The Efficiency of some Nonparametric Rank-Based Competitors to Correlogram Methods

Marc Hallin *and Olivier Tribel †

I.S.R.O., E.C.A.R.E., and Département de Mathématique
Université Libre de Bruxelles
Campus de la Plaine C.P. 210
B 1050 Bruxelles
BELGIUM

Abstract

Hodges and Lehmann (1956) have shown that the asymptotic relative efficiency of Wilcoxon rank tests with respect to Student's t tests, in location models with independent observations, never falls below 0.864—a lower bound which is attained at a parabolic density related with the so-called Epanechnikov kernel. This result actually holds for under general linear models with independent observations. A similar result is proved here, in a time series context, for the so-called Spearman-Wald-Wolfowitz autocorrelation coefficients. It is shown that the asymptotic relative efficiency of the corresponding tests, with respect to the classical everyday practice based on traditional autocorrelations, is never less than 0.856; the bound is attained at a cosine density.

*I.S.R.O., E.C.A.R.E.S., and Département de Mathématique. Research supported by an A.R.C. contract of the Communauté française de Belgique, the European Human Capital contract ERB CHRX-CT 940693 and the Fonds d'Encouragement à la Recherche de l'Université Libre de Bruxelles.

†Département de Physique. Research supported by a F.R.I.A. grant.

1 Hodges and Lehmann's .864 bound

Hodges and Lehmann (1956), in a famous paper that had quite an impact in the history of rank-based inference, established the striking fact that the asymptotic relative efficiency (ARE) of Wilcoxon tests (this includes the one-sample signed rank test, the two-sample Wilcoxon test, as well as the Kruskal-Wallis test for one-way analysis of variance) with respect to their normal-theory competitors (one- and two-sample Student tests and F-tests) never falls below 0.864. This result apparently was unexpected, and caused quite a surprise in the statistical community. As stated by E.L. Lehmann himself in his 1984 *Statistical Science* interview (De Groot, 1986), "*this paper was influential in the sense that it dispelled the belief that while nonparametric [rank-based] techniques are very convenient because you don't have many assumptions, they have so little power that they are no good*".

Interestingly enough, thirty years later, E.L. Lehmann still considers his ".864 *paper with Joe*" [Hodges] as one of his favourite. Along with the Chernoff-Savage (1958) result—showing that the ARE, with respect to the same normal-theory competitors, of normal-score rank and signed rank tests is never less than one—the ".864 *paper*" certainly played a triggering role in the subsequent developments of rank-based inference.

In a more precise setting, Hodges and Lehmann's result can be described as follows. Consider the linear rank statistic (of the Wilcoxon type)

$$S^{(n)} \stackrel{def}{=} \sum_{i=1}^{n} \left(c_i^{(n)} - \bar{c}^{(n)} \right) R_i^{(n)}, \tag{1.1}$$

where

(i) $c_i^{(n)}$, $i = 1, \ldots, n$ denotes a triangular array of *regression* constants, with mean $\bar{c}^{(n)} \stackrel{def}{=} \sum_{i=1}^{n} c_i^{(n)}$

(ii) $\mathbf{R}^{(n)} = (R_1^{(n)}, \ldots, R_n^{(n)})$ is the vector of ranks computed from some n-tuple of *residuals* $\mathbf{Z}^{(n)} = (Z_1^{(n)}, \ldots, Z_n^{(n)})$ which, under the null hypothesis to be tested, reduce to *white noise*—an i.i.d. sequence with unspecified density function.

Similarly, define the linear signed rank statistic (of the Wilcoxon type)

$$S_+^{(n)} \stackrel{def}{=} \sum_{i=1}^{n} c_i^{(n)} \mathrm{sgn}(Z_i^{(n)}) R_{+;i}^{(n)}, \tag{1.2}$$

where

(iv) $\mathrm{sgn}(Z_i^{(n)})$ is the sign of $Z_i^{(n)}$, and

(v) $\mathbf{R}_+^{(n)} = (R_{+;1}^{(n)}, \dots, R_{+;n}^{(n)})'$ is the vector of ranks computed from the absolute values of the $Z_i^{(n)}$'s.

The normal-theory statistic associated with (1.1) and (1.2) is

$$T^{(n)} \stackrel{def}{=} \sum_{i=1}^{n} \left(c_i^{(n)} - \bar{c}^{(n)}\right) \frac{Z_i^{(n)}}{\sigma^{(n)}}, \tag{1.3}$$

where

(vi) $(\sigma^{(n)})^2 \stackrel{def}{=} n^{-1} \sum_{i=1}^{n} \left(Z_i^{(n)} - n^{-1} \sum_{i=1}^{n} Z_i^{(n)}\right)^2$ denotes the empirical variance of the residuals.

Under very general assumptions on the regression constants (see Hájek and Šidák 1965) and the density f of the residuals (which, for the signed-rank statistic (1.2), always include symmetry with respect to the origin), $S^{(n)}$, $S_+^{(n)}$, and $T^{(n)}$, after suitable standardization, are asymptotically normal, as $n \to \infty$, under any sequence of null hypotheses for which the residuals $Z_i^{(n)}$ are white noise.

These distributional results can be extended, under adequate conditions, to *aligned rank* statistics (see, e.g., Puri and Sen 1985), and also to *regression rank score* statistics (Gutenbrunner, Jurečková, Koenker, and Portnoy 1993), yielding a complete toolkit of rank tests of the Wilcoxon type which basically allow for testing arbitrary sets of linear constraints on the parameters of general linear models with independent errors. The asymptotic relative efficiency (ARE), under error density f and with respect to the corresponding normal-theory tests, of the Wilcoxon rank tests based on (1.1) or (1.2), or based on quadratic forms involving vectors of such statistics (cf. the Kruskal-Wallis test for one-way analysis of variance) is

$$12 \left[\int_{-\infty}^{\infty} f_1^2(x)\mathrm{d}x\right]^2, \tag{1.4}$$

where f_1 stands for the standard version of f.

The Hodges-Lehmann result then states that

$$\inf_f 12 \left[\int_{-\infty}^{\infty} f_1^2(x)\mathrm{d}x\right]^2 = 108/125 = 0.864, \tag{1.5}$$

a lower bound which is attained at the *parabolic* density

$$f_0(x) \stackrel{def}{=} \frac{3\sqrt{5}}{20}\left(5 - x^2\right) I\left[-\sqrt{5} \leq x \leq \sqrt{5}\right]. \tag{1.6}$$

This *parabolic* density also enjoys remarquable properties, as a kernel, in density estimation problem, where it is known as the Epanechnikov kernel (Epanechnikov 1969). Actually, such densities can be traces back as far as Lagrange (1776) and Daniel Bernoulli (1778). Lagrange in a long paper had it as one density that he could cope with via a "Laplace transform"; see page 228 of Volume 2 of Lagrange's *Oeuvres* (1868). Bernoulli (1778) used it as a density to find a MLE; see Kendall (1961), and Stigler (1997).

Two simple applications of Hodges and Lehmann's result are

(a) the *one-sample location model*, where a n-tuple of i.i.d. observations $X_1^{(n)}, \ldots, X_n^{(n)}$ has unspecified density, symmetric with respect to some $\mu \in \mathbb{R}$, and the null hypothesis to be tested is $\mu = \mu_0$; letting $Z_i^{(n)} \stackrel{def}{=} X_i^{(n)} - \mu_0$ and $c_i^{(n)} = 1$, a test can be based on (1.2), yielding the classical Wilcoxon signed rank test, or on (1.3), which, after adequate standardization, reduces to the classical Student test statistic;

(b) the *two-sample location model*: under the null hypothesis to be tested, a n-tuple of independent observations $X_1^{(n)}, \ldots, X_m^{(n)}, \ldots, X_n^{(n)}$ are i.i.d., whereas, under the alternative, $X_1^{(n)}, \ldots, X_m^{(n)}$ and $X_{m+1}^{(n)}, \ldots, X_n^{(n)}$ differ in location; letting $Z_i^{(n)} \stackrel{def}{=} X_i^{(n)}$, $c_i^{(n)} = 1$, $i = 1, \ldots, m$ and $c_i^{(n)} = 0$, $i = m+1, \ldots, n$, a test can be based on (1.1), yielding (after due standardization) the classical Wilcoxon rank sum test, or on (1.3), which (still, after standardization) reduces to two-sample Student statistic.

The intuitive interpretation of (1.5) in these two cases is that using the Wilcoxon tests instead of Student's traditional t-tests never entails a serious loss of efficiency: the *cost*, in terms of additional observations, indeed never exceeds $1 - 0.864 = 13.6\%$. On the other hand, the *benefit*, still in terms of sample sizes, can be infinite.

2 Spearman correlograms

Rank-based inference for a long time has been essentially limited to the context of independent observations. The invariance arguments (invariance here is with respect to the group of order-preserving transformations acting on residuals) leading to rank-based methods in linear models with independent error terms however still hold under much more general situations under which the observed series constitutes the realization of some stochastic process, and, more particularly, in time series analysis: see Hallin and Puri (1992), or Hallin and Werker (1999) for a review. A systematic treatment of rank and signed rank testing methods for autoregressive-moving average (ARMA) models has been developed in a series of papers by Hallin and Puri (1988, 1991, 1994). In the presentation given here, we avoid giving precise lists of technical assumptions, for which the interested reader is referred to the literature.

The normal theory in linear time series models leads to test statistics which are linear or quadratic forms in residual autocorrelations, of the form

$$r_k^{(n)} \stackrel{def}{=} (n-k)^{-1} \sum_{t=k+1}^{n} \frac{Z_t^{(n)} Z_{t-k}^{(n)}}{(\sigma^{(n)})^2}, \tag{2.1}$$

where

(i) $Z_t^{(n)}$, $t = 1, \ldots, n$ again denotes a n-tuple of residuals which, under the null hypothesis to be tested (linear constraints on the parameters of the underlying ARMA model), are centered white noise, and

(ii) $(\sigma^{(n)})^2$ is defined as in (1.3), so that $(n-k)^{1/2} r_k^{(n)}$ is asymptotically standardized.

Under Gaussian assumptions, residual correlograms indeed are *locally asymptotically sufficient* in the sense of Le Cam (namely, $(r_k^{(n)}, k = 1, 2, \ldots)$-measurable central sequences exist), so that locally asymptotically optimal tests can be based on residual autocorrelations.

Rank tests for the same problems can be based on a rank-based generalization of (2.1), of the form

$$\underset{\sim}{r}{}_{J_1 J_2; k}^{(n)} \stackrel{def}{=} (n-k)^{-1} \left[\sum_{t=k+1}^{n} J_1\left(\frac{R_t^{(n)}}{n+1}\right) J_2\left(\frac{R_{t-k}^{(n)}}{n+1}\right) - m_{J_1 J_2}^{(n)} \right] (s_{J_1 J_2}^{(n)})^{-1}, \tag{2.2}$$

where

(iii) $R_t^{(n)}$, $t = 1, \ldots, n$ denotes the rank of $Z_t^{(n)}$ among $Z_t^{(n)}$, $t = 1, \ldots, n$,

(iv) J_1 and J_2: (0, 1) $\to \mathbb{R}$ are nondecreasing, continuous *score functions* satisfying

$$\int_0^1 J_i^2(u)\mathrm{d}u < \infty,\ i = 1, 2,$$

(v) $m^{(n)}_{J_1 J_2}$ and $s^{(n)}_{J_1 J_2}$ are exact standardizing constants.

When the underlying innovation density f can be assumed symmetric, signed rank residual autocorrelation coefficients also can be substituted for the unsigned ones (2.2). These signed rank autocorrelations are of the general form

$$\underset{\sim}{r}^{(n)}_{+;J_1^+J_2^+;k} \stackrel{def}{=} (n-k)^{-1} \times \left[\sum_{t=k+1}^{n} J_1^+\left(\frac{\mathrm{sgn}(Z_t^{(n)})R^{(n)}_{+;t}}{n+1}\right) J_2^+\left(\frac{\mathrm{sgn}(Z_{t-k}^{(n)})R^{(n)}_{+;t-k}}{n+1}\right)\right] \times \left(s^{(n)}_{+;J_1^+J_2^+}\right)^{-1}, \tag{2.3}$$

where

(vi) the score functions J_1^+ and J_2^+: $(-1, 1) \to \mathbb{R}$ are odd and square-integrable,

(vii) the ranks $R^{(n)}_{+;t}$ are those of the absolute values of the residuals $Z_t^{(n)}$, and

(viii) $s^{(n)}_{+;J_1 J_2}$ is an exact standardizing constant.

The score functions J_1 and J_2 (or J_1^+ and J_2^+) can be adjusted to the underlying innovation density f in such a way that the corresponding correlograms ($\underset{\sim}{r}^{(n)}_{J_1J_2;k}$ or $\underset{\sim}{r}^{(n)}_{+;J_1^+J_2^+;k}$), just as the traditional one $(r_k^{(n)})$ under Gaussian densities, be *locally asymptotically sufficient* in Le Cam's sense. This yields signed or unsigned normal-score (van der Waerden), Wilcoxon or Laplace rank-based correlograms, associated with Gaussian, logistic or double-exponential innovation densities, respectively. Unsigned Wilcoxon correlograms, for instance, are characterized by $J_1(u) = u$—the usual Wilcoxon score function—and $J_2(v) = F_{\mathcal{L}}^{-1}(v)$, where $F_{\mathcal{L}}$ stands for the standard logistic distribution function.

Alternative, much simpler, extensions of the nonserial Wilcoxon statistics (1.1) and (1.2) are the Spearman-Wald-Wolfowitz autocorrelations Wolfowitz 1943)

$$\underset{\sim}{r}^{(n)}_{\mathcal{S};k} \stackrel{def}{=} \left[(n-k)^{-1} \sum_{t=k+1}^{n} R_t^{(n)} R_{t-k}^{(n)} - \frac{1}{12}(n+1)(3n+2)\right] \left(s_{\mathcal{S}}^{(n)}\right)^{-1}, \tag{2.4}$$

with exact standardizing constants $(s_{\mathcal{S}}^{(n)})^2$ satisfying $(s_{\mathcal{S}}^{(n)})^2 = (n^4/144) + O(n^3)$ as $n \to \infty$, and the signed Spearman-Wald-Wolfowitz autocorrelations

$$\underset{\sim}{r}^{(n)}_{+;\mathcal{S};k} \stackrel{def}{=} \frac{180 \sum_{t=k+1}^{n} \mathrm{sgn}(Z_t^{(n)} Z_{t-k}^{(n)}) R_{+;t}^{(n)} R_{+;t-K}^{(n)}}{[(n-k)(n+k)(20n^3+24n^2-5n-6)]^{1/2}}. \tag{2.5}$$

Both $(n-k)^{1/2}\, \underset{\sim}{r}^{(n)}_{\mathcal{S};k}$ and $(n-k)^{1/2}\, \underset{\sim}{r}^{(n)}_{+;\mathcal{S};k}$ are exactly standardized, and asymptotically standard normal, when $Z_t^{(n)}$ is white noise; clearly, $\underset{\sim}{r}^{(n)}_{\mathcal{S};k}$ can be interpreted as a serial version of the classical Spearman correlation coefficient for bivariate i.i.d. samples—an idea that first appears in Wald and Wolfowitz (1943), whence the terminology.

When (2.4) or (2.5) (the latter requiring, as usual, a symmetric innovation density) are substituted for the traditional autocorrelations (2.1) optimal Gaussian test statistics (typically, the Gaussian Lagrange multiplier test statistics), the ARE, under innovation density f and with respect to the corresponding Gaussian procedures, of the resulting tests is

$$144 \left[\int_0^1 u F_1^{-1}(u) \mathrm{d}u \int_0^1 u \frac{f_1'(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} \mathrm{d}u\right]^2 \tag{2.6}$$

(f_1 stands for the standard version of f, F_1 for the corresponding distribution function; the technical assumptions on f, which are not explicitated here, include absolute regularity, with a.e. derivative f').

For fixed f, this ARE is strictly smaller than the corresponding nonserial one (1.4). Indeed, integration by parts yields

$$12 \left[\int_0^1 u \frac{f_1'(F_1^{-1}(u))}{f(F_1^{-1}(u))} \mathrm{d}u\right]^2 = 12 \left[\int_{-\infty}^{\infty} F_1(x) f_1'(x) \mathrm{d}x\right]^2 = 12 \left[\int_{-\infty}^{\infty} f_1^2(x) \mathrm{d}x\right]^2,$$

whereas the Cauchy-Schwarz inequality implies

$$12\left[\int_0^1 uF_1^{-1}(u)\mathrm{d}u\right]^2 < 12\int_0^1\left(u-\frac{1}{2}\right)^2\mathrm{d}u\int_{-\infty}^{\infty}x^2 f_1(x)\mathrm{d}x = 1.$$

The infimum, over all possible densities f, of (2.6) thus has to be less than 0.864. A natural question then arises: is this new lower bound substantially less than 0.864? The answer, which is provided in Section 3, fortunately is negative.

3 From .864 to .856

The infimum, over all possible densities f, of the asymptotic relative efficiency, with respect to traditional correlogram-based methods, of testing procedures based on the signed or unsigned Spearman-Wald-Wolfowitz autocorrelation coefficients, is

$$\inf_f 144\left[\int_0^1 uF_1^{-1}(u)\mathrm{d}u\int_0^1 u\frac{f_1'(F_1^{-1}(u))}{f_1(F_1^{-1}(u))}\mathrm{d}u\right]^2 = \left[\frac{3\pi^2}{32}\right]^2 \approx 0.856\,. \qquad (3.1)$$

This infimum is not a minimum: strictly, it is not attained at any density satisfying the required technical assumptions required in the derivation of (2.6) as an ARE. But, if the density f in (2.6) is chosen as the *cosine density*

$$f_1^*(x) \stackrel{def}{=} \frac{\sqrt{\pi^2-8}}{4}\cos\left(\frac{\sqrt{\pi^2-8}}{2}x\right) I\left[-\frac{\pi}{\sqrt{\pi^2-8}} \leq x \leq \frac{\pi}{\sqrt{\pi^2-8}}\right], \qquad (3.2)$$

with

$$\frac{f_1^{*\prime}(x)}{f_1^*(x)} = \frac{\sqrt{\pi^2-8}}{2}\tan\left(\frac{\sqrt{\pi^2-8}}{2}x\right) I[-\frac{\pi}{\sqrt{\pi^2-8}} \leq x \leq \frac{\pi}{\sqrt{\pi^2-8}}], \qquad (3.3)$$

then the integral expression in (2.6) yields $\left[\frac{3\pi^2}{32}\right]^2 \approx 0.856$, which, as we shall see, is the infimum we are looking for. However, (3.3) is not square-integrable, so that the cosine density has infinite Fisher information. And, a finite Fisher information was one of the assumptions required in the derivation and the interpretation of (2.6) as an ARE value. Nevertheless, the cosine density (3.2) can be obtained as the limit of a sequence of densities satisfying all the assumptions required in the derivation of (2.6), and thus can be considered as a limiting case.

The proof of (3.1) relies on simple variational arguments. Put

$$\mathcal{I}(f) \stackrel{def}{=} \int_0^1 uF_1^{-1}(u)\mathrm{d}u \int_0^1 u\frac{f_1'(F_1^{-1}(u))}{f_1(F_1^{-1}(u))}\mathrm{d}u$$
$$= \int_{-\infty}^{\infty} f_1^2(x)\mathrm{d}x \int_{-\infty}^{\infty} xf_1(x)F_1(x)\mathrm{d}x. \tag{3.4}$$

¿From the definition of $\mathcal{I}(f)$, it is clear that the infimum is to be obtained for a density f_1 such that either both F_1^{-1} and $\frac{f_1'}{f_1}$ are symmetric or antisymmetric with respect to $u = \frac{1}{2}$. Since F_1^{-1} is nondecreasing, it only can be antisymmetric. This corresponds to a density f_1 which is symmetric with respect to zero, which automatically implies an antisymmetric $\frac{f_1'}{f_1}$ Hence, we can restrict ourselves to symmetric densities f_1.

The Euler-Lagrange equation associated with (3.4) is

$$-2f_1'(x)\int zf_1(z)F_1(z)\mathrm{d}z - F_1(x)\int f_1^2(z)\mathrm{d}z - 2z\lambda_1 - \lambda_2 = 0; \tag{3.5}$$

λ_1 and λ_2 are the Lagrange multipliers of the problem. In (3.5), let

$$y(x) \stackrel{def}{=} F_1(x), \quad \dot{y}(x) \stackrel{def}{=} f_1(x) \quad \text{and} \quad \ddot{y}(x) \stackrel{def}{=} f_1'(x).$$

Also, define

$$\omega^2 \stackrel{def}{=} \frac{1}{2}\frac{\int f_1^2(z)\mathrm{d}z}{\int zf_1(z)F_1(z)\mathrm{d}z}, \qquad \alpha \stackrel{def}{=} -\frac{2\lambda_1}{\int f_1^2(z)\mathrm{d}z}$$

and

$$\beta \stackrel{def}{=} -\frac{\lambda_2}{\int f_1^2(z)\mathrm{d}z}.$$

Equation (3.5) then takes the form

$$\ddot{y} + \omega^2\dot{y} - \alpha\omega^2 x - \beta\omega^2 = 0, \tag{3.6}$$

the real solutions of which are

$$y = A\sin(\omega x) + B\cos(\omega x) + \alpha x + \beta,$$

with the constraints that y should be nonnegative, monotone nondecreasing, twice differentiable, and such that

$$y(-\infty) = 0, \qquad y(\infty) = 1, \tag{3.7}$$

$$\int z\dot{y}(z)\mathrm{d}z = 0 \quad \text{and} \quad \int z^2\dot{y}(z)\mathrm{d}z = 1. \tag{3.8}$$

Assume that f_1 has a bounded (symmetric) support $[-a;\, a]$. The continuity of $\dot{y} = f_1$ at $\pm a$ implies

$$\cos(\omega a) = -\frac{\alpha}{A\omega} \quad \text{and} \quad B\sin(\omega a) = 0.$$

Condition (3.7) yields $y(-a) = 0$ and $y(a) = 1$, whence $\beta = \frac{1}{2}$, and

$$\sin(\omega a) = \frac{1 - 2\alpha a}{2A}.$$

Due to the symmetry of $\dot{y}$, the first part of condition (3.8) is automatically satisfied; the second part implies that

$$a^2\left(1 - \frac{4}{3}\alpha a\right) = \frac{2}{\omega^2} + 1.$$

Summing up,

$$\begin{array}{ll} \alpha = 0 & \omega^2 = \frac{\pi^2-8}{4} \approx (0.683)^2 \\ A = \frac{1}{2} & a = \sqrt{\frac{\pi^2}{\pi^2-8}} \approx 2.3\,, \end{array}$$

yielding the cosine density f_1^* given in (3.2). It is easily checked that

$$144\left[\int_{-\infty}^{\infty} f_1^{*2}(x)\mathrm{d}x \int_{-\infty}^{\infty} x f_1^*(x) F_1^*(x)\mathrm{d}x\right]^2 = \left[\frac{3\pi^2}{32}\right]^2 \approx 0.856,$$

which establishes the desired result.

4 Conclusions

Rank tests, which are daily practice in biostatistics, psychometrics or the planning of experiments, are almost totally ignored in such other fields as econometrics, hydrology or environmental statistics, where the data essentially take the form of time series. All the attractive properties—distribution-freeness (exact or asymptotic), similarity and unbiasedness, robustness, ...—that make rank tests so attractive in the presence of independent observations appear even more appealing, though, in the time-series context.

The only serious objection against rank-based methods in this area thus would be their eventual low power, when compared with the traditional correlogram-based methods. The power of a rank-based test of course depends on the score functions considered, and on the underlying density (in a time-series context, the innovation density). The result which is proved here shows that, when the very simple Wilcoxon scores are adopted, yielding the so-called Spearman-Wald-Wolfowitz autocorrelations, the ARE of the resulting tests with respect to their traditional counterparts (e.g., Gaussian Lagrange multipliers), based on classical correlograms, is never less than 0.856—a lower bound which is hardly worse than the celebrated Hodges-Lehmann 0.864 obtained for Wilcoxon rank tests in the i.i.d. context. This lower bound is approached when sequences of innovation densities converging to the cosine density (3.2) are considered. Some other ARE values are (see Hallin and Werker 1999): 0.912 under Gaussian densities, 1.000 under logistic densities and 1.226 under double-exponential densities. The upper bound is $+\infty$.

Another result (Hallin 1994), extends to the time-series context the well-known Chernoff-Savage (1958) property of normal-score (van der Waerden) tests, whose ARE efficiency (with respect to Student's t tests in the i.i.d. case, to classical correlogram-based methods in the time-series case) is always larger than or equal to one.

These two results, of an asymptotic nature, are confirmed by small sample simulation studies: see Hallin and Mélard (1988) or Garel and Hallin (1999). The fear that rank-based methods would have low power is thus completely unjustified, quite on the contrary, and rank-based methods definitely deserve entering time-series practice.

Acknowledgement. The authors are grateful to Lucien Le Cam and Stephen Stigler for their historical comments on Laplace and Bernoulli, in connection with the so-called Epanechnikov kernel.

References

[1] Bernoulli, D. (1778). Dijudicatio maxime probabilis plurium observationum discrepantium atque verisimillima inductio inde formanda. *Acta Academiae Scientiarum Imperialis Petropolitanae 1777*, 3-23. Reprinted in Bernoulli (1982). English translation in Kendall (1961), 3-13. Translation reprinted in Pearson and Kendall (1970), 157-167.

[2] Bernoulli, D. (1982). *Die Werke von Daniel Bernoulli: Band 2, Analysis, Wahrscheinlichkeitsrechnung*, Birkhaeuser, Basel.

[3] Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric tests. *Annals of Mathematical Statistics* **29**, 972-994.

[4] DeGroot, M.H. (1986). A conversation with Erich L. Lehmann. *Statistical Science* **1**, 243-258.

[5] Epanechnikov, V. (1969). Nonparametric estimation of a multidimensional probability density. *Theory of Probability and their Applications* **14**, 153-158.

[6] Garel, B. and M. Hallin (1999). Rank-based *AR* order identification. *Journal of the American Statistical Association*, to appear.

[7] Gutenbrunner, C., J. Jurečková, R. Koenker and S. Portnoy (1993). Tests of linear hypotheses based on regression rank scores. *Journal of Nonparametric Statistics* **2**, 307-331.

[8] Hájek, J. and Z. Šidák (1967). *Theory of Rank Tests.* Academic Press, New York.

[9] Hallin, M. (1994). On the Pitman-nonadmissibility of correlogram-based methods. *Journal of Time Series Analysis* **15**, 607-612.

[10] Hallin, M. and G. Mélard (1988). Rank-based tests for randomness against first-order serial dependence. *Journal of the American Statistical Association* **83**, 1117-1129.

[11] Hallin, M. and M.L. Puri (1988). Optimal rank-based procedures for time-series analysis: testing an *ARMA* model against other *ARMA* models. *Annals of Statistics* **16**, 402-432.

[12] Hallin, M. and M.L. Puri (1991). Time-series analysis via rank-order theory: signed-rank tests for *ARMA* models. *Journal of Multivariate Analysis* **39**, 1-29.

[13] Hallin, M. and M.L. Puri (1992). Rank tests for time series analysis : a survey. In *New Directions In Time Series Analysis*, D. Brillinger, E. Parzen and M. Rosenblatt, Eds., Springer-Verlag, New York, 111-154.

[14] Hallin, M. and M.L. Puri (1994). Aligned rank tests for linear models with autocorrelated error terms. *Journal of Multivariate Analysis* **50**, 175-237.

[15] Hallin, M. and B. Werker (1999). Optimal testing for semi-parametric AR models: from Lagrange multipliers to autoregression rank scores and adaptive tests. In *Asymptotics, Nonparametrics and Time Series*, S. Ghosh Ed., M. Dekker, New York, 295-350.

[16] Hodges, J.L. and E.L. Lehmann (1956). The efficiency of some nonparametric competitors of the t-test. *Annals of Mathematical Statistics* **58**, 324-335.

[17] Kendall, M. G. (1961). Daniel Bernoulli on maximum likelihood. *Biometrika* **48**, 1-18. Reprinted in Pearson and Kendall (1970), 155-172.

[18] Lagrange, J.-L. (1776). Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations; dans lequel on examine les avantages de cette méthode par le calcul des probabilités, et où l'on résoud differens problèmes relatifs à cette matière, *Miscellanea Taurinensia* **5**, 167-232. Reprinted in Lagrange (1868) Vol. 2, 173-236.

[19] Lagrange, J.-L. (1868). *Oeuvres de Lagrange*, Vol. 2, Gauthier-Villars, Paris.

[20] Pearson, E. S. and M. G. Kendall, Eds. (1970). *Studies in the History of Statistics and Probability.* Charles Griffin, London.

[21] Puri, M.L. and P.K. Sen (1985). *Nonparametric Methods in General Linear Models.* Wiley, New York.

[22] Stigler, S.(1997). Daniel Bernoulli, Leonhard Euler, and Maximum Likelihood. In *Festschrift for Lucien LeCam*, D. Pollard, E. Torgersen, and G. Yang, Eds., Springer-Verlag, New York, 345-367.

[23] Wald, A. and J. Wolfowitz (1943). An exact test for randomness in the nonparametric case based on serial correlation. *Annals of Mathematical Statistics* **14**, 378-388.