

## LINEAR SERIAL RANK TESTS FOR RANDOMNESS AGAINST ARMA ALTERNATIVES

BY MARC HALLIN,<sup>1</sup> JEAN-FRANCOIS INGENBLEEK AND MADAN L. PURI<sup>2</sup>

*Université Libre de Bruxelles and Indiana University*

In this paper we introduce a class of linear serial rank statistics for the problem of testing white noise against alternatives of ARMA serial dependence. The asymptotic normality of the proposed statistics is established, both under the null as well as alternative hypotheses, using LeCam's notion of contiguity. The efficiency properties of the proposed statistics are investigated, and an explicit formulation of the asymptotically most efficient score-generating functions is provided. Finally, we study the asymptotic relative efficiency of the proposed procedures with respect to their normal theory counterparts based on sample autocorrelations.

**1. Introduction.** Nonparametric methods have been developed for the analysis of univariate, and, much later, for the analysis of multivariate observations as a reaction against the distributional assumptions (mostly, that of normality) on which much of the classical inference relies.

The need for nonparametric procedures is even stronger in the area of (univariate and multivariate) time-series analysis: indeed, while there are several quite tractable inference procedures for independently distributed nonnormal observations, almost nothing exists, in time series, except for methods based on the normal theory likelihood ratio approach.

Although many rank-based procedures have been developed in (non-time-series context) for testing randomness against different types of trend alternatives (see, e.g., Dufour et al. (1982)), one of the most significant problems for practical applications, viz. the problem of testing randomness against serial dependence using rank methods has not received much attention—apart from some scattered results here and there. Of course, even if not always specifically time series-oriented, some of the “historical” nonparametric tests, such as the *run test* and the *turning point test* can be used for testing the hypothesis of an independent time series against serial dependence—in fact, the history of the problem can be traced back to Hotelling and Pabst (1936), and Wald and Wolfowitz (1943) who proposed tests based on serial versions of Spearman's correlation coefficient referred to below as *Spearman's autocorrelation coefficient* (see Section 5.3). Jogdeo (1968)—also in a non-time-series context—studied the distribution (under independence) of a class of statistics which is not unrelated with our  $S_n$  statistics (to be defined below); however, because of the restrictions he puts on the score functions, his statistics are not adapted to time-series situations.

---

Received February 1984; revised March 1985.

<sup>1</sup> Research supported by a grant of the Belgium F.N.R.S.

<sup>2</sup> Research supported by National Science Foundation Grant MCS 8301409.

AMS 1980 subject classifications. Primary 62M10, 62G10.

*Key words and phrases.* Linear serial rank statistics, autoregressive, moving average, ARMA models, asymptotic relative efficiency.

Moreover, in none of these early papers is any particular alternative of serial dependence considered, nor any optimality question addressed.

The first attempt to investigate the power of serial rank procedures against specific alternatives of serial dependence is due to Knoke (1977), who studies the asymptotic relative efficiency of several tests based on rank statistics (e.g., *Spearman's first-order autocorrelation coefficient* or the *turning point statistic*) with respect to the first-order sample autocorrelation coefficient, for autoregressive alternatives of order one. This, in some sense, was the first step towards the application of rank methods in time-series analysis problems—but the statistics that he studies are not new, and are not specifically devised for the alternative considered. Recently, in a rather restrictive set-up, Gupta and Govindarajulu (1980) considered the first-order autoregressive and moving average *normal* alternatives, and derived a locally most powerful rank statistic which is a particular case of our *linear serial rank statistic* (cf. (1.1)). Aiyar (1981) proposed, on heuristic considerations, a test (still for first-order alternatives) based on a statistic which is also a particular case of an extended *van der Waerden statistic* which we arrive at in (5.4). Bartels (1982) introduces a rank version of *von Neumann's ratio test*, still for first-order autoregressive alternatives, which he shows to be more efficient than the run test and parametric von Neumann test (under normality).

In a somewhat different direction, a highly systematic and theoretically-based approach is provided in Bell et al. (1970), where the Pitman functions generating the most powerful distribution-free tests are derived for a null hypothesis of randomness against the alternative of a process with stationary independent increments and several other alternatives of dependence.

More recently, letting  $Z_t = X_t X_{t-1}$ , where  $(X_1 \cdots X_n)$  denotes the observed series, Dufour (1982) applied to  $Z_t$  some well-known procedures, such as the *sign*, *Wilcoxon*, *signed rank* and *van der Waerden* tests for symmetry about zero. Although the statistics he introduces are new, they still are an adaptation of existing ones: no optimality considerations are made and whether multivariate extensions are possible is unknown. A sign test of the type studied in Dufour (1982) was also considered by Govindarajulu and Dwass (1983). An overall review of some of these procedures is given in Govindarajulu (1983).

In spite of the growing interest in time-series problems, there exists thus no systematic, coherent and unified rank order theory, on the model of the one developed in Hájek and Šidák (1967) and Puri and Sen (1971) for independent observations; the few available results constitute only a very incomplete and piecemeal approach, and the subject remains, to a large extent, unexplored.

The purpose of this paper is to undertake a first systematic time-series oriented study of this important problem of testing for randomness against the types of serial dependence usually considered in time-series analysis (autoregressive, moving average or mixed ARMA dependence). We propose to do so by considering the statistics of the form

$$(1.1) \quad S_n = (n - p)^{-1} \sum_{t=p+1}^n a_n(R_t^{(n)}, R_{t-1}^{(n)}, \dots, R_{t-p}^{(n)}),$$

where  $a_n(\dots)$  is some given *score function* and  $R_t^{(n)}$  is the rank of the observation

made at time  $t$  in an observed series of length  $n$ . These statistics are what we call *linear serial rank statistics*. There are many reasons for considering this class of statistics. For example,

- (a) it is intuitively a quite natural idea to account for serial dependence (of order 1 through  $p$ ) by considering the ranks of successive  $(p + 1)$ -tuples of observations;
- (b) special cases of the score functions  $a_n$  give traditional test statistics, such as the *run statistic* (with respect to the median)

$$a_n(i_1, i_2) = \begin{cases} 1 & \text{if } (2i_1 - n - 1)(2i_2 - n - 1) < 0 \\ 0 & \text{if } (2i_1 - n - 1)(2i_2 - n - 1) \geq 0, \end{cases}$$

the *turning point* statistic

$$a_n(i_1, i_2, i_3) = \begin{cases} 1 & \text{if } i_1 > i_2 < i_3 \\ 1 & \text{if } i_1 < i_2 > i_3 \\ 0 & \text{elsewhere,} \end{cases}$$

*Spearman's rank correlation coefficient of order  $p$*  (up to additive and multiplicative constants)

$$a_n(i_1, i_2, \dots, i_{p+1}) = i_1 i_{p+1} / (n + 1)^2, \text{ etc.};$$

- (c) it can be shown that the locally most powerful rank statistic for testing randomness against autocorrelation (AR(1)) alternatives belongs to this class of statistics (cf. Ingenbleek (1980));
- (d) multivariate versions can be considered (hence, among others, multivariate extensions of the run, turning point, . . . statistics); this latter argument is very important, since the need for results is particularly strong in the multivariate case.

We mainly develop here the asymptotic theory of the class of tests based on linear serial rank statistics, and our approach largely relies on Hájek's projection method and LeCam's results on *contiguous sequences* of hypotheses.

Denote by  $H_0^{(n)}$  the null hypothesis under which the observed series (of length  $n$ ) is *white noise*, and by  $H_1^{(n)}$  the alternative under which it is generated by an ARMA( $p_1, p_2$ ) model of the form

$$(1.2) \quad X_t - n^{-1/2} \sum_{i=1}^{p_1} a_i X_{t-i} = \varepsilon_t + n^{-1/2} \sum_{i=1}^{p_2} b_i \varepsilon_{t-i}.$$

Under fairly mild assumptions (Section 2), we show in Section 3 (Proposition 3.1) that the sequence of hypotheses  $H_1^{(n)}$  is contiguous to  $H_0^{(n)}$ .

This result, together with a lemma by LeCam, allows us to derive the asymptotic distribution of  $S_n$  under the alternatives  $H_1^{(n)}$  from the knowledge of the asymptotic joint distribution of  $S_n$  and the log-likelihood ratio under  $H_0^{(n)}$ ; these distributions are obtained (Section 4) by introducing an adequate *U-statistic*.

The last section (Section 5) is devoted to *efficiency* considerations. We investigate the *asymptotic relative efficiency* (ARE) of linear serial rank statistics and provide (Proposition 5.1) an explicit form of the asymptotically most efficient one—hence, for particular density types (normal, logistic, double exponential, . . .), time-series versions of Wilcoxon's, van der Waerden's, median, . . . tests.

Next, we (Proposition 5.3) compare our nonparametric statistics with their classical (normal theory) competitors, based on autocorrelation coefficients. This comparison seems to be quite favorable to serial rank statistics: the ARE of the “optimal” linear serial rank test with respect to that based on its “optimal” normal theory competitor depends on the Fisher information associated with the white noise density type, taking value one in the case of normal distribution,  $\pi^2/9$  in the case of logistic distribution and two in the case of double exponential distribution. Finally, a more detailed study of the first-order case (ARMA(1, 1), AR(1) or MA(1) alternatives) is given, the “optimal” serial rank statistic (in the sense of Proposition 5.1) having then the important additional property that it is *uniformly* “optimal” (over the whole range of ARMA(1, 1), AR(1) and MA(1) alternatives, and within the class of linear serial rank statistics).

**2. Notation and basic assumptions.** Throughout the paper, we denote by  $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ ,  $\mathbb{R}$  and  $\mathbb{C}$ , respectively, the sets of integer, real and complex numbers.

Let  $\{\varepsilon_t; t \in \mathbb{Z}\}$  be a discrete-time stationary *white noise*, i.e., a sequence of independent and identically distributed random variables with means  $E[\varepsilon_t] = 0$ ,  $t \in \mathbb{Z}$ ; assume that it has a density  $f(x)$ , and that the following conditions are satisfied.

- (i)  $\varepsilon_t$  has finite moments up to the third order; denote its variance by  $\sigma^2$ .
- (ii)  $f(x)$  is a.e. derivable, and its derivative  $f'(x)$  satisfies  $\int_{-\infty}^{+\infty} |f'(x)| dx < \infty$ .
- (iii)  $f(x)$  has finite Fisher’s information  $I(f)$ , i.e.,  $f(x)$  is *absolutely continuous on finite intervals*, and

$$0 < I(f) = \int_{-\infty}^{+\infty} \left( \frac{f'(x)}{f(x)} \right)^2 f(x) dx < \infty$$

(as usual,  $I(f)$  is the Fisher information related to the location parameter family  $\{f_\theta(x) = f(x - \theta) \mid \theta \in \mathbb{R}\}$ ).

- (iv) Let  $F(x)$  be  $\varepsilon_t$ ’s distribution function, and  $F^{-1}(u) = \inf\{x \mid F(x) \geq u\}$ ,  $0 < u < 1$ . Put

$$\phi(F^{-1}(u)) = -f'(F^{-1}(u))/f(F^{-1}(u)) \quad 0 < u < 1.$$

This function can also be written a.e. as  $\phi(x) = -f'(x)/f(x)$ ,  $x \in \mathbb{R}$ . Assume  $\phi(x)$  is a.e. derivable, and its derivative  $\phi'(x)$  satisfies (a.e.) a Lipschitz condition  $|\phi'(x) - \phi'(y)| < A|x - y|$ .

These conditions are satisfied by most of the densities considered in the theory of rank tests. However, they do not include Cauchy and stable distributions.

Under these conditions, we have (Hájek and Šidák (1967) Chapter I)

$$(2.1) \quad \int_{-\infty}^{+\infty} \phi(x)f(x) dx = 0$$

and

$$(2.2) \quad \int_{-\infty}^{+\infty} \phi^2(x)f(x) dx = I(f).$$

It is also easy to check that

$$(2.3) \quad \int_{-\infty}^{+\infty} x\phi(x)f(x) dx = 1$$

and

$$(2.4) \quad \int_{-\infty}^{+\infty} \phi'(x)f(x) dx = I(f).$$

Note also that  $\sigma^2 I(f)$  is independent of the scale transformation (i.e., if we put  $f_\sigma(x) = (1/\sigma)f_1(x/\sigma)$ , then  $\sigma^2 I(f_\sigma) = I(f_1)$ ).

### 3. Asymptotic distribution of likelihood ratios.

3.1. *The hypotheses.* Let  $a_1, \dots, a_{p_1}, b_1, \dots, b_{p_2}$  be an arbitrary  $(p_1 + p_2)$ -tuple of real numbers, and consider the sequence

$$(3.1) \quad X_t^{(n)} - n^{-1/2} \sum_{i=1}^{p_1} a_i X_{t-i}^{(n)} = \varepsilon_t + n^{-1/2} \sum_{i=1}^{p_2} b_i \varepsilon_{t-i}, \quad t \in \mathbb{Z}, \quad n = 1, 2, \dots$$

of stochastic difference equations. For  $n$  sufficiently large, all the roots of the characteristic equation

$$z^{p_1} - n^{-1/2} \sum_{i=1}^{p_1} a_i z^{p_1-i} = 0, \quad z \in \mathbb{C}$$

lie inside the unit-circle, and (3.1) (see, for example, Wold, 1954, page 99) generates a sequence of stationary processes  $\{X_t^{(n)}; t \in \mathbb{Z}\}$ . Denote by  $\mathbf{x}^{(n)} = (x_1^{(n)}, \dots, x_n^{(n)})$  an observed realization of  $\mathbf{X}^{(n)} = (X_1^{(n)}, \dots, X_n^{(n)})$ .

If  $a_1 = a_2 = \dots = a_{p_1} = b_1 = b_2 = \dots = b_{p_2} = 0$ , the processes  $\{X_t^{(n)}\}$  of course all coincide with the generating white noise process  $\{\varepsilon_t\}$ . The likelihood function of the observation  $\mathbf{x}^{(n)}$  is then

$$(3.2) \quad \mathcal{L}_n^0(\mathbf{x}^{(n)}) = \prod_{t=1}^n f(x_t^{(n)}).$$

Denote by  $H_0^{(n)}$  the sequence of simple (null) hypotheses consisting of these densities.

If (at least)  $a_{p_1}$  and  $b_{p_2}$  are both different from zero, the processes  $\{X_t^{(n)}\}$  are stationary *autoregressive-moving average processes* of order  $p_1$  and  $p_2$  (shortly, ARMA( $p_1, p_2$ ) processes). Denote then by  $G_{ab}^{(n)}(x_{t+1}, \dots, x_{t+p_1})$  the distribution function of  $p_1$  successive values of  $\{X_t^{(n)}\}$ . The likelihood function of  $\mathbf{x}^{(n)}$  is

$$(3.3) \quad \begin{aligned} \mathcal{L}_n^1(\mathbf{x}^{(n)}) = & \int \prod_{t=1}^n f(x_t^{(n)}) - n^{-1/2} \sum_{i=1}^{p_1} a_i x_{t-i}^{(n)} + \sum_{u=1}^{t-1} g_u^{(n)}(x_{t-u}^{(n)} - n^{-1/2} \sum_{i=1}^{p_1} a_i x_{t-u-i}^{(n)}) \\ & + \sum_{u=t}^{t+p_2-1} g_u^{(n)}(e_{t-u}) dG_{ab}^{(n)}(x_{-p_1+1}, \dots, x_0) f(e_{-p_2+1}, \dots, f(e_0) de_{-p_2+1}, \dots, de_0, \end{aligned}$$

where the coefficients  $g_u^{(n)}$  are the *Green's functions* associated with the moving average difference operator ( $L$  denotes the lag operator)  $1 + n^{-1/2} \sum_{i=1}^{p_2} b_i L^i$  (cf. Appendix 1). Denote by  $H_1^{(n)}$  the corresponding sequence of hypotheses:  $H_1^{(n)}$  will be our sequence of alternatives (ARMA dependence);  $H_0^{(n)}$  is the sequence of null hypotheses (randomness).

Particular cases of  $H_1^{(n)}$  are

- (a) the autoregressive case of order  $p_1$  ( $b_1 = b_2 = \dots = b_{p_2} = 0; a_{p_1} \neq 0$ )
- (b) the moving average case of order  $p_2$  ( $a_1 = a_2 = \dots = a_{p_1} = 0; b_{p_2} \neq 0$ ).

The corresponding likelihood functions can be obtained by letting, respectively,  $g_u^{(n)} = 0 \forall n, u = 1, 2, \dots$  or  $a_i = 0, i = 1, \dots, p_1$  in (3.3).

As we shall see, however, the asymptotic behaviour of the likelihood function  $\ell_n^1(\mathbf{x}^{(n)})$  is quite similar under the three types of dependence considered (mixed ARMA, pure AR or pure MA).

3.2. *Contiguity.* In what follows, we shall omit the superscripts  $(n)$ , and write  $\mathbf{x}, x_t, \mathbf{X}$  and  $X_t$  for  $\mathbf{x}^{(n)}, x_t^{(n)}, \mathbf{X}^{(n)}$  and  $X_t^{(n)}$  respectively. Consider the likelihood ratio

$$(3.4) \quad L_n(\mathbf{x}) = \begin{cases} \ell_n^1(\mathbf{x})/\ell_n^0(\mathbf{x}) & \text{if } \ell_n^0(\mathbf{x}) > 0 \\ 1 & \text{if } \ell_n^1(\mathbf{x}) = \ell_n^0(\mathbf{x}) = 0 \\ \infty & \text{if } \ell_n^1(\mathbf{x}) > \ell_n^0(\mathbf{x}) = 0. \end{cases}$$

It follows from LeCam’s first lemma (we adopt here Hájek and Šidák’s (1967) terminology) that, in order to establish that a sequence of hypotheses  $H_1^{(n)}$  is contiguous to a sequence  $H_0^{(n)}$ , it is sufficient to show that  $\log L_n(\mathbf{X})$  is asymptotically normal (under  $H_0^{(n)}$ ), with mean  $-d^2/2$  and variance  $d^2$ .

The following proposition therefore implies that  $H_1^{(n)}$  is contiguous to  $H_0^{(n)}$ .

PROPOSITION 3.1. Under  $H_0^{(n)}$ ,  $\log L_n(X_1, \dots, X_n) = \mathcal{L}_n^0(X_1, \dots, X_n) - d^2/2 + o_p$ , where

$$(3.5) \quad \mathcal{L}_n^0(\mathbf{X}) = n^{-1/2} \sum_{t=p+1}^n \phi(X_t) \sum_{i=1}^p d_i X_{t-i}$$

$$(3.6) \quad d_i = \begin{cases} a_i + b_i & 1 \leq i \leq \min(p_1, p_2) \\ a_i & p_2 < i \leq p_1 \text{ if } p_2 < p_1 \\ b_i & p_1 < i \leq p_2 \text{ if } p_1 < p_2 \end{cases}$$

$$(3.7) \quad p = \max(p_1, p_2) \quad \text{and} \quad d^2 = \sum_{i=1}^p d_i^2 \sigma^2 I(f).$$

Moreover,  $\mathcal{L}_n^0(\mathbf{X})$  is asymptotically normal, with mean zero and variance  $d^2$ .

PROOF. See Appendix 2.

The form of this asymptotic distribution shows that, for  $n$  sufficiently large, there will be little difference, from a statistical point of view, between AR, MA and ARMA models; as an example, the samples  $\mathbf{x}^{(n)}$  generated by the models

$$x_t - n^{-1/2} \sum_{i=1}^p a_i X_{t-i} = \varepsilon_t, \quad X_t = \varepsilon_t + n^{-1/2} \sum_{i=1}^p a_i \varepsilon_{t-i}, \\ X_t - n^{-1/2} \sum_{i=1}^p a_i X_{t-i}/2 = \varepsilon_t + n^{-1/2} \sum_{i=1}^p a_i \varepsilon_{t-i}/2$$

have the same asymptotic behaviour.

**4. Asymptotic distribution of linear serial rank statistics.** In this section, we study the asymptotic distribution under  $H_1^{(n)}$  of the linear serial rank statistics  $S_n$  defined in (1.1). To this end, we derive the joint asymptotic normality of

$$\begin{pmatrix} n^{1/2}(S_n - m_n) \\ \log L_n \end{pmatrix}$$

under  $H_0^{(n)}$ , where

$$m_n = E[S_n | H_0^{(n)}] = \frac{1}{n(n-1) \cdots (n-p)} \sum_{1 \leq i_1 \neq \dots \neq i_{p+1} \leq n} \cdots \sum a_n(i_1, \dots, i_{p+1}).$$

The asymptotic normality under  $H_1^{(n)}$  will then follow from LeCam's third lemma (Proposition 4.2).

Put  $U_t = F(X_t)$  and  $\underline{U} = (U_1, \dots, U_n)$ . We assume that the score functions  $a_n(\dots)$  are such that there exists a function  $J = J(v_{p+1}, v_p, \dots, v_1)$ , defined over  $[0, 1]^{p+1}$ , such that

$$(4.1) \quad 0 < \int_{[0,1]^{p+1}} J^2(v_{p+1} \cdots v_1) dv_{p+1} \cdots dv_1 < \infty$$

and

$$(4.2) \quad \lim_{n \rightarrow \infty} E[(J(U_{p+1}, \dots, U_1) - a_n(R_{p+1}^{(n)}, \dots, R_1^{(n)}))^2 | H_0^{(n)}] = 0$$

(this assumption is satisfied most of the time when  $a_n$  is of the form

$$a_n(i_1, i_2, \dots, i_{p+1}) = J(i_1/(n+1), i_2/(n+1), \dots, i_{p+1}/(n+1)).$$

Such a function  $J$  will be called a *score-generating function* (associated with the serial rank statistic  $S_n$ ).

We proceed as follows. First (Section 4.1), we establish the asymptotic equivalence of  $(n-p)^{1/2}(S_n - m_n)$  with  $\mathcal{S}_n - \mathcal{E}_n$ , where

$$(4.3) \quad \mathcal{S}_n(\mathbf{X}) = (n-p)^{-1/2} \sum_{t=p+1}^n J(F(X_t), F(X_{t-1}), \dots, F(X_{t-p}))$$

and

$$(4.4) \quad \mathcal{E}_n(\mathbf{X}) = \frac{(n-p)^{1/2}}{n(n-1) \cdots (n-p)} \sum_{1 \leq t_1 \neq \dots \neq t_{p+1} \leq n} \cdots \sum J(F(X_{t_1}) \cdots F(X_{t_{p+1}}));$$

then (Section 4.2) we show that  $n^{-1/2}(\mathcal{S}_n - \mathcal{E}_n)$  and  $n^{-1/2}\mathcal{L}_n^0$  are themselves asymptotically equivalent to  $U$ -statistics. An appropriate convergence theorem for  $U$ -statistics finally establishes (Section 4.3) the desired result.

**4.1. Asymptotic equivalence of  $(n-p)^{1/2}(S_n - m_n)$  and  $(\mathcal{S}_n - \mathcal{E}_n)$  (under  $H_0^{(n)}$ ).** Let

$$(4.5) \quad \Delta_n(\mathbf{X}) = (n-p)^{1/2}(S_n(\mathbf{X}) - m_n) - (\mathcal{S}_n(\mathbf{X}) - \mathcal{E}_n(\mathbf{X})).$$

Then it suffices to show that  $\lim_{n \rightarrow \infty} E[\Delta_n^2(\mathbf{X})] = 0$ .

First, note that

$$(n - p)^{1/2}S_n - \mathcal{L}_n = (n - p)^{-1/2} \sum_{t=p+1}^n [a_n(R_t^{(n)}, \dots, R_{t-p}^{(n)}) - J(U_{(R_t)}, \dots, U_{(R_{t-p})})]$$

can be written as

$$(n - p)^{1/2} \sum_{t=p+1}^n \alpha_n^{U_{(\cdot)}}(R_t^{(n)}, \dots, R_{t-p}^{(n)}) = S_n^\alpha(\mathbf{X}; U_{(\cdot)}),$$

where  $U_{(\cdot)}$  is the vector  $(U_{(1)}, U_{(2)}, \dots, U_{(n)})$ . Conditionally, given the vector  $U_{(\cdot)}$ ,  $\alpha_n^{U_{(\cdot)}}(\dots) = a_n(\dots) - J(\dots)$  is indeed a score function, and, consequently,  $S_n^\alpha(\mathbf{X}; U_{(\cdot)})$  is a linear serial rank statistic.

Now

$$E[\Delta_n^2(\mathbf{X})] = E_{U_{(\cdot)}}[E\{((n - p)^{1/2}S_n^\alpha(\mathbf{X}; U_{(\cdot)}) - ((n - p)^{1/2}m_n - \mathcal{E}_n(\mathbf{X}))\}^2 | U_{(\cdot)})],$$

and

$$\begin{aligned} & E[(n - p)^{1/2}S_n^\alpha(\mathbf{X}; U_{(\cdot)}) | U_{(\cdot)}] \\ &= (n - p)^{1/2}[Ea_n(R_{p+1}^{(n)}, \dots, R_1^{(n)}) - EJ(U_{R_{p+1}^{(n)}}), \dots, U_{R_1^{(n)}})] \\ &= (n - p)^{1/2} \left[ m_n - \frac{1}{n(n - 1) \dots (n - p)} \sum_{1 \leq t_1 \neq \dots \neq t_{p+1} \leq n} J(U_{t_1} \dots U_{t_{p+1}}) \right] \\ &= (n - p)^{1/2}m_n - \mathcal{E}_n. \end{aligned}$$

Hence, denoting by  $D^2(S_n^\alpha | U_{(\cdot)})$  the conditional variance of  $S_n^\alpha(\mathbf{X}; U_{(\cdot)})$ , we have

$$E[\Delta_n^2] = (n - p)E_{U_{(\cdot)}}[D^2(S_n^\alpha | U_{(\cdot)})].$$

In view of Lemmas 2 and 4 (Appendix 3), we obtain

$$\begin{aligned} E[\Delta_n^2] &\leq E_{U_{(\cdot)}}[D^2(\alpha_n^{U_{(\cdot)}} | U_{(\cdot)}) (2p + 1) \\ &\quad + n \text{Cov}(\alpha_n^{U_{(\cdot)}}(R_{p+1}^{(n)}, R_p^{(n)}, \dots, R_1^{(n)}), \alpha_n^{U_{(\cdot)}}(R_{2p+2}^{(n)}, R_{2p+1}^{(n)}, \dots, R_{p+2}^{(n)}) | U_{(\cdot)})] \\ &\leq (2p + 1 + K)E_{U_{(\cdot)}}[E\{(\alpha_n^{U_{(\cdot)}}(R_{p+1}^{(n)}, \dots, R_1^{(n)}))^2 | U_{(\cdot)}\}] \\ &= (2p + 1 + K)E\{(\alpha_n^{U_{(\cdot)}}(R_{p+1}^{(n)}, \dots, R_1^{(n)}))^2\}, \end{aligned}$$

which by (4.2) converges to 0 (as  $n \rightarrow \infty$ ).

4.2. We now show that  $n^{-1/2}(\mathcal{L}_n - \mathcal{E}_n)$  and  $n^{-1/2}\mathcal{L}_n^0$  are asymptotically equivalent to  $U$ -statistics.

Define the  $(p + 1)$ -dimensional random variables

$$(4.6) \quad \mathbf{Y}_t = \begin{pmatrix} Y_{t,1} \\ \vdots \\ Y_{t,p+1} \end{pmatrix} = \begin{pmatrix} U_t \\ \vdots \\ U_{t-p} \end{pmatrix}, \quad p + 1 \leq t \leq n.$$

The  $\mathbf{Y}_t$ 's are identically distributed (uniformly over  $[0, 1]^{p+1}$ , under  $H_0^{(n)}$ ); of course, they are not independent—but, being  $p$ -dependent, they constitute an *absolutely regular process* (Yoshihara, 1976).



Consider the function  $G(\mathbf{y})$  (of  $(p + 1)$  arguments  $\mathbf{y} = (y_1, \dots, y_{p+1}) \in [0, 1]^{p+1}$ ) where

$$(4.7) \quad G(\mathbf{y}) = G(y_1, \dots, y_{p+1}) = \phi(F^{-1}(y_1)) \sum_{i=1}^p d_i F^{-1}(y_{i+1}).$$

$G(\mathbf{Y})$  could serve as a kernel (of degree 1) converting the variables  $n^{-1/2} \mathcal{L}_n^0$  into a sequence of  $U$ -statistics. However, in order to obtain a  $U$ -statistic approximation of  $n^{-1/2} \mathcal{E}_n$  we shall need a kernel of degree  $(p + 1)$ . Therefore, put

$$(4.8) \quad \Phi^{\mathcal{L}}(\mathbf{Y}_{t_1}, \dots, \mathbf{Y}_{t_{p+1}}) = \sum_{j=1}^{p+1} G(\mathbf{Y}_{t_j}) / (p + 1),$$

then the corresponding  $U$ -statistic is

$$\binom{n-p}{p+1}^{-1} \sum_{p+1 \leq t_1 < \dots < t_{p+1} \leq n} \Phi^{\mathcal{L}}(\mathbf{Y}_{t_1}, \dots, \mathbf{Y}_{t_{p+1}}) = n^{-1/2} \mathcal{L}_n^0 + o_p(n^{-1/2}).$$

In a similar way, the kernel

$$(4.9) \quad \begin{aligned} \Phi^{\mathcal{J}}(\mathbf{Y}_{t_1}, \dots, \mathbf{Y}_{t_{p+1}}) &= \sum_{j=1}^{p+1} J(\mathbf{Y}_{t_j}) / (p + 1) \\ &= \sum_{j=1}^{p+1} J(Y_{t_j,1}, \dots, Y_{t_j,p+1}) / (p + 1) \end{aligned}$$

defines a  $U$ -statistic which is asymptotically equivalent to  $n^{-1/2} \mathcal{J}_n$ :

$$\binom{n-p}{p+1}^{-1} \sum_{p+1 \leq t_1 < \dots < t_{p+1} \leq n} \Phi^{\mathcal{J}}(\mathbf{Y}_{t_1}, \dots, \mathbf{Y}_{t_{p+1}}) = n^{-1/2} \mathcal{J}_n + o_p(n^{-1/2}).$$

As for  $\mathcal{E}_n$ , let us consider the kernel

$$(4.10) \quad \Phi^{\mathcal{E}}(\mathbf{Y}_{t_1}, \dots, \mathbf{Y}_{t_{p+1}}) = (1/(p + 1)!) \sum_j J(Y_{j_1,1}, \dots, Y_{j_{p+1},1}),$$

where the summation  $\sum_j$  extends over all possible  $(p + 1)!$  permutations  $(j_1, \dots, j_{p+1})$  of  $(t_1, \dots, t_{p+1})$ . The corresponding  $U$ -statistic is

$$\begin{aligned} \mathcal{U}_n^{\mathcal{E}} &= \binom{n-p}{p+1}^{-1} \sum_{p+1 \leq t_1 < \dots < t_{p+1} \leq n} \Phi^{\mathcal{E}}(\mathbf{Y}_{t_1}, \dots, \mathbf{Y}_{t_{p+1}}) \\ &= (1/(n-p) \dots (n-2p)) \sum_{p+1 \leq t_1 \neq \dots \neq t_{p+1} \leq n} J(U_{t_1} \dots U_{t_{p+1}}). \end{aligned}$$

From (4.4), we obtain

$$(n-p)^{-1/2} \mathcal{E}_n = ((n-p) \dots (n-2p)) / (n(n-1) \dots (n-p)) \mathcal{U}_n^{\mathcal{E}} + \delta_n,$$

where

$$\begin{aligned} \delta_n &= (1/n(n-1) \dots (n-p)) \sum_{r=1}^p \sum_{1 \leq j_1 \neq \dots \neq j_r \leq p} \sum_{1 \leq k_1 < \dots < k_r < p+1} \sum_{p+1 \leq t_1 \neq \dots \neq t_{p+1-r} \leq n} \\ &\quad J(U_{t_1}, \dots, U_{t_{k_1-1}} U_{j_1} U_{t_{k_1}}, \dots, U_{t_{k_2-2}} U_{j_2} U_{t_{k_2-1}}, \dots, U_{t_{k_r-r}} U_{j_r} U_{t_{k_r-r+1}}, \dots, U_{t_{p+1-r}}). \end{aligned}$$

The number of terms, in  $\delta_n$ , is

$$\begin{aligned} & \sum_{\ell=1}^p p(p-1) \cdots \\ & (p-\ell+1) \frac{(p+1) \cdots (p+2-\ell)}{\ell!} (n-p) \cdots (n-2p+\ell) \\ & \leq p!(p+1)!(n-p) \cdots (n-2p+1). \end{aligned}$$

Thus,

$$E|\delta_n| \leq p!(p+1)! \frac{(n-p) \cdots (n-2p+1)}{n(n-1) \cdots (n-p)} E|J(U_{p+1} \cdots U_1)| = O(n^{-1}),$$

and

$$n^{-1/2} \mathcal{E}_n = \mathcal{U}_n^{\mathcal{E}} + o_p(n^{-1/2}).$$

4.3. *Asymptotic normality.* It follows from the conclusions of the preceding section that,  $\alpha$  and  $\beta$  being arbitrary coefficients,  $\mathcal{U}_{\alpha\beta}^n = n^{-1/2}[\alpha(\mathcal{S}_n - \mathcal{E}_n) + \beta \mathcal{L}_n^0]$  is (up to  $o_p(n^{-1/2})$  terms) a sequence of  $U$ -statistics, with kernel

$$(4.11) \quad \Phi_{\alpha\beta} = \alpha(\Phi^{\mathcal{S}} - \Phi^{\mathcal{E}}) + \beta\Phi^{\mathcal{L}}.$$

A number of results are available concerning the asymptotic behaviour of  $U$ -statistics in case the sample observations (here, the  $Y_i$ 's) are not independent. The one we are using in Proposition 4.1 below is a theorem established by Yoshihara (1976) for the class of stationary absolutely regular processes—from which the  $p$ -dependent processes constitute a subclass.

Let

$$(4.12) \quad \begin{aligned} J^*(u_{p+1}, \dots, u_1) &= J(u_{p+1}, \dots, u_1) \\ &- \sum_{k=1}^{p+1} \int_{[0,1]^p} J(v_p, \dots, v_k, u_1, v_{k-1}, \dots, v_1) dv_1 \cdots dv_p \\ &+ p \int_{[0,1]^{p+1}} J(v_{p+1}, \dots, v_1) dv_1 \cdots dv_{p+1}. \end{aligned}$$

Obviously,  $E[J^*(U_{p+1}, \dots, U_1)] = 0$ . We have then the following results.

PROPOSITION 4.1. Under  $H_0^{(n)}$ ,

$$\begin{pmatrix} \sqrt{n}(S_n - m_n) \\ \log L_n \end{pmatrix}$$

is asymptotically normal, with mean

$$\begin{pmatrix} 0 \\ -1/2 \sum_{i=1}^p d_i^2 \sigma^2 I(f) \end{pmatrix}$$

and covariance matrix

$$\begin{pmatrix} V^2 & \sum_{i=1}^p d_i C_i \\ \sum_{i=1}^p d_i C_i & \sum_{i=1}^p d_i^2 \sigma^2 I(f) \end{pmatrix},$$

where

$$(4.13) \quad V^2 = \int_{[0,1]^{p+1}} J^{*2}(v_{p+1}, \dots, v_1) dv_1 \cdots dv_{p+1} + 2 \sum_{j=1}^p \int_{[0,1]^{p+1+j}} J^*(v_{p+1}, \dots, v_1) J^*(v_{p+1+j}, \dots, v_{1+j}) dv_1 \cdots dv_{p+1+j}$$

and

$$(4.14) \quad C_i = \int_{[0,1]^{p+1}} J^*(v_{p+1}, \dots, v_1) \sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}) dv_1 \cdots dv_{p+1}.$$

PROPOSITION 4.2. Under  $H_1^{(n)}$ ,  $\sqrt{n}(S_n - m_n)$  is asymptotically normal, with mean  $\sum_{i=1}^p d_i C_i$  and variance  $V^2$ .

The proof of Proposition 4.1 is given in Appendix 4; Proposition 4.2 follows immediately from Proposition 4.1, by application of LeCam’s third lemma. Note that the asymptotic variance  $V^2$  depends only on the score-generating function  $J(\dots)$ , not on the ARMA model coefficients  $a_i$  and  $b_i$  (cf. (3.7)), nor on the white noise density  $f$ . The mean, on the contrary, depends on both  $f$  and the coefficients  $a_i$  and  $b_i$ ; however, it remains invariant under scale transformations because of its dependence on  $f$  only through  $\phi(F^{-1}(u))F^{-1}(v)$ .

**5. Asymptotic efficiency of linear serial rank statistics.**

5.1. ARE of two linear serial rank statistics—Optimal scores. The results of Proposition 4.2 allow for an explicit form of the asymptotic relative efficiency of two linear serial rank statistics.

Let  $S_n^{(1)}$  and  $S_n^{(2)}$  have asymptotic normal distributions under  $H_1^{(n)}$ , with means  $\sum d_i C_i^{(1)}$  and  $\sum d_i C_i^{(2)}$  and with variances  $V_{(1)}^2$  and  $V_{(2)}^2$ . Then, the ARE of  $S_n^{(1)}$  with respect to  $S_n^{(2)}$  is

$$(5.1) \quad e(S_n^{(1)}, S_n^{(2)}) = (V_{(2)} \sum_{i=1}^p d_i C_i^{(1)} / V_{(1)} \sum_{i=1}^p d_i C_i^{(2)})^2.$$

A test statistic  $\bar{S}_n$  such that  $e(\bar{S}_n, S_n) \geq 1$  for any linear serial rank statistic  $S_n$  will be asymptotically the most efficient statistic (in Pitman’s sense) within the class of linear serial rank statistics for testing randomness (i.e.,  $H_0^{(n)}$ ) against ARMA dependence (i.e.,  $H_1^{(n)}$ ). (A test based on  $\bar{S}_n$  will be called the asymptotically optimal test against  $H_1^{(n)}$ .) We shall denote by  $H_d^{(n)}$ , a sequence of alternatives characterized by the coefficients  $\underline{d} = (d_1, \dots, d_p)$  (cf. (1.2) and (3.6)).

The following result provides an explicit formulation of the optimal test statistic.

PROPOSITION 5.1. An asymptotically optimal linear serial rank test for  $H_0^{(n)}$  against  $H_d^{(n)}$  is provided by any statistic  $S_n^{\underline{d}}$  with score-generating function (up to additive and multiplicative constants) given by

$$(5.2) \quad J^{\underline{d}}(v_{p+1}, \dots, v_1) = \sum_{i=1}^p (d_i / (p + 1 - i)) \sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}).$$

Under  $H_h^{(n)}$  ( $h \in \mathbb{R}^p$ ),  $n^{1/2}(S_n^d - m_n^d)$  is asymptotically normal, with mean  $\sum_{i=1}^p h_i d_i \sigma^2 I(f)$  and variance  $V_d^2 = \sum_{i=1}^p d_i^2 \sigma^2 I(f)$ .

This optimality result relies on the following lemma whose proof is given in Appendix 5.

LEMMA 5.1. Let  $S_n$  be a linear rank statistic with score-generating function  $J^*(v_{p+1}, \dots, v_1)$  (cf. (4.12)), and let

$$(5.3) \quad \begin{aligned} J_0^*(v_{p+1}, \dots, v_1) \\ = (\sigma^2 I(f))^{-1} \sum_{i=1}^p (C_i / (p + 1 - i)) \sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}). \end{aligned}$$

Denote by  $S_n^0$  a linear serial rank statistic associated with  $J_0^*$ . Then  $e(S_n, S_n^0) \leq 1$  for any alternative  $H_n^d$ .

$J_0^*(v_{p+1}, \dots, v_1)$  is actually the projection of  $J^*(v_{p+1}, \dots, v_1)$  onto the linear  $L^2$ -space spanned by  $\{\sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}); i = 1, \dots, p\}$ . This immediately follows from (4.14) and

$$\int [\sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i})]^2 dv_{p+1} \dots dv_1 = (p + 1 - i) \sigma^2 I(f).$$

Similarly,  $(S_n^0 - m_n^0)$  is the projection (under  $H_0^{(n)}$  and up to  $o_p(n^{-1/2})$  terms) of  $(S_n - m_n)$  onto the second-order linear space spanned by  $p$  linear serial rank statistics of orders 1 through  $p$ , associated with the score-generating functions

$$\sum_{j=1}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}), \quad i = 1, \dots, p.$$

PROOF OF PROPOSITION 5.1. Because of Lemma 5.1, we may restrict our attention to score functions of the form (5.3). The corresponding statistic has an asymptotic variance

$$\begin{aligned} V_0^2 = \sum_{i=1}^p \sum_{j=0}^{p-i} c_i^2 \int [\phi(F^{-1}(u)) F^{-1}(v)]^2 du dv \\ + 2 \sum_{s=1}^p \sum_{i=1}^p \sum_{j=0}^{p-i} \sum_{k=1}^p \sum_{\ell=0}^{p-k} c_i c_k \int \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}) \\ \cdot \phi(F^{-1}(v_{p+1+s-\ell})) F^{-1}(v_{p+1+s-\ell-k}) dv \end{aligned}$$

where  $c_i = (\sigma^2 I(f))^{-1} C_i / (p + 1 - i)$ .

Omitting routine computations, we obtain

$$(5.4) \quad \begin{aligned} V_0^2 = \sigma^2 I(f) [\sum_{i=1}^p (p + 1 - i) c_i^2 + 2 \sum_{i=1}^p c_i^2 \sum_{j=1}^{p-i} (p + 1 - j - i)] \\ = \sigma^2 I(f) \sum_{i=1}^p c_i^2 (p + 1 - i)^2 = (\sigma^2 I(f))^{-1} \sum_{i=1}^p C_i^2. \end{aligned}$$

The values of the  $C_i$ 's that maximize

$$(\sum_{i=1}^p d_i C_i)^2 / V_0^2 = (\sigma^2 I(f)) (\sum_{i=1}^p d_i C_i)^2 / \sum_{i=1}^p C_i^2$$

are thus proportional to the  $d_i$ 's, and this establishes the desired result.  $\square$

Note that the optimal scores (5.2) could also be taken as

$$J^d(v_{p+1}, \dots, v_1) = \sum_{i=1}^p (d_i/(p + 1 - i)) \sum_{j=0}^{p-i} F^{-1}(v_{p+1-j})\phi(F^{-1}(v_{p+1-j-i})).$$

This is in accordance with the fact that, due to stationarity, if  $\{x_t\}$  is generated by an ARMA model of the form (3.1), it can also be seen as generated by an ARMA model of the form

$$X_t - n^{-1/2} \sum_{i=1}^{p_1} a_i X_{t+i} = \varepsilon_t + n^{-1/2} \sum_{i=1}^{p_2} b_i \varepsilon_{t+i}.$$

An additional consequence of Proposition 5.1 is that the “optimal” serial rank statistic against an ARMA( $p_1, p_2$ ) alternative is a serial rank statistic of order  $p = \max(p_1, p_2)$ . Using a statistic of the wrong order—whether too low or too high—always results into a loss of efficiency.

**EXAMPLES.** The optimal score functions in (5.2) depend only on the *type* of the density of the generating white noise; examples are:

(i) Gaussian white noise (van der Waerden scores)

$$(5.5) \quad J(v_{p+1}, \dots, v_1) = \sum_{i=1}^p (d_i/(p + 1 - i)) \sum_{j=0}^{p-i} \Phi^{-1}(v_{p+1-j})\Phi^{-1}(v_{p+1-j-i})$$

$$\left( \text{with } \Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x e^{-u^2/2} du \right).$$

(ii) Logistic white noise (Wilcoxon scores)

$$(5.6) \quad J(v_{p+1}, \dots, v_1) = \sum_{i=1}^p (d_i/(p + 1 - i)) \sum_{j=0}^{p-i} (2 v_{p+1-j} - 1)\log(v_{p+1-j-i}/(1 - v_{p+1-j-i})).$$

(iii) Double exponential white noise (median test scores)

$$(5.7) \quad J(v_{p+1}, \dots, v_1) = \sum_{i=1}^p d_i/(p + 1 - i) \sum_{j=0}^{p-i} \text{sgn}(v_{p+1-j} - 1/2)F_e^{-1}(v_{p+1-j-i})$$

$$\left( \text{with } F_e(x) = \frac{1}{2} \int_{-\infty}^x e^{-|u|} du \right).$$

As an example of an optimal test statistic, let us consider the problem of testing randomness against the sequence of ARMA(2, 1) alternatives

$$X_t + (1.4)n^{-1/2}X_{t-1} + (0.5)n^{-1/2}X_{t-2} = \varepsilon_t + (0.2)n^{-1/2}\varepsilon_{t-1}, \quad t \in \mathbb{Z}$$

where  $\{\varepsilon_t\}$  has a logistic density (with unspecified variance). The following serial

version of Wilcoxon’s test is then optimal:

$$\begin{aligned}
 S_n &= (n - 2)^{-1} \sum_{t=3}^n \left\{ \frac{-1.2}{2} \left[ \left( 2 \frac{R_t^{(n)}}{n + 1} - 1 \right) \log \left( \frac{R_{t-1}^{(n)}}{n - R_{t-1}^{(n)} + 1} \right) \right. \right. \\
 &\quad \left. \left. + \left( 2 \frac{R_{t-1}^{(n)}}{n + 1} - 1 \right) \log \left( \frac{R_{t-2}^{(n)}}{n - R_{t-2}^{(n)} + 1} \right) \right] \right. \\
 &\quad \left. - 0.5 \left( 2 \frac{R_t^{(n)}}{n + 1} - 1 \right) \log \left( \frac{R_{t-2}^{(n)}}{n - R_{t-2}^{(n)} + 1} \right) \right\} \\
 &= -(1.2)(n - 1)^{-1} \sum_{t=2}^n \left( 2 \frac{R_t^{(n)}}{n + 1} - 1 \right) \log \left( \frac{R_{t-1}^{(n)}}{n - R_{t-1}^{(n)} + 1} \right) \\
 &\quad - (0.5)(n - 2)^{-1} \sum_{t=3}^n \left( 2 \frac{R_t^{(n)}}{n + 1} - 1 \right) \log \left( \frac{R_{t-2}^{(n)}}{n - R_{t-2}^{(n)} + 1} \right) + O_p(n^{-1}).
 \end{aligned}$$

5.2. *Comparison with classical parametric procedures.* The parametric version of our problem is treated, in the Gaussian case, by Anderson (1971, Chapter 6). For the problem of testing for randomness against an alternative of AR( $p$ ) dependence

$$(5.8) \quad X_t - \sum_{i=1}^p \beta_i X_{t-i} = \varepsilon_t, \quad t \in \mathbb{Z}$$

(with  $\beta_p \neq 0$ ), Anderson (1971, page 266) shows that the most powerful similar test is based on a test statistic

$$(5.9) \quad \sum_{k=1}^p \gamma_k r_k^{(n)},$$

where

$$(5.10) \quad \gamma_k = \beta_k + \sum_{i=1}^{p-k} \beta_i \beta_{i+k},$$

$r_k^{(n)}$  being a serial correlation coefficient of order  $k$ . These coefficients are such that

$$(5.11) \quad r_k^{(n)} = r_k^{(n)} + o_p(n^{-1/2}),$$

where  $r_k^{(n)}$  is the sample autocorrelation coefficient of order  $k$ , i.e.,

$$(5.12) \quad r_k^{(n)} = (n/(n - k)) \sum_{t=k+1}^n X_t X_{t-k} / \sum_{t=1}^n X_t^2.$$

It is therefore natural to compare our linear serial rank statistics (1.1) with linear combinations of the autocorrelation coefficients (also, time-series analysts generally give much attention to the *correlogram* of their series). In order to do this, we need the asymptotic joint distribution of those linear combinations and  $\log L_n$  under  $H_0^{(n)}$ .

PROPOSITION 5.2. Under  $H_0^{(n)}$ ,  $(n^{1/2}r_1^{(n)}, \dots, n^{1/2}r_p^{(n)}, \log L_n)'$  is asymptotically normal, with mean  $(0, \dots, 0, -1/2 \sum_{i=1}^p d_i^2 \sigma^2 I(f))'$  and covariance matrix

$$\begin{pmatrix} & & & d_1 \\ & & & \vdots \\ I & & & \\ & & & d_p \\ d_1 \cdots d_p & \sum d_i^2 \sigma^2 I(f) & & \end{pmatrix}$$

( $I$  being the  $p \times p$  identity matrix).

COROLLARY 5.1. Under  $H_d^{(n)}$ ,  $n^{1/2} \sum_{k=1}^p \alpha_k r_k^{(n)}$  is asymptotically normal with mean  $\sum_{i=1}^p \alpha_i d_i$  and variance  $\sum_{i=1}^p \alpha_i^2$ .

COROLLARY 5.2. The asymptotically most efficient (in Pitman's sense) linear combination of the  $r_k^{(n)}$ 's against  $H_d^{(n)}$  is  $\sum_{k=1}^p d_k r_k^{(n)}$ .

PROOF. See Appendix 6.  $\square$

Letting  $\beta_i = \beta_i^{(n)} = n^{-1/2} a_i$  in (5.10), and by denoting by  $\gamma_k^{(n)}$  the corresponding values of  $\gamma_k$ , we obtain in the pure AR( $p$ ) case, that

$$\sum_{k=1}^p d_k r_k^{(n)} = n^{1/2} \sum_{k=1}^p \gamma_k^{(n)} \dot{r}_k^{(n)} + o_p.$$

The asymptotically optimal test statistic in Corollary 5.2 is thus asymptotically equivalent to the one providing the most powerful similar test (which is of course not surprising!).

Now, if we denote by  $e$  the ARE of the asymptotically optimal serial rank statistic defined in Proposition 5.1 with respect to the asymptotically optimal combination of autocorrelations  $\sum_{k=1}^p d_k r_k^{(n)}$ , we obtain the following result.

PROPOSITION 5.3.

$$(5.13) \quad e = \sigma^2 I(f).$$

PROOF. On account of Proposition 5.1 and Corollary 5.1, we have

$$e = \frac{(\sum_{i=1}^p d_i^2 \sigma^2 I(f))^2 / \sum_{i=1}^p d_i^2 \sigma^2 I(f)}{(\sum_{i=1}^p d_i^2)^2 / \sum_{i=1}^p d_i^2} = \sigma^2 I(f). \quad \square$$

This efficiency thus depends on the density  $f$ . For normal, logistic, and double exponential densities, it takes the values 1,  $\pi^2/9$ , and 2, respectively.

A more detailed study for the case  $p = 1$  is given in the next section.

5.3. *The first-order case (AR(1), MA(1) or ARMA(1, 1) alternatives).* The asymptotic mean of any linear serial rank statistic has the form of an inner product  $\sum_{i=1}^p d_i C_i$ ; any of these statistics has thus the drawback that, when

$p \geq 2$ , there exists, for any type of density  $f$ , a family of alternatives against which this statistic has a zero efficiency—namely, the alternatives characterized by a vector of coefficients  $\underline{d} = (d_i)$  (cf. 3.7)) orthogonal to  $\underline{C} = (C_i)$  (4.14) (recall that  $C_i$  depends only on  $J$  and  $f$ ).

This does not happen, however, in the first-order case ( $p = 1$ ); indeed the ARE of a linear serial rank statistic  $S_n^{(1)}$  with respect to another  $S_n^{(2)}$  is given by

$$e(S_n^{(1)}, S_n^{(2)}) = \left( \frac{C_{(1)}/V_{(1)}}{C_{(2)}/V_{(2)}} \right)^2$$

and hence does not depend on the alternative. As a consequence, the statistic

$$(5.14) \quad S'_n = (n - 1)^{-1} \sum_{t=2}^n \phi \left( F^{-1} \left( \frac{R_t^{(n)}}{n + 1} \right) \right) F^{-1} \left( \frac{R_{t-1}^{(n)}}{n + 1} \right)$$

is uniformly asymptotically most efficient against the whole family of ARMA(1, 1) alternatives where  $f$  is a strongly unimodal density (cf. Hájek and Sidák, 1967).

The classical parametric test statistic in this problem is the first-order autocorrelation coefficient  $r_1^{(n)}$ . Under the usual gaussian assumptions, this statistic (cf. Anderson, 1971, Chapter 6) provides the (one-sided) uniformly most powerful similar and (two-sided) unbiased tests. It is therefore interesting to investigate the AREs  $e(S'_n)$  of (5.14) with respect to  $r_1^{(n)}$ .

In Table 1, we give, for various score functions (van der Waerden, Wilcoxon, median test, Spearman) and various density types (normal, logistic and double exponential), the values of  $e(S'_n)$ .

Because of its popularity, we also include in this comparative study the Spearman autocorrelation coefficient of order 1, viz.

$$r_S^n = \frac{(n - 1)^{-1} \sum_{t=2}^n R_t^{(n)} R_{t-1}^{(n)} - (n + 1)/2}{(n^2 - 1)/12}.$$

$r_S^n$  is equivalent (for strongly unimodal densities  $f(\cdot)$ ) to the linear serial rank statistic  $S_n^S$  induced by the score-generating function  $J(u, v) = uv$ . Indeed, it can be shown that

$$r_S^n = (S_n^S - 1/4)/(1/12(n - 1)/(n + 1)),$$

with

$$m_n^S = E(S_n^S) = (3n^2 - n - 2)/12(n^2 - 1)$$

and

$$V^S = \int_{[0,1]^2} \left( uv - \frac{u + v}{2} + \frac{1}{4} \right)^2 du dv = (144)^{-1}.$$

Consequently,  $n^{1/2}(S_n^S - m_n^S)$  is asymptotically normal, with mean 0 and variance 1/144 (under  $H_0^{(n)}$ ). This confirms a classical result (cf., for example, Kendall and Stuart, 1968) that  $n^{1/2}r_S^n$  is asymptotically normal, with mean 0 and variance 1.



TABLE 1  
Asymptotic relative efficiencies of some linear serial rank statistics with respect to the first-order sample autocorrelation

Score Generating Functions $J(u, v)$	$V^2$	Types of Densities		
		Normal	Logistic	Double Exponential
van der Waerden $\Phi^{-1}(u)\Phi^{-1}(v)$	1	1	$\frac{1}{\Pi^3} \left[ \int_{-\infty}^{+\infty} e^{-x^2} (\Phi(x))^{-1} dx \right]^2$ $\approx 1.047$	$\frac{2}{\Pi^3} \left[ \int_{-\infty}^0 e^{-x^2} (\Phi(x))^{-1} dx \right]^2$ $\approx 1.226$
Wilcoxon $(2u-1)\log(v/1-v)$	$\frac{\Pi^2}{9}$	$\frac{9}{\Pi^5} \left[ \int_{-\infty}^{+\infty} e^{-x^2} (\Phi(x))^{-1} dx \right]^2$ $\approx 0.948$	$\frac{\Pi^2}{9} \approx 1.097$	$\frac{9}{4\Pi^2} \left[ \frac{\Pi^2}{6} - (\log 2)^2 + \log 4 \right]^2$ $\approx 1.483$
Median Test $\text{sgn}(u - \frac{1}{2})F_2^{-1}(v)$	2	$\frac{1}{\Pi^3} \left[ \int_{-\infty}^0 e^{-x^2} (\Phi(x))^{-1} dx \right]^2$ $\approx 0.613$	$\frac{1}{8} \left[ \frac{\Pi^2}{6} - (\log 2)^2 + \log 4 \right]^2$ $\approx 0.813$	2
Spearman $uv$	$(144)^{-1}$	$\left(\frac{3}{\Pi}\right)^2 \approx 0.912$	1	$\frac{1}{4}(\frac{3}{2})^4 \approx 1.226$
Maximal ARE		1	1.097	2

APPENDIX 1

*The Green's functions.* Denoting by  $L$  the lag operator ( $L\psi_t = \psi_{t-1}$ ), consider the linear difference operator of order  $p$ :  $C(L) = 1 + \sum_{i=1}^p c_i L^i$ , where the  $c_i$ 's are real and  $c_p \neq 0$ . The Green's function  $g_u$  associated with this operator is (cf. Miller, 1968, Chapter 2.2 or Hallin, 1984) the value in  $t = u$  of the solution of the homogeneous difference equation  $\psi_t + \sum_{i=1}^p c_i \psi_{t-i} = 0$ ,  $t \in \mathbb{Z}$ , taking on initial values  $\psi_0 = 1, \psi_{-1} = \dots = \psi_{-p+1} = 0$ .

Let  $g_u^{(n)}$  be the Green's function associated with the operator  $B^{(n)}(L) = 1 + n^{-1/2} \sum_{i=1}^p b_i L^i$ . The following lemma will be helpful for the proof of Proposition 3.1.

LEMMA 1.

$$\sum_{u=v}^w |g_u^{(n)}| \leq \sum_{u=[(v-1)/p]+1}^w (pb_M n^{-1/2})^u \quad \forall w \geq v \geq 0$$

where  $b_M = \max_i |b_i|$

COROLLARY 1. For  $n > 4(pb_M)^2$ ,

$$(A1) \quad \sum_{u=pv+1}^\infty |g_u^{(n)}| \leq 2(pb_M n^{-1/2})^{v+1} = o(n^{-v/2}) \quad \forall v \geq 0.$$

PROOF.  $g_u^{(n)}$  is a sum of terms of the form

$$(A2) \quad (-1)^k (b_{i_{(1)}} b_{i_{(2)}} \dots b_{i_{(k)}}) n^{-k/2},$$

where the indexes  $i_{(1)}, \dots, i_{(k)} \in \{1, \dots, p\}$  need not be distinct, and  $[(u-1)/p] + 1 \leq k \leq u$  ( $[z]$  stands for the largest integer  $\leq z$ ). Each of these terms is absolutely bounded by  $(b_M n^{-1/2})^k$ .

Denote by  $\nu_u^k$  the maximal (some might collapse) number of these terms in  $g_u^{(n)}$ .  $\nu_u^k$  does not depend on  $n$ . Put  $\nu_0^k = \delta_{k0}$ , and  $\nu_u^k = 0$  for  $u < 0$  or  $k \notin [[(u-1)/p] + 1; u]$ . Obviously,  $\nu_u^k = \sum_{i=1}^p \nu_{u-i}^{k-1}$ . Hence, letting  $\nu^k = \sum_{u=0}^\infty \nu_u^k = \sum_{u=k}^{p^k} \nu_u^k$ , we get  $\nu^k = \sum_{i=1}^p \sum_{u=0}^\infty \nu_{u-i}^{k-1} = p\nu^{k-1}$  and, since  $\nu^0 = 1$ ,  $\nu^k = p^k$ ; it follows that

$$\sum_{u=v}^w |g_u^{(n)}| \leq \sum_{k=[(v-1)/p]+1}^w \nu^k (b_M n^{-1/2})^k = \sum_k (pb_M n^{-1/2})^k.$$

The proof of Corollary 1 is straightforward.  $\square$

APPENDIX 2

PROOF OF PROPOSITION 3.1. To avoid unnecessary details of computations involved in the proof for the general ARMA( $p_1, p_2$ ) case, we successively consider the pure AR( $p$ ) and MA( $p$ ) cases, since these two cases cover all the problems raised in the general case.

A2.1. *The AR( $p$ ) case.* Let  $\mathbf{X}_0 = (X_{-p+1}, \dots, X_0)$ , and

$$\prod_{i=1}^p f(\mathbf{X}_t - n^{-1/2} \sum_{i=1}^p a_i \mathbf{X}_{t-i}) = \lambda^{(n)}(\mathbf{X}; \mathbf{X}_0).$$

Then, we have (a.e.)

$$(A3) \quad \begin{aligned} & \log \lambda^{(n)}(\mathbf{X}; \mathbf{X}_0) \\ &= \sum_{t=1}^n \log f(X_t) + n^{-1/2} \sum_{t=1}^n \phi(X_t) \sum_{i=1}^p a_i X_{t-i} \\ & \quad - (2n)^{-1} \sum_{t=1}^n \phi'(X_t - \theta n^{-1/2} \sum_{i=1}^p a_i X_{t-i}) \sum_{i,j=1}^p a_i a_j X_{t-i} X_{t-j}, \end{aligned}$$

with  $\theta = \theta(X_n, \dots, X_{-p+1}) \in [0, 1]$ . The first-order term in (A3) is

$$(A4) \quad \mathcal{L}_n^0(\mathbf{X}) + n^{-1/2} \sum_{t=1}^n \phi(X_t) \sum_{i=1}^p a_i X_{t-i} = \mathcal{L}_n^0(\mathbf{X}) + o_p.$$

$\mathcal{L}_n^0$  is itself of the form  $n^{-1/2} \sum_{t=p+1}^n Z_t$ , where the  $Z_t$ 's are  $p$ -dependent. It is therefore (cf., for example, Anderson, 1971, page 427) asymptotically normal with mean  $E[Z_t] = 0$  and variance

$$E[Z_t^2] + 2 \sum_{i=1}^p E[Z_t Z_{t+i}] = (\sum_{i=1}^p a_i^2) \sigma^2 I(f) + 0 = d^2.$$

Now consider the second-order term in (A3). Put  $|a_M| = \max_i |a_i|$  and  $Q_t = \sum_{i,j=1}^p a_i a_j X_{t-i} X_{t-j}$ .  $\phi'$  being Lipschitzian (assumption (iv)), we have

$$\begin{aligned} & \left| \sum_{t=1}^n Q_t \phi'(X_t - \theta n^{-1/2} \sum_{i=1}^p a_i X_{t-i}) - \sum_{t=1}^n Q_t \phi'(X_t) \right| \\ & \leq n^{-1/2} A |a_M| \sum_{t=1}^n |Q_t| \sum_{i=1}^p |X_{t-i}| \\ & \leq n^{-1/2} A |a_M|^3 \sum_{t=1}^n \sum_{i,j,k=1}^p |X_{t-i} X_{t-j} X_{t-k}|. \end{aligned}$$

Since its summands are  $p$ -dependent, and since (assumption (i))  $E |X_r X_s X_t| < \infty \forall r, s, t$ , this upper bound is  $O_p(n^{1/2})$ . It follows that the second-order term in (A3) can be written as  $-(2n)^{-1} \sum_{t=1}^n \phi'(X_t) \sum_{i,j=1}^p a_i a_j X_{t-i} X_{t-j} + o_p$ , and hence converges in probability to  $-1/2 \sum_{i=1}^p a_i^2 E[\phi'(X_t) X_{t-i}^2] = -d^2/2$ . Summing up,

$$\log \lambda^{(n)}(\mathbf{X}; \mathbf{X}_0) = \log \mathcal{L}_0^{(n)}(\mathbf{X}) + \mathcal{L}_n^0(\mathbf{X}) - (d^2/2) + R^{(n)}(X_p, \dots, X_0, \dots, X_{-p+1}),$$

where  $R^{(n)}$  is  $o_p$ . Hence

$$\begin{aligned} \log L^{(n)}(\mathbf{X}) &= \log E[\exp(\log \lambda^{(n)}(\mathbf{X}; \mathbf{X}_0) - \log \mathcal{L}_0^{(n)}(\mathbf{X})) | \mathbf{X}] \\ &= (\mathcal{L}_n^0(\mathbf{X}) - d^2/2) + \log E[\exp(R^{(n)}) | \mathbf{X}] = (\mathcal{L}_n^0(\mathbf{X}) - d^2/2) + o_p, \end{aligned}$$

where the latter expectation is to be taken with respect to the joint distribution  $G_{a,b}^{(n)}$  of  $X_{-p+1}, \dots, X_0$ .

Proposition 3.1 now readily follows from the asymptotic normality of  $\mathcal{L}_n^0(\mathbf{X})$ .

**A2.2 The MA( $p$ ) case.** Define here  $\log \lambda^{(n)}(\mathbf{X}, \varepsilon_0)$  where  $\varepsilon_0 = (\varepsilon_{-p+1}, \dots, \varepsilon_0)$  as

$$(A5) \quad \begin{aligned} \log \lambda^{(n)}(\mathbf{X}, \varepsilon) &= \sum_{t=1}^n \log f(X_t + \sum_{u=1}^{t-1} g_u^{(n)} X_{t-u} + \sum_{u=t}^{t+p-1} g_u^{(n)} \varepsilon_{t-u}) \\ &= \sum_{t=1}^n \log f(X_t) - \sum_{t=1}^n (\sum_{u=1}^{t-1} g_u^{(n)} X_{t-u} + \sum_{u=t}^{t+p-1} g_u^{(n)} \varepsilon_{t-u}) \phi(X_t) \\ & \quad - 1/2 \sum_{t=1}^n (\sum_{u=1}^{t-1} g_u^{(n)} X_{t-u} + \sum_{u=t}^{t+p-1} g_u^{(n)} \varepsilon_{t-u})^2 \\ & \quad \cdot \phi'(X_t + \theta \sum_{u=1}^{t-1} g_u^{(n)} X_{t-u} + \theta \sum_{u=t}^{t+p-1} g_u^{(n)} \varepsilon_{t-u}) \end{aligned}$$

with  $\theta = \theta(\mathbf{X}, \varepsilon) \in [0, 1]$ .

The first-order term in (A5) is

$$-\sum_{u=1}^{n-1} g_u^{(n)} \sum_{t=u+1}^n \phi(X_t) X_{t-u} - \sum_{u=1}^{n+p-1} g_u^{(n)} \sum_{t=u-p+1}^u \phi(X_t) \varepsilon_{t-u} = -\sum_{(1)} - \sum_{(2)}.$$

Considering  $-\sum_{(1)}$ , decompose

$$\begin{aligned} -\sum_{(1)} &= n^{-1/2} \sum_{i=1}^p b_i \sum_{t=i+1}^n \phi(X_t)X_{t-i} - \sum_{u=1}^p (g_u^{(n)} + b_u n^{-1/2}) \sum_{t=u+1}^n \phi(X_t)X_{t-u} \\ &\quad - \sum_{u=p+1}^{2p} g_u^{(n)} \sum_{t=u+1}^n \phi(X_t)X_{t-u} - \sum_{u=2p+1}^{n-1} g_u^{(n)} \sum_{t=u+1}^n \phi(X_t)X_{t-u} \\ &= \sum_{(11)} - \sum_{(12)} - \sum_{(13)} - \sum_{(14)}. \end{aligned}$$

Up to a finite number of  $o_p$ 's  $\sum_{(11)} = \mathcal{L}_n^0(\mathbf{X})$  (as in (A4), replacing the  $a_i$ 's with  $b_i$ 's), and is therefore asymptotically normal, with mean 0 and variance  $d^2 = \sum_{i=1}^p b_i^2 \sigma^2 I(f)$ . For  $1 \leq u \leq p$ ,  $g_u^{(n)}$  is a finite sum of terms of the form

$$-b_u n^{-1/2}, b_{i(1)} b_{i(2)} n^{-2/2}, \dots, (-1)^u b_{i(1)} b_{i(2)} \dots b_{i(u)} n^{-u/2}.$$

$(g_u^{(n)} - b_u n^{-1/2})$  is thus  $O(n^{-1})$ , and, since  $E[\phi(X_t)X_{t-u}] = 0$ ,  $\sum_{(12)}$  converges to 0 in probability. Identical arguments apply to the case of  $\sum_{(13)}$ . As for  $\sum_{(14)}$ , it follows from Lemma 1 that

$$\begin{aligned} E | \sum_{(14)} | &\leq \sum_{u=2p+1}^{n-1} |g_u^{(n)}| (n - 2p) E |X_t| E | \phi(X_t) | \\ &\leq 2(qb_M)^3 n^{-3/2} (n - 2p) E |X_t| E | \phi(X_t) |. \end{aligned}$$

Hence  $\sum_{(14)}$  also converges to 0.

Similar arguments can be used to show that  $\sum_{(2)}$  is also  $o_p$ .

Let  $Q_t = | \sum_{u=1}^{t+1} g_u^{(n)} X_{t-u} + \sum_{u=t}^{t+p-1} g_u^{(n)} \varepsilon_{t-u} |$ , and consider the second-order term in (A5). Since  $X_t$  and  $\varepsilon_t$  have, under  $H_0^{(n)}$ , exactly the same distribution, and since our purpose is to prove that this second-order term converges in probability to a constant, we may write  $Q_t$  under the more convenient form  $Q_t = | \sum_{u=1}^{t+p-1} g_u^{(n)} X_{t-u} |$ .

$\phi'$  being (assumption (iv)) Lipschitzian, this second-order term is of the form

$$-1/2 \sum_{t=1}^n Q_t^2 \phi'(X_t) + 1/2 T_n(\mathbf{X}, \varepsilon) \quad \text{where} \quad |T_n(\mathbf{X}; \varepsilon)| \leq A\theta \sum_{t=1}^n Q_t^3.$$

Now

$$\begin{aligned} \sum_{t=1}^n E Q_t^3 &= \sum_{t=1}^n E | \sum_{u=1}^{t+p-1} g_u^{(n)} X_{t-u} |^3 \leq M \sum_{t=1}^n (\sum_{u=1}^{t+p-1} |g_u^{(n)}|)^3 \\ &\leq nM (\sum_{u=1}^{\infty} |g_u^{(n)}|)^3 \leq M(2pb_M)^3 n^{-1/2}. \end{aligned}$$

Hence  $T_n(\mathbf{X}, \varepsilon)$  is  $o_p$ . Finally,

$$\begin{aligned} -1/2 \sum_{t=1}^n Q_t^2 \phi'(X_t) &= -1/2 \sum_{u=1}^p (g_u^{(n)})^2 \sum_{t=u-p+1}^n X_{t-u}^2 \phi'(X_t) \\ &\quad - 1/2 \sum_{u=p+1}^{n+p-1} (g_u^{(n)})^2 \sum_{t=u-p+1}^n X_{t-u}^2 \phi'(X_t) \\ &\quad - \sum_{u=2}^{p+1} \sum_{v=1}^{u-1} g_u^{(n)} g_v^{(n)} \sum_{t=u-p+1}^n X_{t-u} X_{t-v} \phi'(X_t) \\ &\quad - \sum_{u=p+2}^{n+p-1} \sum_{v=1}^{u-1} g_u^{(n)} g_v^{(n)} \sum_{t=u-p+1}^n X_{t-u} X_{t-v} \phi'(X_t) \\ &= -1/2 \sum_{(21)} - 1/2 \sum_{(22)} - \sum_{(23)} - \sum_{(24)}. \end{aligned}$$

$$\begin{aligned} \sum_{(21)} &= \sum_{u=1}^p (b_u/\sqrt{n})^2 \sum_{t=u-p+1}^n X_{t-u}^2 \phi'(X_t) \\ &\quad + \sum_{u=1}^p (g_u^{(n)} - b_u n^{-1/2})^2 \sum_{t=u-p+1}^n X_{t-u}^2 \phi'(X_t) \\ &\quad + 2 \sum_{u=1}^p (g_u^{(n)} - b_u n^{-1/2}) b_u n^{-1/2} \sum_{t=u-p+1}^n X_{t-u}^2 \phi'(X_t), \end{aligned}$$

but  $g_u^{(n)^2} - b_u^2 n^{-1}$ , as we already know, is  $O(n^{-2})$ , and thus

$$\sum_{(21)} = \sum_{i=1}^p b_i^2 E[X_i^2] E[\phi'(X_t)] + o_p = \sum_{i=1}^p b_i^2 \sigma^2 I(f) + o_p.$$

Also

$$\begin{aligned} E|\sum_{(22)}| &\leq \sum_{u=p+1}^{n+p-1} (g_u^{(n)})^2 (n-u+p) \sigma^2 I(f) \\ &\leq 2(n-1) \sigma^2 I(f) (pb_M n^{-1/2})^{2(p+1)}, \\ E|\sum_{(24)}| &\leq (E|X_t|)^2 I(f) (n+p-2) \sum_{u=2}^{n+p-1} \sum_{v=1}^{u-1} |g_u^{(n)} g_v^{(n)}| \\ &\leq (E|X_t|)^2 I(f) (n+p-2) 4(pb_M n^{-1/2})^4, \end{aligned}$$

and, since  $g_u^{(n)} g_v^{(n)}$  is  $O(n^{-1})$ ,  $\sum_{(23)}$  is  $o_p$ .

So far, we have proved that  $\log \lambda^{(n)}(\mathbf{X}, \varepsilon) = \log \mathcal{L}_0^{(n)}(\mathbf{X}) + \mathcal{L}_n^0(\mathbf{X}) - (d^2/2) + R_n(\mathbf{X}, \varepsilon)$  where  $R_n(\dots)$  is  $o_p$ . Since  $L^{(n)}$  is again a conditional expectation, viz.

$$L^{(n)} = E_\varepsilon[\exp(\log \lambda^{(n)}(\mathbf{X}, \varepsilon) - \log \mathcal{L}_0^{(n)}(\mathbf{X})) | \mathbf{X}],$$

the proof follows as in the AR( $p$ ) case.

**A2.3. The ARMA( $p_1, p_2$ ) case.** We briefly sketch here the proof; it is easy to see that it will lead to the same problems as in the AR and MA cases treated above.

Assume  $p_2 \leq p_1 = p$ . Denoting by  $\lambda^{(n)}(\mathbf{X}; \mathbf{X}_0; \varepsilon)$ , the integrand in (3.3), we have

$$\begin{aligned} \log \lambda^{(n)} &= \sum_{t=1}^n \log f(X_t) \\ &+ \sum_{t=1}^n \phi(X_t) [\sum_{i=1}^p X_{t-i} (n^{-1/2} a_i - g_i^{(n)}) + n^{-1/2} \sum_{u=1}^p g_u^{(n)} \sum_{i=1}^p a_i X_{t-u-i} \\ &\quad - \sum_{u=p+1}^{t-1} g_u^{(n)} (X_{t-u} - n^{-1/2} \sum_{i=1}^p a_i X_{t-u-i}) - \sum_{u=t}^{t+p_2-1} g_u^{(n)} \varepsilon_{t-u}] \\ &- 1/2 \sum_{t=1}^n [n^{-1/2} \sum_{i=1}^p a_i X_{t-i} - \sum_{u=1}^{t-1} g_u^{(n)} (X_{t-u} - n^{-1/2} \sum_{i=1}^p a_i X_{t-u}) \\ &\quad - \sum_{u=t}^{t+p_2-1} g_u^{(n)} \varepsilon_{t-u}]^2 \phi'(X_t - \theta[\dots]). \end{aligned}$$

The term  $\sum_{t=1}^n \phi(X_t) \sum_{i=1}^p X_{t-i} (n^{-1/2} a_i - g_i^{(n)})$  is of the form (3.5), up to  $o_p$  terms. The second-order term converges in probability to  $-1/2 [\sum_{i=1}^{p_2} (a_i + b_i)^2 + \sum_{i=p_2+1}^p a_i^2]$ . The proof follows.  $\square$

### APPENDIX 3

*Some properties of linear serial rank statistics.*

**LEMMA 2.** *The variance of  $(n-p)S_n = \sum_{i=p+1}^n a(R_i^{(n)}, R_{i-1}^{(n)}, \dots, R_{i-p}^{(n)})$  is*

$$\begin{aligned} D^2((n-p)S_n) &= (n-p) \text{Var}(a(R_{p+1}^{(n)}, \dots, R_1^{(n)})) \\ &+ 2 \sum_{i=1}^p (n-p-i) \text{Cov}(a(R_{p+1+i}^{(n)}, \dots, R_{1+i}^{(n)}), a(R_{p+1}^{(n)}, \dots, R_1^{(n)})) \\ &+ [(n-3p)(n-3p-1) + p(2n-5p-1)] \\ &\quad \cdot \text{Cov}(a(R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)}), a(R_{p+1}^{(n)}, \dots, R_1^{(n)})). \end{aligned}$$

COROLLARY 2.

$$D^2((n-p)S_n) \leq \text{Var}(a(R_{p+1}^{(n)}, \dots, R_1^{(n)}))(n-p)(2p+1) + n \text{Cov}(a(R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)}), a(R_{p+1}^{(n)}, \dots, R_1^{(n)})).$$

PROOF. The proof is straightforwardly obtained by expanding  $E[\sum a(\dots)]^2$ .  $\square$

LEMMA 3.

$$E[a(R_{p+1}^{(n)}, \dots, R_1^{(n)})] = \frac{(n-p-1)(n-p-2) \dots (n-2p-1)}{n(n-1) \dots (n-p)} E[a(R_{p+1}^{(n)}, \dots, R_1^{(n)}) | R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)}] + \sum_{\ell=1}^{p+1} \frac{(n-p-1)(n-p-2) \dots (n-2p-1+\ell)}{n(n-1) \dots (n-p)} \sum_{p+2 \leq j_1 \neq \dots \neq j_\ell \leq 2p+2} \sum_{1 \leq k_1 < k_2 < \dots < k_\ell \leq p+1} E[a(R_{p+1}^{(n)}, \dots, R_{k_\ell+1}^{(n)} R_{j_\ell}^{(n)} R_{k_\ell-1}^{(n)}, \dots, R_{k_1+1}^{(n)} R_{j_1}^{(n)} R_{k_1-1}^{(n)}, \dots, R_1^{(n)}) | R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)}].$$

PROOF. Again, the proof follows from a tedious enumeration of the  $n(n-1) \dots (n-p)$  terms of  $\sum_{1 \leq i_1 \neq \dots \neq i_p \leq n} a(i_1, \dots, i_p)$ , which contains  $(n-p-1) \dots (n-2p-1)$  terms with  $\# \{i_1 \neq \dots \neq i_p\} \cap \{R_{p+2}^{(n)}, \dots, R_{2p+2}^{(n)}\} = \emptyset$ ,  $(n-p-1), \dots, (n-2p)$  terms with  $\# \{i_1 \neq \dots \neq i_p\} \cap \{R_{p+2}^{(n)}, \dots, R_{2p+2}^{(n)}\} = 1, \dots, (n-p-1), \dots, (n-2p-1+\ell)$  terms with  $\# \{i_1 \neq \dots \neq i_p\} \cap \{R_{p+2}^{(n)}, \dots, R_{2p+2}^{(n)}\} = \ell, \dots$ , and  $(n-p+1)$  terms with  $\# \{i_1 \neq \dots \neq i_p\} \cap \{R_{p+2}^{(n)}, \dots, R_{2p+2}^{(n)}\} = p+1$ .  $\square$

LEMMA 4.

$$n | \text{Cov}(a(R_{p+1}^{(n)}, \dots, R_1^{(n)}), a(R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)})) | \leq KE[a^2(R_{p+1}^{(n)}, \dots, R_1^{(n)})]$$

(where the constant  $K$  does not depend on  $n$ ).

PROOF. Applying Lemma 3,

$$E[a(R_{p+1}^{(n)}, \dots, R_1^{(n)})a(R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)})] = E[a(R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)})E[a(R_{p+1}^{(n)}, \dots, R_1^{(n)}) | R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)}]] = \frac{n(n-1) \dots (n-p)}{(n-p-1) \dots (n-2p-1)} E^2[a(R_{p+1}^{(n)}, \dots, R_1^{(n)})] - \sum_{\ell=1}^{p+1} [(n-2p-2+\ell) \dots (n-2p-1)]^{-1} \sum_{p+2 \leq j_1 \neq \dots \neq j_\ell \leq 2p+2} \sum_{1 \leq k_1 < \dots < k_\ell \leq p+1} E[a(R_{p+1}^{(n)}, \dots, R_{k_\ell+1}^{(n)} R_{j_\ell}^{(n)}, \dots, R_{j_1}^{(n)} R_{k_1-1}^{(n)}, \dots, R_1^{(n)})a(R_{2p+2}^{(n)}, \dots, R_{p+2}^{(n)})].$$

For every fixed value of  $\ell$ , this latter sum contains

$$((p+1) \dots (p+2-\ell))^2 / (n-2p-2+\ell) \dots (n-2p-1) \ell!$$

terms and since  $|E[a(\dots)a(\dots)]| \leq E[a^2(\dots)]$ , this sum is bounded by

$$\begin{aligned} \sum_{\ell=1}^{p+1} \frac{((p+1) \cdots (p+2-\ell))^2}{(n-2p-2+\ell) \cdots (n-2p-1)\ell!} E[a^2(R_{p+1}^{(n)}, \dots, R_1^{(n)})], \\ = C_n E[a^2(R_{p+1}^{(n)}, \dots, R_1^{(n)})], \end{aligned}$$

where  $C_n$  is  $O(n^{-1})$ . Thus

$$\begin{aligned} \text{Cov}[a(R_{p+1}^{(n)}, \dots, R_1^{(n)}), a(R_{2p+2}^{(n)}, \dots, R_{p+1}^{(n)})] \\ \leq \left( \frac{n(n-1) \cdots (n-p)}{(n-p-1) \cdots (n-2p-1)} - 1 \right) (Ea(R_{p+1}^{(n)}, \dots, R_1^{(n)}))^2 \\ + C_n E[a^2(R_{p+1}^{(n)}, \dots, R_1^{(n)})] \\ \leq C'_n E[a^2(R_{p+1}^{(n)}, \dots, R_1^{(n)})], \end{aligned}$$

with  $C'_n$  another  $O(n^{-1})$ .  $\square$

#### APPENDIX 4

**PROOF OF PROPOSITION 4.1.** The  $g_1$  function in Yoshihara's (1976) Theorem 1 is here

$$g_{\alpha\beta}(\mathbf{Y}_t) = \int_{[0,1]^{p(p+1)}} \Phi_{\alpha\beta}(\mathbf{Y}_t, \mathbf{y}_2, \dots, \mathbf{y}_{p+1}) d\mathbf{y}_2 \cdots d\mathbf{y}_{p+1},$$

where  $d\mathbf{y}_i$  stands for  $dy_{i,1}, dy_{i,2}, \dots, dy_{i,p+1}$ . Since

$$\begin{aligned} \int_{[0,1]^{p(p+1)}} \Phi^{\mathcal{L}}(\mathbf{Y}_t, \mathbf{y}_2, \dots, \mathbf{y}_{p+1}) d\mathbf{y}_2 \cdots d\mathbf{y}_{p+1} \\ = (p+1)^{-1}(G(\mathbf{Y}_t) + pEG(\mathbf{Y}_1)) = (p+1)^{-1}\phi(F^{-1}(U_t)) \sum_{i=1}^p d_i F^{-1}(U_{t-i}), \\ \int_{[0,1]^{p(p+1)}} \Phi^{\mathcal{J}}(\mathbf{Y}_t, \mathbf{y}_2, \dots, \mathbf{y}_{p+1}) d\mathbf{y}_2 \cdots d\mathbf{y}_{p+1} \\ = (p+1)^{-1}(J(\mathbf{Y}_t) + p \int_{[0,1]^{p+1}} J(u_{p+1}, \dots, u_1) du_1 \cdots du_{p+1}), \end{aligned}$$

and

$$\begin{aligned} \int \Phi^{\mathcal{E}}(\mathbf{Y}_t, \mathbf{y}_2, \dots, \mathbf{y}_{p+1}) d\mathbf{y}_2 \cdots d\mathbf{y}_{p+1} \\ = \frac{1}{(p+1)} \sum_{k=1}^{p+1} \int_{[0,1]^p} J(u_p, \dots, u_k, U_t, u_{k-1}, \dots, u_1) du_1 \cdots du_p, \end{aligned}$$

we obtain

$$\begin{aligned} g_{\alpha\beta}(\mathbf{Y}_t) &= g_{\alpha\beta}(U_t, \dots, U_{t-p}) \\ &= (p+1)^{-1}[\alpha J^*(U_t, \dots, U_{t-p}) + \beta\phi(F^{-1}(U_t)) \sum_{i=1}^p d_i F^{-1}(U_{t-i})]. \end{aligned}$$

Clearly,  $E[g_{\alpha\beta}(\mathbf{Y}_t)] = 0$ ; the variance of  $g_{\alpha\beta}(\mathbf{Y}_t)$  is

$$D^2(g_{\alpha\beta}(\mathbf{Y}_t)) = (p + 1)^{-2} \left\{ \alpha^2 \int_{[0,1]^{p+1}} \mathbf{J}^{*2}(v_{p+1}, \dots, v_1) dv_1 \cdots dv_{p+1} + 2\alpha\beta \sum_{i=1}^p d_i \int_{[0,1]^{p+1}} \mathbf{J}^*(v_{p+1}, \dots, v_1) \phi(F^{-1}(v_{p+1})) F^{-1}(v_{p+1-i}) dv_1 \cdots dv_{p+1} + \beta^2 \sum_{i=1}^p d_i^2 \sigma^2 I(f) \right\};$$

the covariances between  $g_{\alpha\beta}(\mathbf{Y}_t)$  and  $g_{\alpha\beta}(\mathbf{Y}_{t+j})$ ,  $1 \leq j \leq p$ , are

$$E[g_{\alpha\beta}(\mathbf{Y}_t) g_{\alpha\beta}(\mathbf{Y}_{t+j})] = (p + 1)^{-2} \left[ \alpha^2 \int_{[0,1]^{p+1+j}} \mathbf{J}^*(v_{p+1}, \dots, v_1) \mathbf{J}^*(v_{p+1+j}, \dots, v_{1+j}) dv_1 \cdots dv_{p+1+j} + \alpha\beta \sum_{i=1}^{p-j} d_i \int_{[0,1]^{p+1}} F^{-1}(v_{p+1-i}) \phi(F^{-1}(v_{p+1})) \mathbf{J}^*(v_{j+1+p}, \dots, v_{j+1}) dv_{j+1} \cdots dv_{j+1+p} \right]$$

(with the convention that, for  $j = p$ , the sum  $\sum_{i=1}^{p-j}$  vanishes). On account of assumption (iii) and (4.1),

$$\begin{aligned} \sigma_{\alpha\beta}^2 &= D^2(g_{\alpha\beta}(\mathbf{Y}_t)) + 2 \sum_{j=1}^p E[g_{\alpha\beta}(\mathbf{Y}_t) g_{\alpha\beta}(\mathbf{Y}_{t+j})] \\ &= (p + 1)^{-2} \{ \alpha^2 V^2 + 2\alpha\beta \sum_{i=1}^p d_i C_i + \beta^2 \sum_{i=1}^p d_i^2 \sigma^2 I(f) \} \end{aligned}$$

(with  $V^2$  and  $C_i$  given in (4.13) and (4.14)) is never zero; the conditions of Yoshihara's (1976) Theorem 1 are therefore satisfied; and  $n^{1/2}(\mathcal{Z}_{\alpha\beta}^n - E\mathcal{Z}_{\alpha\beta}^n)$  is asymptotically normal, with mean zero and variance  $(p + 1)^2 \sigma_{\alpha\beta}^2$ . This completes the proof, since

$$n^{1/2}(\mathcal{Z}_{\alpha\beta}^n - E\mathcal{Z}_{\alpha\beta}^n) = \alpha\sqrt{n}(S_n - m_n) + \beta(\log L_n + 1/2 \sum d_i^2 \sigma^2 I(f)) + o_p. \quad \square$$

### APPENDIX 5

PROOF OF LEMMA 5.1. It is easy to see that

$$\{ \sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}); i = 1, \dots, p \}$$

constitutes a  $p$ -tuple of  $L^2$ -orthogonal functions, with norms

$$[(p + 1 - i) \sigma^2 I(f)]^{1/2}, i = 1, \dots, p.$$

Any square-integrable score-generating function  $\mathbf{J}^*$  can therefore be decomposed into  $\mathbf{J}^*(v_{p+1}, \dots, v_1) = \mathbf{J}_0^*(v_{p+1}, \dots, v_1) + \mathbf{J}_1^*(v_{p+1}, \dots, v_1)$ , where  $\mathbf{J}_0^*$  is given by (5.3) and

$$(A6) \quad \int \mathbf{J}_1^*(v_{p+1}, \dots, v_1) \sum_{j=0}^{p-i} \phi(F^{-1}(v_{p+1-j})) F^{-1}(v_{p+1-j-i}) dv_{p+1}, \dots, dv_1 = 0, \quad i = 1, \dots, p.$$



Denote by  $S_n, S_n^0,$  and  $S_n^\perp$  linear serial rank statistics associated with  $J^*, J_0^*,$  and  $J_\perp^*$ , respectively; denote by  $V^2, V_0^2,$  and  $V_\perp^2$  the asymptotic variances of  $n^{1/2}(S_n - m_n), n^{1/2}(S_n^0 - m_n^0)$  and  $n^{1/2}(S_n^\perp - m_n^\perp),$  respectively. Then, it is easy to check that

$$n^{1/2}(S_n - m_n) = n^{1/2}(S_n^0 - m_n^0) + n^{1/2}(S_n^\perp - m_n^\perp) + o_p(1),$$

with  $\lim_{n \rightarrow \infty} nE[(S_n^0 - m_n^0)(S_n^\perp - m_n^\perp)] = 0$  (under  $H_0^{(n)}$ ).

Thus, we have  $V^2 = V_0^2 + V_\perp^2.$  Now, since using  $J^*$  or  $J_0^*$  in (4.14) leads to the same values of the  $C_i$ 's, we have

$$e(S_n, S_n^0) = \frac{(\sum_{i=1}^p d_i C_i)^2 / (V_0^2 + V_\perp^2)}{(\sum_{i=1}^p d_i C_i)^2 / V_0^2} \leq 1,$$

which completes the proof.  $\square$

### APPENDIX 6

*Asymptotic normality of  $(n^{1/2}r_1^{(n)}, \dots, n^{1/2}r_p^{(n)}, \log L_n)'$ .*

**PROOF OF PROPOSITION 5.2.** Consider the linear combination (with  $\sum_{i=1}^p \alpha_i^2 = 1$ )

$$\begin{aligned} & \sqrt{n} \sum_{i=1}^p \alpha_i r_i^{(n)} + \beta \log L^n \\ &= [\sigma^{-2} \sum_{i=1}^p \alpha_i n^{-1/2} (n-k)^{-1} \sum_{t=i+1}^n X_t X_{t-i} + \mathcal{L}_n^0(\mathbf{X})] - 1/2 \beta^2 \sigma^2 I(f) \sum_{i=1}^p d_i^2 \\ (A7) \quad & - \sum_{i=1}^p \alpha_i (\sigma^{-2} - ((1/n) \sum_{t=1}^n X_t^2)^{-1}) (n^{1/2} / (n-i)) n^{1/2} \sum_{t=i+1}^n X_t X_{t-i} + o_p \\ &= \sum_{(1)} - 1/2 \beta^2 \sum_i d_i^2 \sigma^2 I(f) - \sum_{(2)}. \end{aligned}$$

Since  $n^{1/2} \sum X_t X_{t-k}$  is asymptotically normal, and since  $(\sigma^{-2} - ((1/n) \sum_t X_t^2)^{-1})$  converges to 0 in probability,  $\sum_{(2)}$  is  $o_p.$  Hence

$$\sum_{(1)} = n^{-1/2} \sum_{t=1}^{n-p} X_t [\sigma^{-2} \sum_{i=1}^p \alpha_i X_{t+i} + \beta d_i \phi X_{t+i}] + o_p = n^{-1/2} \sum_{t=1}^{n-p} Z_t + o_p.$$

An immediate computation gives

$$(A8) \quad E(Z_t^2) = \sum_{i=1}^p \alpha_i^2 + \beta^2 \sum_{i=1}^p d_i^2 \sigma^2 I(f) + 2\beta \sum_{i=1}^p d_i \alpha_i$$

and

$$E(Z_t Z_{t+k}) = 0.$$

Consequently, (A7) is asymptotically normal, with mean  $-1/2 \beta^2 \sum d_i^2 \sigma^2 I(f)$  and variance (A8). This proves Proposition 5.2.  $\square$

Corollary 5.1 is an immediate consequence of LeCam's third lemma and Corollary 5.2 follows from the Cauchy-Schwarz inequality.  $\square$

**Acknowledgement.** The authors gratefully acknowledge the valuable comments and suggestions of the Editor, Professor Michael Perlman, the Associate Editor and the referee.

## REFERENCES

- AIYAR, R. J. (1981). Asymptotic efficiency of rank tests of randomness against autocorrelation. *Ann. Inst. Statist. Math.* **33 A** 255–262.
- ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.
- BARTELS, R. (1982). The rank version of von Neumann's ratio test for randomness. *J. Amer. Statist. Assoc.* **77** 40–46.
- BELL, C. B., WOODROOFE, M. and AVADHANI, T. V. (1970). Some nonparametric tests for stochastic processes. In *Nonparametric Techniques in Statistical Inference*. (M. L. Puri, ed.) 215–258. Cambridge University Press.
- DUFOUR, J.-M. (1982). Rank tests for serial dependence. *J. Time Series Anal.* **2** 117–128.
- DUFOUR, J.-M., LEPAGE, Y. and ZEIDAN, H. (1982). Nonparametric testing for time series: a bibliography. *Canad. J. Statist.* **10** 1–38.
- GOVINDARAJULU, Z. (1983). Rank tests for randomness against autocorrelated alternatives. In *Time Series Analysis: Theory and Practice 4*. (O. D. Anderson, ed.) 65–73. North-Holland, Amsterdam and New York.
- GOVINDARAJULU, Z. and DWASS, M. (1983). Simultaneous tests for randomness and location of symmetry. *Proc. Second Prague Conference on Asymptotic Statistics*, 181–197. Hradec Karlove, Czechoslovakia.
- GUPTA, G. D. and GOVINDARAJULU, Z. (1980). Nonparametric tests of randomness against autocorrelated normal alternatives. *Biometrika* **67** 375–379.
- HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests*. Academic, New York.
- HALLIN, M. (1984). Spectral factorization of nonstationary moving average processes. *Ann. Statist.* **12** 172–192.
- HALLIN, M., INGENBLEEK, J.-FR., and PURI, M. L. (1985). Tests de Rangs Linéaires pour une Hypothèse de Bruit Blanc. *C. R. Acad. Sc. Paris* **301** Série 1, 49–52.
- HOTELLING, H. and PABST, M. R. (1936). Rank correlation and tests of significance involving no assumption of normality. *Ann. Math. Statist.* **7** 29–43.
- INGENBLEEK, J.-FR. (1980). *Statistiques de rangs sérielles univariées et multivariées*. Unpublished D.Sc. dissertation, Université Libre de Bruxelles, Brussels.
- JOGDEO, K. (1968). Asymptotic normality in nonparametric methods. *Ann. Math. Statist.* **39** 905–922.
- KENDALL, M. and STUART, A. (1968). *The Advanced Theory of Statistics*. Vol. 3. Griffin, London.
- KNOKE, D. J. (1977). Testing for randomness against autocorrelation: alternative tests. *Biometrika* **64** 523–529.
- LECAM, L. (1960). Locally asymptotically normal families of distributions. *Univ. Calif. Publ. Statist.* **3** 37–98.
- MILLER, K. S. (1968). *Linear Difference Equations*. Benjamin, New York.
- PURI, M. L. and SEN, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*. Wiley, New York.
- YOSHIHARA, K. (1976). Limiting behavior of  $U$ -statistics for stationary absolutely regular processes. *Z. Wahrsch. verw. Gebiete* **35** 237–252.
- WALD, A. and WOLFOWITZ, J. (1943). On exact tests for randomness in the nonparametric case based on serial correlation. *Ann. Math. Statist.* **14** 378–388.
- WOLD, H. (1954). *A Study in the Analysis of Stationary Time Series*. Almqvist and Wiksell, Stockholm.

M. HALLIN  
 J.-FR. INGENBLEEK  
 INSTITUT DE STATISTIQUE  
 C.P. 210 - CAMPUS DE LA PLAINE  
 UNIVERSITÉ LIBRE DE BRUXELLES  
 B1050 BRUXELLES  
 BELGIUM

M. L. PURI  
 DEPARTMENT OF MATHEMATICS  
 SWAIN HALL EAST  
 INDIANA UNIVERSITY  
 BLOOMINGTON, INDIANA 47405