8	23.8	27.6	43.8	45.0	43.9
	50	87.0	80.3	75.8	87.0	80.7	77.1	87.0	76.9	68.4	87.0	75.7	64.5	87.0	86.7	84.5
	100	99.5	98.8	97.8	99.5	99.0	98.2	99.5	98.2	95.4	99.5	98.4	96.3	99.5	99.3	98.8
0.6	20	57.7	49.0	39.5	57.7	47.2	36.5	57.7	44.6	41.7	57.7	37.2	35.8	57.7	60.1	55.3
	50	95.9	93.0	90.4	95.9	93.4	91.1	95.9	91.4	86.5	95.9	90.9	84.4	95.9	95.6	94.7
	100	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.6	99.9	99.8	99.6	99.9	100	99.9
0.7	20	68.9	61.4	50.5	68.9	59.8	49.0	68.9	57.1	54.3	68.9	50.2	48.8	68.9	71.6	66.2
	50	99.0	97.8	96.7	99.0	98.1	96.8	99.0	97.6	94.7	99.0	97.2	93.5	99.0	98.9	98.3
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
0.8	20	78.4	71.4	62.5	78.4	70.5	60.7	78.4	68.7	64.5	78.4	62.3	60.2	78.4	80.3	75.2
	50	99.8	99.6	99.4	99.8	99.5	99.4	99.8	99.4	98.8	99.8	99.2	98.8	99.8	99.8	99.9
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
0.9	20	79.5	72.5	63.3	79.5	71.3	61.4	79.5	69.6	65.6	79.5	62.9	61.3	79.5	80.4	75.1
	50	99.8	99.5	99.1	99.8	99.6	99.1	99.8	99.4	98.5	99.8	99.4	98.3	99.8	99.8	99.7
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

statistic based on sample correlations has power of about 0.05 for all sample sizes.

5.2. *Comparison with Delgado's Test.* In this section a simulation study is carried out to compare the power of the tests statistics presented here to that of Delgado (1996). For sake of comparison with Delgado (1996), the simulation is done using his two alternatives.

First a sequence of observations following an $AR(1)$ model $U_t = b * U_{t-1} + \delta_t$ is considered, then a second sequence where $U_t = b\delta_{t-1}^2 + \delta_t$ is used. In both situations δ_t; $t \geq 1$ are independent $N(0, 1)$ random variables. For each of these studies, 5000 Monte-Carlo replicates are generated and the percentage of time the independence hypothesis is rejected is recorded. The cut-off values for all tests were obtained by simulations. Table VI provides the results for the first model and Table VII for the second model. It can be seen that there is no clear winner, that is Delgado statistic performs a little better for the $AR(1)$ setting. But V, $\bar{V}$, V^* or $\bar{V}^*$ are more powerful in detecting the nonlinear alternative considered in the second study.

TABLE VII

Percentage of Rejection of $U_t = b\delta_{t-1}^2 + \delta_t$

		Statistic														
		$V_{n,p}$			$\bar{V}_{n,p}$			$V^*{}_{n,p}$			$\bar{V}^*_{n,p}$			Delgado		
		p			p			p			p			p		
b	n	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4
0.0	20	4.8	5.4	5.4	4.8	5.4	5.4	4.8	5.7	5.7	4.8	5.6	5.3	4.8	4.9	5.8
	50	5.2	4.8	5.0	5.2	4.9	5.6	5.2	4.7	5.7	5.2	4.6	5.1	5.2	4.9	5.3
	100	5.5	5.0	4.9	5.5	5.2	5.3	5.5	5.1	5.0	5.5	5.3	5.1	5.5	5.5	4.9
0.1	20	4.5	5.3	4.5	4.5	5.0	4.7	4.5	5.4	4.7	4.5	5.3	4.9	4.5	5.0	4.8
	50	5.0	5.5	5.9	5.0	5.5	5.1	5.0	6.0	5.6	5.0	5.8	5.9	5.0	4.5	4.7
	100	6.1	6.3	5.5	6.1	6.7	5.7	6.1	6.0	5.8	6.1	6.1	5.5	6.1	5.0	4.5
0.2	20	5.6	6.2	6.0	5.6	5.9	5.6	5.6	6.6	5.3	5.6	6.7	5.4	5.6	5.1	4.4
	50	7.2	6.1	5.8	7.2	6.1	5.5	7.2	7.7	6.3	7.2	6.7	6.3	7.2	4.7	5.2
	100	10.5	9.1	7.3	10.5	8.1	6.3	10.5	8.6	8.2	10.5	7.0	6.7	10.5	6.4	5.6
0.3	20	6.5	5.6	5.6	6.5	5.6	5.3	6.5	6.4	5.7	6.5	6.2	5.9	6.5	4.9	4.7
	50	9.2	7.8	7.9	9.2	7.3	6.1	9.2	9.4	9.3	9.2	8.4	8.3	9.2	5.0	5.7
	100	18.0	14.4	10.5	18.0	12.9	8.9	18.0	14.7	12.8	18.0	10.9	9.1	18.0	8.3	6.7
0.4	20	7.5	6.0	5.6	7.5	5.9	5.2	7.5	7.5	7.1	7.5	7.2	7.0	7.5	5.6	4.7
	50	11.6	8.3	9.1	11.6	8.3	7.7	11.6	10.8	10.8	11.6	9.2	8.6	11.6	5.8	6.0
	100	17.1	18.5	10.9	17.1	17.3	9.5	17.1	18.5	13.2	17.1	13.8	9.2	17.1	10.2	6.8
0.5	20	7.3	6.1	5.7	7.3	6.3	5.8	7.3	7.6	7.5	7.3	8.0	7.3	7.3	5.4	5.8
	50	13.8	10.8	9.4	13.8	10.3	7.5	13.8	13.9	12.2	13.8	11.0	10.3	13.8	6.6	6.7
	100	32.9	23.5	18.3	32.9	22.1	16.1	32.9	23.8	21.8	32.9	17.6	14.0	32.9	11.9	9.8
0.6	20	7.9	7.6	6.8	7.9	7.1	6.1	7.9	8.6	8.2	7.9	8.1	7.8	7.9	6.2	5.5
	50	15.5	11.7	11.3	15.5	11.0	8.9	15.5	13.9	13.7	15.5	11.8	11.2	15.5	7.5	7.6
	100	38.6	27.5	21.4	38.6	25.8	19.4	38.6	27.6	25.2	38.6	20.7	17.2	38.6	13.4	10.6
0.7	20	7.3	7.9	6.7	7.3	7.4	6.0	7.3	9.6	8.2	7.3	8.7	7.6	7.3	6.9	5.5
	50	17.4	12.9	11.3	17.4	12.0	9.5	17.4	15.7	14.6	17.4	12.6	11.9	17.4	7.8	7.5
	100	42.3	29.6	22.0	42.3	29.3	21.1	42.3	30.0	27.8	42.3	22.8	17.6	42.3	14.4	11.0
0.8	20	8.9	7.5	6.5	8.9	7.2	5.5	8.9	8.9	8.5	8.9	9.4	8.1	8.9	6.9	5.6
	50	17.2	13.9	11.7	17.2	12.8	9.4	17.2	16.5	15.1	17.2	13.4	13.0	17.2	8.2	7.7
	100	43.0	31.3	23.2	43.0	30.9	20.2	43.0	31.4	27.2	43.0	24.5	17.2	43.0	16.0	11.0
0.9	20	9.0	8.3	6.9	9.0	7.8	6.1	9.0	9.9	9.0	9.0	9.1	8.2	9.0	6.8	6.2
	50	17.3	13.1	11.7	17.3	12.1	9.7	17.3	16.3	15.4	17.3	12.7	12.5	17.3	7.8	7.7
	100	44.9	31.4	23.2	44.9	30.6	21.9	44.9	31.0	27.6	44.9	24.8	18.3	44.9	15.2	12.5

6. PROOFS

This section provides the proofs of the results stated earlier in the manuscript. Each subsection is devoted to one proof. Most of the results stated in this paper involve the covariance function of the processes R_A, which can be easily manipulated using the following extension of the binomial formula.

PROPOSITION 6.1. (Multinomial Formula) *Let A be a nonempty set and let $u, v \in \mathbb{R}^{|A|}$. Then*

$$\sum_{B \subset A} \left(\prod_{i \in B} u^{(i)} \right) \left(\prod_{j \in A \setminus B} v^{(j)} \right) = \prod_{i \in A} (u^{(i)} + v^{(i)}). \tag{8}$$

6.1. *Proof of Proposition* 2.1. Since for any $i \neq j$, $\mu_{i,j}(t) = P(U_i \leqslant t^{(i)}, U_j \leqslant t^{(j)}) - K^{(i)}(t^{(i)}) K^{(j)}(t^{(j)})$, $1 \leqslant i, j \leqslant p$, the property $\mu_{i,j} \equiv 0$ yields the independence of U_i and U_j. Next if $\{U_i; i \in B\}$ are independent for all $B \subset I_p$ with $|B| \leqslant k$, then this is also true for all sets $A \subset I_p$ with $|A| = k+1$, because $\mu_A \equiv 0$ implies that for all $t \in \mathbb{R}^p$,

$$\begin{aligned} 0 &= \mu_A(t) \\ &= K_p(t^A) + \sum_{B \subset A,\, B \neq A} (-1)^{|A \setminus B|} K_p(t^B) \prod_{j \in A \setminus B} K^{(j)}(t^{(j)}) \\ &= K_p(t^A) + \prod_{j \in A} K^{(j)}(t^{(j)}) \sum_{B \subset A,\, B \neq A} (-1)^{|A \setminus B|} \\ &= K_p(t^A) - \prod_{j \in A} K^{(j)}(t^{(j)}), \end{aligned}$$

proving that $\{U_i; i \in A\}$ are independent.

6.2. *Proof of Theorem* 2.1. First, define

$$\tilde{R}_{n,A}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{k \in A} [\mathbb{I}\{\varepsilon_i^{(k)} \leqslant t^{(k)}\} - K^{(k)}(t^{(k)})], \tag{9}$$

where $K^{(k)}$ denote the kth marginal distribution of ε_1. Observe that $\mu_A(t)$ is the expectation of $\tilde{R}_{n,A}(t)$. The proof of the Theorem proceeds as follows. First it will be shown that the processes $\tilde{R}_{n,A}(t)$, $A \subset I_p$ converge to the limiting processes R_A's given in the statement. Next it will be established that $\sup_t |R_{n,A}(t) - \tilde{R}_{n,A}(t)|$ converges in probability to zero.

For the asymptotic behaviour of $\tilde{R}_{n,A}$, note that

$$\tilde{R}_{n,A}(t) = \sum_{B \subset A} (-1)^{|A \setminus B|} \prod_{j \in A \setminus B} K^{(j)}(t^{(j)}) \times \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\prod_{j \in B} \mathbb{I}\{\varepsilon_i^{(j)} \leqslant t^{(j)}\} - \prod_{j \in B} K^{(j)}(t^{(j)}) \right].$$

For every fixed B, the summand in the above expression is an empirical process obtained from a sequence of independent and identically distributed random vectors and is therefore tight. Since there is only a finite number of B, the sequence of processes $(\tilde{R}_{n,A})$ is therefore tight. The convergence of the finite dimensional distributions to Gaussian limit is also easy to establish. To complete the proof one just need to verify the expression of the covariance function. For, let $A, B \subset I_p$ and let $t, s \in \mathbb{R}^p$ using representation (9) one gets

$$\operatorname{Cov}(\tilde{R}_{n,A}(s), \tilde{R}_{n,B}(t)) = \begin{cases} 0 & \text{if } A \neq B \\ \prod_{j \in A} \{K(\min(t^{(j)}, s^{(j)})) - K(t^{(j)})\, K(s^{(j)})\} & \text{if } A = B. \end{cases}$$

Finally, observe that

$$\begin{aligned} |R_{n,A}(t) - \tilde{R}_{n,A}(t)| &= \Bigg| \sum_{B \subset A,\, B \neq \varnothing} (-1)^{|B|} \prod_{k \in B} \{K_n^{(k)}(t^{(k)}) - K(t^{(k)})\} \\ &\quad \times \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{j \in A \setminus B} [\mathbb{I}\{\varepsilon_i^{(j)} \leqslant t^{(j)}\} - K^{(j)}(t^{(j)})] \Bigg| \\ &\leqslant \sum_{B \subset A,\, B \neq \varnothing} \prod_{k \in B} |K_n^{(k)}(t^{(k)}) - K^{(k)}(t^{(k)})|\, |\tilde{R}_{n,A \setminus B}(t^{A \setminus B})|, \end{aligned}$$

which goes to zero in probability by the Glivenko–Cantelli lemma and the fact that $\tilde{R}_{n,A \setminus B}$ is tight by the above arguments.

6.3. *Proof of Theorem* 2.2. Once again, define

$$\tilde{R}_{n,A}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{k \in A} [\mathbb{I}\{\varepsilon_i^{(k)} \leqslant t^{(k)}\} - K(t^{(k)})].$$

The proof proceeds exactly like the one of Theorem 2.1. First the asymptotic behaviour of $\tilde{R}_{n,A}$ is established, then it is shown that $\tilde{R}_{n,A}(x)$ and $R_{n,A}$ are asymptotically equivalent.

For the first step, let $[.]$ denote the integer part, set

$$r_n(t)=\sum_{h=1}^{n-p[n/p]}\frac{1}{\sqrt{n}}\prod_{j\in A}\left[\mathbb{I}\{\varepsilon^{(j)}_{p[n/p]+h}\leqslant t^{(j)}\}-K(t^{(j)})\right]$$

and observe that it is uniformly bounded by $p/\sqrt{n}$. Next

$$\begin{aligned}\tilde{R}_{n,A}(t)-r_n(t)&=\sum_{h=1}^{p}\frac{1}{\sqrt{n}}\sum_{i=0}^{[n/p]-1}\prod_{j\in A}\left[\mathbb{I}\{\varepsilon^{(j)}_{pi+h}\leqslant t^{(j)}\}-K(t^{(j)})\right]\\&=\sum_{h=1}^{p}\sum_{B\subset A}(-1)^{|A\backslash B|}\prod_{j\in A\backslash B}K(t^{(j)})\\&\quad\times\frac{1}{\sqrt{n}}\sum_{i=0}^{[n/p]-1}\left[\prod_{j\in B}\mathbb{I}\{\varepsilon^{(j)}_{pi+h}\leqslant t^{(j)}\}-\prod_{j\in B}K(t^{(j)})\right].\end{aligned}$$

In the above representation for every fixed h and B, the above sum over i is a p dimensional empirical process of a sequence of i.i.d random vectors and is therefore tight. Since there is only a finite number of B's and h's, the sequence $(\tilde{R}_{n,A})$ is therefore tight. To complete the proof one must consider the finite dimensional distribution of $\tilde{R}_{n,A}$. First, note that for all $A\neq B\in\mathscr{A}_p$, $\mathrm{Cov}(\tilde{R}_{n,A}(t),\tilde{R}_{n,B}(s))=0$ and

$$\begin{aligned}\mathrm{Cov}(\tilde{R}_{n,A}(t),\tilde{R}_{n,A}(s))&=\prod_{j\in A}E[\mathbb{I}\{\varepsilon^{(j)}_1\leqslant t^{(j)}\}-K(t^{(j)})]\\&\quad\times[\mathbb{I}\{\varepsilon^{(j)}_1\leqslant s^{(j)}\}-K(s^{(j)})]\\&=\prod_{j\in A}\{[(\min(t^{(j)},s^{(j)}))-K(t^{(j)})\,K(s^{(j)})\}.\end{aligned}$$

Moreover, for any fixed t_1, ..., t_k the central limit theorem for p dependent sequence (Billingsley, 1968) applies and yields the desired Gaussian limit.

For the second step, note that the same argument as that given in the previous proof yields

$$\begin{aligned}|R_{n,A}(t)-\tilde{R}_{n,A}(t)|\leqslant&\sum_{B\subset A,\,B\neq\varnothing}\prod_{k\in B}|K_{n,1}(t^{(k)})-K(t^{(k)})|\\&\times|\tilde{R}_{n,A\backslash B}(t^{A\backslash B})|+\frac{p}{\sqrt{n}}\end{aligned}$$

which goes to zero in probability by the Glivenko–Cantelli lemma and since $\tilde{R}_{n,A}$ is tight by the first step.

6.4. *Proof of Corollary* 2.2. Once more, redefine

$$\tilde{R}_{n,A}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{k \in A} [\mathbb{I}\{e_i^{(k)} \leqslant t^{(k)}\} - K(t^{(k)})].$$

To prove this corollary one needs to show that $\tilde{R}_{n,A}$ converges to the specified limit and that the processes $R_{n,A}$ and $\tilde{R}_{n,A}$ are asymptotically equivalent, that is, $\sup_t |R_{n,A}(t) - \tilde{R}_{n,A}(t)|$ converges in probability to zero as n goes to infinity.

The convergence of $\tilde{R}_{n,A}$ to the specified limit is a consequence of Theorem 2.3 and the representation

$$\tilde{R}_{n,A}(t) = \sum_{B \subset A} (-1)^{|A \backslash B|} \tilde{\beta}^*_{n,p}(t^B) \prod_{i \in A \backslash B} K(t^{(i)}).$$

For the asymptotic equivalence of $R_{n,A}$ and $\tilde{R}_{n,A}$, the same argument as in the proof of Theorem 2.1 yields

$$\begin{aligned}
&\sup_t |R_{n,A}(t) - \tilde{R}_{n,A}(t)| \\
&\leqslant \sum_{B \subset A,\, B \neq \varnothing} \prod_{k \in B} |\tilde{K}_{n,1}(t^{(k)}) - K(t^{(k)})| \, |\tilde{R}_{n,A \backslash B}(t^{A \backslash B})| + \frac{p}{\sqrt{n}} \\
&= \frac{1}{n^{|B|/2}} \sum_{B \subset A,\, B \neq \varnothing} \prod_{k \in B} |\tilde{\beta}^*_{n,1}(t^{(k)})| \, |\tilde{R}_{n,A \backslash B}(t^{A \backslash B})| + \frac{p}{\sqrt{n}}
\end{aligned}$$

which goes to zero in probability, since by Theorem 2.3, $\tilde{\beta}^*_{n,p}$ is tight and by the above argument $\tilde{R}_{n,A \backslash B}$ is tight.

6.5. *Proof of Proposition* 2.2. First it shall be shown that the hypotheses of Theorem 2.3 are satisfied. With the regression setting, one easily verifies that Condition (R1) is verified whenever $E(\|Z\|)$ is finite. Condition (R3) holds if U admits a continuous bounded density k. To show the first part of (R2) write

$$\begin{aligned}
&\alpha_{q,n,j,\psi \circ g}(s,t) \\
&\simeq \frac{1}{\sqrt{n}} \sum_{i=0}^{[(n/p]-1} \Big[\psi(g(X_{pi+q+j})) \\
&\quad \times \mathbb{I}\{\varepsilon_{pi+q}^{(j)} \leqslant t^{(j)} + sr(X_{pi+q+j})\} \prod_{k \neq j} \mathbb{I}\{\varepsilon_{pi+q}^{(k)} \leqslant t^{(k)}\} \\
&\quad - E\Big\{\psi(g(X_j)) \, \mathbb{I}\{\varepsilon_1^{(j)} \leqslant t^{(j)} + sr(X_j)\} \prod_{k \neq j} \mathbb{I}\{\varepsilon_1^{(k)} \leqslant t^{(k)}\}\Big\}\Big],
\end{aligned}$$

and observe that

$$|\alpha_{n,\,j,\,\psi\circ g}(s/\sqrt{n},\,t)-\alpha_{n,\,j,\,\psi\circ g}(0,\,t)| \leqslant \sum_{q=0}^{p-1} |\alpha_{q,\,n,\,j,\,\psi\circ g}(s/\sqrt{n},\,t)-\alpha_{q,\,n,\,j,\,\psi\circ g}(0,\,t)|.$$

That is, (3) will follow if

$$\sup_{t\in C} |\alpha_{q,\,n,\,j,\,\psi\circ g}(s/\sqrt{n},\,t)-\alpha_{q,\,n,\,j,\,\psi\circ g}(0,\,t)|$$

goes to zero in probability for each $q=0, ..., p-1$. Since $g(Y, Z)=g(Z)$ is independent of U and because R1 and R3 are satisfied the above follows from Lemma 7.2 of Ghoudi and Rémillard (1998b) whenever $E(\|Z\|^2)$ is finite. The second part of (R2) is quite easy if a_n and b_n are the least square estimates of a and b. To complete the proof of this proposition it suffices to show that the representation of $\tilde{R}_A$ given in Corollary 2.2 reduces to the R_A's defined in Theorem 2.2. Using the definition of $\tilde{\beta}_p$ one obtains

$$\begin{aligned}
\tilde{R}_A(t) &= \sum_{B\subset A} (-1)^{|A\backslash B|}\, \tilde{\beta}_p(t^B) \prod_{i\in A\backslash B} K(t^{(i)}) \\
&= \sum_{B\subset A} (-1)^{|A\backslash B|}\, \beta_p(t^B) \prod_{i\in A\backslash B} K(t^{(i)}) \\
&\quad - \sum_{B\subset A} (-1)^{|A\backslash B|} \sum_{j=1}^{p} \mu_j(t^B, IH) \prod_{i\in A\backslash B} K(t^{(i)}) \\
&= R_A(t) + \sum_{j=1}^{p} \{A+B'E(Z)\}\, k(t^{(j)}) \\
&\quad \times \prod_{i\in A\backslash\{j\}} K(t^{(i)}) \sum_{B\subset A\backslash\{j\}} (-1)^{|A|-1-|B|} \\
&= R_A(t).
\end{aligned}$$

6.6. *Proof of Proposition* 3.3. Note that $D_A(s)-W(s_A)=\sum_{B\subset A,\, B\neq A} (-1)^{|A\backslash B|}\, W(s_B) \prod_{j\in A\backslash B} s^{(j)}$, which is $\mathscr{K}_A$-measurable. The rest of the proof is achieved in two steps. In the first Step, one must prove that the process $D_A(s)$ is orthogonal to $W(t_C)$, for any $t\in[0,1]^p$ and for any $C\subset A$, $C\neq A$. In the second Step one shows that D_A and $\tilde{D}_A$ have the same covariance functions.

Step 1. Let C be a subset of $\{1, ..., p\}$.

$$
\begin{aligned}
E\{D_A(s)\, W(t_C)\} &= \sum_{B \subset A} (-1)^{|A \setminus B|} \prod_{i \in A \setminus B} s^{(i)} \\
&\quad \times \prod_{i \in B \cap C} (s^{(i)} \wedge t^{(i)}) \prod_{i \in B \setminus C} s^{(i)} \prod_{i \in C \setminus B} t^{(i)} \\
&= \sum_{A_0 \subset A \cap C} \sum_{A_1 \subset A \setminus C} (-1)^{|A| - |A_0| - |A_1|} \prod_{i \in A \setminus A_0} s^{(i)} \\
&\quad \times \prod_{i \in A_0} s^{(i)} \wedge t^{(i)} \prod_{i \in C \setminus A_0} t^{(i)} \\
&= \sum_{A_0 \subset A \cap C} (-1)^{|C \setminus A_0|} \prod_{i \in A \setminus A_0} s^{(i)} \prod_{i \in A_0} s^{(i)} \wedge t^{(i)} \\
&\quad \times \prod_{i \in C \setminus A_0} t^{(i)} \sum_{A_1 \subset A \setminus C} (-1)^{|A \setminus C| - |A_1|} \\
&= \begin{cases} \prod_{i \in A} (s^{(i)} \wedge t^{(i)} - s^{(i)} t^{(i)}) \prod_{i \in C \setminus A} t^{(i)}, & A \subset C \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

It follows that $D_A(s) = W(s_A) - E\{ W(s_A) \mid \mathscr{K}_A\}$.

Step 2. This step is dedicated to the computation of the covariance between D_A and D_B. Straightforward computations show that

$$
\mathrm{Cov}\{D_A(s), D_B(t)\} = \sum_{B_0 \subset B} (-1)^{|B \setminus B_0|} \mathrm{Cov}\{D_A(s), W(t_{B_0})\} \textstyle\prod_{i \in B \setminus B_0} t^{(i)}. \tag{10}
$$

From Step 1, this is equal to zero unless $A \subset B$, Inverting the roles of A and B in the above implies that $\mathrm{Cov}\{D_A(s), D_B(t)\} = 0$ if $A \neq B$. For $A = B$, Equation (10) reduces to

$$
\begin{aligned}
\mathrm{Cov}\{D_A(s), D_A(t)\} &= \mathrm{Cov}\{D_A(s), W(t_A)\} \\
&= \prod_{i \in A} (s^{(i)} \wedge t^{(i)} - s^{(i)} t^{(i)}) = C_A(s, t).
\end{aligned}
$$

6.7. *Proof of Lemma* 4.1. First notice that using Donsker's invariance principle (Donsker, 1952, or Billingsley, 1968), $\int R^2_{n,A}(t)\, dK_p(t)$ converges in distribution to $\int R^2_A(t)\, dK_p(t) = \xi_{|A|}$. Next, using the fact that if K is continuous $T_{n,A}$ is distribution free and repeating the argument of the proof of the Lemma in Section 2 of Kiefer (1959) one concludes that

$$\int R_{n,A}^2(t)\, dK_{n,p}(t) - \int R_{n,A}^2(t)\, dK_p(t)$$

converges in probability to zero.

ACKNOWLEDGMENTS

The authors are supported in part by the Fonds Institutionnel de Recherche, Université du Québec à Trois-Rivières, the Fonds pour la formation de chercheurs et l'aide à la recherche du Gouvernement du Québec, and by the Natural Sciences and Engineering Research Council of Canada, Grants OGP0184065, OGP0005724, and OGP0042137.

REFERENCES

1. M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions," National Bureau of Standards, Washington, DC, 1964.
2. P. Barbe, C. Genest, K. Ghoudi, and B. Remillard, On Kendall's process, *J. Multivariate Anal.* **58** (1996), 197–229.
3. P. Billingsley, "Convergence of Probability Measures," Wiley, New York, 1968.
4. J. R. Blum, J. Kiefer, and M. Rosenblatt, Distribution free tests of independence based on the sample distribution function, *Ann. Math. Statist.* **32** (1961), 485–498.
5. H. W. Block and Z. Fang, A multivariate extension of Hoeffding's lemma, *Ann. Probab.* **16** (1988), 1803–1820.
6. P. J. Brockwell and R. A. Davis, "Time Series: Theory and Methods," Springer-Verlag, New York, 1991.
7. D. S. Cotterill and M. Csörgő, On the limiting distribution of and critical values for the multivariate Cramér–von Mises Statistic, *Ann. Statist.* **10** (1982), 233–244.
8. D. S. Cotterill and M. Csörgő, On the limiting distribution and critical values for the Hoeffding, Blum, Kiefer, Rosenblatt independence criterion, *Statist. & Decisions* **3** (1985), 1–48.
9. P. Deheuvels and G. V. Martynov, Cramér–von Mises type tests with applications to tests of independence for multivariate extreme-value distributions, *Comm. Statist. Theory Methods* **25** (1996), 871–908.
10. M. A. Delgado, Testing serial independence using the sample distribution function, *J. Time Ser. Anal.* **11** (1996), 271–285.
11. M. D. Donsker, Justification and extension of Doob's heuristic approach to the Kolmogorov–Smirnov theorems, *Ann. Math. Statist.* **23** (1952), 277–281.
12. K. Ghoudi and B. Remillard, Empirical processes based on pseudo-observations, *in* "Proceedings of the International Conference on Asymptotic Methods in Statistics" (B. Szyskowitz, Ed.), Elsevier, Amsterdam, 1998.
13. K. Ghoudi and B. Remillard, "Empirical Processes Based on Pseudo-observations. II. The Multivariate Case," Rapport Technique du Laboratoire de Recherche en Probabilités et Statistique, U. Q. T. R., Vol. 9, 1998.

14. W. Hoeffding, A nonparametric test of independence, *Ann. Math. Statist.* **19** (1948), 546–557.
15. Y. Hong, Testing for pairwise serial independence via the empirical distribution function, *J. Roy. Statist. Soc. Ser. B* **60** (1998), 429–453.
16. Y. Hong, Generalized spectral tests for serial dependence, *J. Roy. Statist. Soc. Ser. B* **62** (2000), 557–574.
17. J. P. Imhof, Computing the distribution of quadratic forms in normal variables, *Biometrika* **48** (1961), 419–426.
18. J. D. Jobson, "Applied Multivariate Data Analysis. Vol. I. Regression and Experimental Design," Springer-Verlag, New York, 1991.
19. M. G. Kendall, The analysis of economic time-series. Part I. Prices, *J. Statist. Soc. A* **116** (1953), 11–25.
20. J. Kiefer, K-sample analogues of the Kolmogorov–Smirnov and Cramér–von Mises tests, *Ann. Math. Statist.* **30** (1959), 420–447.
21. R. J. Kulperger and R. A. Lockhart, Tests of independence in time series, *J. Time Ser. Anal.* **19** (1998), 165–185.
22. G. Marsaglia and A. Zaman, "The Kiss Generator," technical report, Florida State University, 1995.
23. H. J. Skaug and D. Tjøstheim, A nonparametric test of serial independence based on the empirical distribution function, *Biometrika* **80** (1993), 591–602.
24. M. A. Stephens, *in* "Goodness of Fit Techniques" (R. B. D'Agostino and M. A. Stephens, Eds.), Dekker, New York, 1986.