Institute of Mathematical Statistics

LECTURE NOTES — MONOGRAPH SERIES

Kolmogorov-Smirnov Tests for AR Models Based on Autoregression Rank Scores

Faouzi El Bantli[1]
I.S.R.O.
Université Libre de Bruxelles
Brussels, Belgium

Marc Hallin[2]
I.S.R.O., E.C.A.R.E.S.,
and Département de Mathématique
Université Libre de Bruxelles
Brussels, Belgium

Abstract

Tests of the Kolmogorov-Smirnov type are constructed for the parameter of an autoregressive model of order p. These tests are based on autoregression rank scores, and extend to the time-series context a method proposed by Jurečková (1991) for regression rank scores and regression models with independent observations. Their asymptotic distributions are derived, and they are shown to coincide with those of classical Kolmogorov-Smirnov statistics, under null hypotheses as well as under contiguous alternatives. Local asymptotic efficiencies are investigated. A Monte Carlo experiment is carried out to illustrate the performance of the proposed tests.

Keywords: Autoregressive models, Autoregression quantiles, Autoregression rank scores, Kolmogorov-Smirnov test, Local asymptotic efficiency.

1 Introduction

Consider the autoregressive model of order p

$$Y_t = \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \varepsilon_t, \qquad t \in \mathbb{Z} \tag{1.1}$$

($AR(p)$ model), where $p \geq 1$ is a fixed integer, $\boldsymbol{\theta} = (\theta_1, ..., \theta_p)'$ a p-dimensional vector of unknown autoregressive coefficients, and $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ a process of

[1]Research supported by the Fonds d'Encouragement à la Recherche de l'Université Libre de Bruxelles.

[2]Research supported by an A.R.C. contract of the Communauté française de Belgique and the Fonds d'Encouragement à la Recherche de l'Université Libre de Bruxelles.

independent and identically distributed random variables, with unspecified cumulative distribution function F and probability density f satisfying

$$\int x\mathrm{d}F(x) = 0 \quad \text{and} \quad 0 < \sigma_f^2 := \int x^2\mathrm{d}F(x) < \infty. \tag{1.2}$$

Letting

$$\boldsymbol{\Theta} := \left\{ \boldsymbol{\theta} \in \mathbb{R}^p \;\middle|\; z^p - \sum_{i=1}^{p} \theta_i z^{p-i} \neq 0 \; \forall z \in \mathbb{C},\ |z| \leq 1 \right\}, \tag{1.3}$$

we shall assume throughout the paper that $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (the usual *causality* assumption) : $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ then is the innovation process, and f the innovation density, of the $AR(p)$ process $\{Y_t,\ t \in \mathbb{Z}\}$. Denote by $\mathbf{Y}_n := (Y_1, \ldots, Y_n)'$ a realization of length n of some solution of (1.1); we do not require $\mathbf{Y}_n$ to be stationary, since all solutions are *asymptotically stationary* (see Hallin and Werker (1998) for a detailed discussion on this issue). Assume furthermore that $(Y_{-p+1}, \ldots, Y_0)$ also are observed; in case they are not, they safely can be put equal to zero without affecting asymptotic results.

Koul and Saleh (1995), extending to the time-series context Koenker and Bassett (1978)'s concept of regression quantiles, defined the *α-autoregression quantile* for model (1.1) as the solution

$$\hat{\boldsymbol{\vartheta}}^{(n)}(\alpha) := \left(\hat{\theta}_0^{(n)}(\alpha), \hat{\boldsymbol{\theta}}^{(n)}(\alpha)\right), \quad \hat{\theta}_0^{(n)}(\alpha) \in \mathbb{R}, \quad \hat{\boldsymbol{\theta}}^{(n)}(\alpha) \in \mathbb{R}^p$$

of the minimization problem

$$\hat{\boldsymbol{\vartheta}}^{(n)}(\alpha) := \arg\min_{\boldsymbol{\vartheta} \in \mathbb{R}^{p+1}} \sum_{t=1}^{n} \varrho_\alpha \left(Y_t - \theta_0 - \mathbf{x}_{t-1}'\boldsymbol{\theta}\right), \tag{1.4}$$

where the minimum is taken over all $\boldsymbol{\vartheta} = (\theta_0, \boldsymbol{\theta})$ such that $\theta_0 \in \mathbb{R}$ and $\boldsymbol{\theta} \in \mathbb{R}^p$,

$$\varrho_\alpha(u) := |u| \left\{\alpha I[u > 0] + (1-\alpha) I[u \leq 0]\right\}, \; u \in \mathbb{R}, \quad \alpha \in [0,1],$$

and $\mathbf{x}_t := (Y_t,\ \ldots,\ Y_{t-p+1})'$, $t = 0, \ldots, n-1$.

This α-autoregression quantile can be obtained as the component $\hat{\boldsymbol{\vartheta}}^{(n)} = \left(\hat{\theta}_0^{(n)}, \hat{\boldsymbol{\theta}}^{(n)}\right) \in \mathbb{R}^{p+1}$ of the optimal solution $\left(\hat{\boldsymbol{\vartheta}}^{(n)}, \hat{\mathbf{r}}^+, \hat{\mathbf{r}}^-\right) \in \mathbb{R}^{2n+p+1}$ of the linear program

$$\begin{cases} \alpha \mathbf{1}_n' \mathbf{r}^+ + (1-\alpha)\mathbf{1}_n' \mathbf{r}^- & := \min \\ \mathbf{Y}_n - \mathbf{1}_n \theta_0 - \mathbf{X}_n \mathbf{z} = \mathbf{r}^+ - \mathbf{r}^- & \\ \mathbf{z} \in \mathbb{R}^p, \quad \mathbf{r}^\pm \in \mathbb{R}_+^n, \quad 0 \leq \alpha \leq 1 & \end{cases}, \tag{1.5}$$

where $\mathbf{1}_n$ stands for the n-dimensional vector $(1, ..., 1)'$, and $\mathbf{X}_n$ is the $n \times p$ matrix with rows $\mathbf{x}_t'$, $0 \leq t \leq n-1$.

The corresponding dual program is

$$\begin{cases} \mathbf{Y}_n'\mathbf{a} & := \max \\ \mathbf{1}_n'\mathbf{a} = n(1-\alpha) \\ \mathbf{X}_n'\mathbf{a} = (1-\alpha)\mathbf{X}_n'\mathbf{1}_n \\ \mathbf{a} \in [0,1]^n, \quad 0 \leq \alpha \leq 1 \end{cases} \tag{1.6}$$

Following the idea of Jurečková and Gutenbrunner (1992), later adapted to the autoregressive context by Koul and Saleh (1995), the solutions $\hat{\mathbf{a}}^{(n)}(\alpha) = \left(\hat{a}_1^{(n)}(\alpha), ..., \hat{a}_n^{(n)}(\alpha)\right)$, $0 \leq \alpha \leq 1$ of (1.6) are called the *α-autoregression rank scores.*

The formal duality between (1.5) and (1.6) implies that, for all $t = 1, ..., n$ and $\alpha \in [0,1]$,

$$\hat{a}_t^{(n)}(\alpha) = \begin{cases} 1 & \text{if} \quad Y_t > \hat{\theta}_0(\alpha) + \mathbf{x}_{t-1}'\hat{\boldsymbol{\theta}}^{(n)}(\alpha) \\ 0 & \text{if} \quad Y_t < \hat{\theta}_0(\alpha) + \mathbf{x}_{t-1}'\hat{\boldsymbol{\theta}}^{(n)}(\alpha) \end{cases},$$

while, for $t \in \left\{ t \,\middle|\, Y_t = \hat{\theta}_0(\alpha) + \boldsymbol{x}_{t-1}'\hat{\boldsymbol{\theta}}^{(n)}(\alpha) \right\}$, the components of $\hat{\mathbf{a}}^{(n)}(\alpha)$ are determined by the equality constraints in (1.6). Clearly, the sample paths $\{\hat{\mathbf{a}}^{(n)}(\alpha), \quad 0 \leq \alpha \leq 1\}$ are continuous, piecewise linear, and such that $\hat{a}_t^{(n)}(0) = 1$, and $\hat{a}_t^{(n)}(1) = 0$. An obvious modification of the algorithms of Koenker and d'Orey (1987 and 1994) allows for an efficient computation of the solutions $\hat{\boldsymbol{\vartheta}}^{(n)}(\alpha)$ and $\hat{\mathbf{a}}^{(n)}(\alpha)$ over the whole interval $[0,1]$.

A crucial property of autoregression rank scores is their *autoregression-invariance*, i.e., denoting by $\hat{a}_t^{(n)}(\alpha, \mathbf{Y}_n)$ the solution of (1.6), the fact that

$$\hat{a}_t^{(n)}(\alpha, \mathbf{Y}_n + z\mathbf{1}_n + \mathbf{X}_n\mathbf{z}) = \hat{a}_t^{(n)}(\alpha, \mathbf{Y}_n), \quad (z, \mathbf{z}) \in \mathbb{R}^{p+1}, \tag{1.7}$$

which immediately follows from (1.6). Some further algebraic relations between autoregression quantiles and the corresponding autoregression rank scores are provided in Lemma 2.1 of Hallin and Jurečková (1999). Quite remarkably, no preliminary estimation of $\boldsymbol{\theta}$ is needed in order to compute autoregression rank score statistics. This is in sharp contrast with the more familiar *aligned rank* methods (Hallin and Puri, 1994), where ranks are computed from estimated residuals; see Jurečková (1991), Gutenbrunner and

Jurečková (1992), Gutenbrunner et al. (1993), Hallin et al. (1997a, 1977b), Harel and Puri (1998), or Hallin and Jurečková (1999) for details and numerical applications.

In the present paper, new tests based on a autoregression rank score version of the traditional Kolmogorov-Smirnov statistic are introduced for model (1.1). The asymptotic behaviour of these tests is investigated in Section 3, where we show that the limiting distributions of the test statistics coincide with those of the classical Kolmogorov-Smirnov statistics, both under the null hypothesis as under contiguous alternatives. Our results extend those of Jurečková (1991) from regression models to autoregression models. The local asymptotic efficiency of these tests is also investigated. Finally, the performance of the proposed tests is illustrated on simulated AR series with Normal, Laplace and Cauchy innovation densities, respectively.

2 Limiting Distributions of Kolmogorov-Smirnov Statistics Based on Autoregression Rank Scores

Assume that the density f of the innovations in the autoregressive model (1.1) remains unspecified within the family $\mathcal{F}$ of exponentially tailed densities satisfying (1.2) and the following conditions (borrowed from Hallin and Jurečková, 1999) :

(F1) $f(x)$ is positive for all $x \in \mathbb{R}$, and absolutely continuous, with a.e. derivative f' and finite Fisher information $\mathcal{I}(f) := \int \left(\frac{f'(x)}{f(x)}\right)^2 f(x)\mathrm{d}x$ $< \infty$; moreover, there exists $K_f \geq 0$ such that, for all $|x| > K_f$, f has two bounded derivatives, f' and f'', respectively;

(F2) f is monotonically decreasing to 0 as $x \longrightarrow \pm\infty$ and, for some $b = b_f > 0$, $r = r_f \geq 1$,

$$\lim_{x\to-\infty} \frac{-\log F(x)}{b|x|^r} = \lim_{x\to\infty} \frac{-\log(1-F(x))}{b|x|^r} = 1.$$

We will focus on the problem of testing null hypotheses of the form

$$\mathcal{H}_0^n : \theta_p = 0, \quad \boldsymbol{\theta}_{(1)} := (\theta_1, ..., \theta_{p-1}) \text{ unspecified},$$

against alternatives

$$\mathcal{H}_1^n : \theta_p \neq 0, \quad \boldsymbol{\theta}_{(1)} := (\theta_1, ..., \theta_{p-1}) \text{ unspecified}.$$

Such tests play a crucial role, for instance, in the order identification process (see Garel and Hallin 1999).

Write the $AR(p)$ model (1.1) as

$$\mathbf{Y}_n = \mathbf{X}_n\boldsymbol{\theta} + \boldsymbol{\varepsilon}_n = \mathbf{X}_{n;1}\boldsymbol{\theta}_{(1)} + \mathbf{X}_{n;2}\theta_p + \boldsymbol{\varepsilon}_n, \tag{2.1}$$

where $\mathbf{X}_n := \left(\mathbf{X}_{n;1}\vdots\mathbf{X}_{n;2}\right)$ is the $(n \times p)$ matrix with rows $\mathbf{x}'_{t-1}$, $1 \le t \le n$, $\mathbf{X}_{n;2} := (Y_{-p+1}, \ldots, Y_{n-p})'$ (hence, $X_{n;2,t} = Y_{t-p}$, $t = 1, \ldots, n$), and $\boldsymbol{\varepsilon}_n := (\varepsilon_1, \ldots, \varepsilon_n)'$. Denote by

$$\mathbf{P}_n := \mathbf{X}_{n;1}(\mathbf{X}'_{n;1}\mathbf{X}_{n;1})^{-1}\mathbf{X}'_{n;1}$$

the (random) matrix projecting $\mathbb{R}^n$ onto the linear space spanned by the columns of $\mathbf{X}_{n;1}$. Define

$$\hat{\mathbf{X}}_{n;2} := \mathbf{P}\mathbf{X}_{n;2} \quad \text{and} \quad \mathbf{X}^{\perp}_{n;2} := \mathbf{X}_{n;2} - \hat{\mathbf{X}}_{n;2} - \left(\bar{X}_{n;2} - \bar{X}^0_{n;2}\right)\mathbf{1}_n$$
$$= [\mathbf{I} - \mathbf{P}_n]\left(\mathbf{X}_{n;2} - \bar{X}_{n;2}\mathbf{1}_n\right),$$

where

$$\bar{X}_{n;2} := n^{-1}\sum_{t=1}^{n} X_{n;2,t} = n^{-1}\sum_{t=-p+1}^{n-p} Y_t =: \bar{Y} \quad \text{and} \quad \bar{X}^0_{n;2} := n^{-1}\sum_{t=1}^{n} \hat{X}_{n;2,t},$$

and let

$$D^2_n := n^{-1}\left(\mathbf{X}^{\perp}_{n;2}\right)'\mathbf{X}^{\perp}_{n;2} = n^{-1}\left(\mathbf{X}_{n;2} - \bar{X}_{n;2}\mathbf{1}_n\right)'[\mathbf{I}_n - \mathbf{P}_n]\left(\mathbf{X}_{n;2} - \bar{X}_{n;2}\mathbf{1}_n\right).$$

Denoting by $\gamma_k(\boldsymbol{\theta})$, $k = 1, \ldots$ the autocovariances of the stationary solution of (1.1), the consistency under $AR(p)$ dependence of empirical autocovariances implies that D^2_n converges in probability, as $n \to \infty$, to

$$D^2 := \gamma_0(\boldsymbol{\theta}) - (\gamma_{p-1}(\boldsymbol{\theta}), \ldots, \gamma_1(\boldsymbol{\theta}))\begin{pmatrix} \gamma_0(\boldsymbol{\theta}) & \gamma_1(\boldsymbol{\theta}) & \ldots & \gamma_{p-2}(\boldsymbol{\theta}) \\ \gamma_1(\boldsymbol{\theta}) & \gamma_0(\boldsymbol{\theta}) & \ldots & \\ \vdots & & & \vdots \\ \gamma_{p-2}(\boldsymbol{\theta}) & \ldots & & \gamma_0(\boldsymbol{\theta}) \end{pmatrix}^{-1}\begin{pmatrix} \gamma_{p-1}(\boldsymbol{\theta}) \\ \vdots \\ \gamma_1(\boldsymbol{\theta}) \end{pmatrix}$$

which, in view of classical Yule-Walker equations, under $\mathcal{H}^n_0$ reduces to

$$D^2 = \gamma_0(\boldsymbol{\theta}) - \sum_{i=1}^{p-1}\theta_i\gamma_i(\boldsymbol{\theta}) = \sigma^2_f,$$

a simple scale factor that does not depend on $\boldsymbol{\theta}$, nor on the shape of the innovation density f.

Let $\hat{\mathbf{a}}^{(n)}(\alpha) = \left(\hat{a}^{(n)}_1(\alpha), \ldots, \hat{a}^{(n)}_n(\alpha)\right)$, $0 \le \alpha \le 1$, be the autoregression rank scores computed under $\mathcal{H}^n_0$, i.e., corresponding to the submodel

$$\mathbf{Y}_n = \mathbf{X}_{n;1}\boldsymbol{\theta}_{(1)} + \boldsymbol{\varepsilon}_n \tag{2.2}$$

(though of course $\hat{\mathbf{a}}^{(n)}(\alpha)$, as a statistic, does not depend on $\boldsymbol{\theta}_{(1)}$). For each n, consider the process $\{T_n(\alpha) : 0 \leq \alpha \leq 1\}$ defined by

$$T_n(\alpha) := n^{-1/2} D_n^{-1} \sum_{t=1}^{n} X_{n;2,t} \hat{a}_t^{(n)}(\alpha), \quad 0 \leq \alpha \leq 1. \tag{2.3}$$

This process has trajectories in the space $C_{[0,1]}$ of continuous functions $\alpha \mapsto c(\alpha)$, $\alpha \in [0,1]$ (as usual, $C_{[0,1]}$ is equipped with the Borel σ-field $\mathcal{C}$ associated with the uniform metric $\|c_1 - c_2\| := \max_{0\leq\alpha\leq 1} |c_1(\alpha) - c_2(\alpha)|$). Following Section V.3.2 of Hájek and Šidák (1967) [for convenience, we consistently refer to the original edition], define on $(C_{[0,1]}, \mathcal{C})$ the continuous functionals

$$h^+(c(\cdot)) := \max_{0\leq\alpha\leq 1} c(\alpha) \qquad \text{and} \qquad h^{\pm}(c(\cdot)) := \max_{0\leq\alpha\leq 1} |c(\alpha)|.$$

The one- and two-sided autoregression rank score-based Kolmogorov-Smirnov statistics we are considering in the problem of testing $\mathcal{H}_0^n$ are

$$K_n^+ := h^+(T_n(\cdot)) = n^{-1/2} D_n^{-1} \max_{0\leq\alpha\leq 1} \sum_{t=1}^{n} X_{n;2,t} \hat{a}_t^{(n)}(\alpha) \tag{2.4}$$

and

$$K_n^{\pm} := h^{\pm}(T_n(\cdot)) = n^{-1/2} D_n^{-1} \max_{0\leq\alpha\leq 1} \left| \sum_{t=1}^{n} X_{n;2,t} \hat{a}_t^{(n)}(\alpha) \right|, \tag{2.5}$$

respectively. The following results provide the asymptotic distributions of K_n^+ and $K_n^{\pm}$ under the null hypothesis $\mathcal{H}_0^n$.

Theorem 2.1 *Assume that (F1)-(F2) are satisfied. Then, under $\mathcal{H}_0^n$,*

$$\lim_{n\to\infty} \mathrm{P}(K_n^+ \leq x) = \begin{cases} 0 & x \leq 0 \\ 1 - \exp(-2x^2) & x > 0 \end{cases} \tag{2.6}$$

and

$$\lim_{n\to\infty} \mathrm{P}(K_n^{\pm} \leq x) = \begin{cases} 0 & x \leq 0 \\ 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1} \exp(-2k^2x^2) & x > 0 \end{cases}. \tag{2.7}$$

The proof is based on the following lemma.

Lemma 2.1 *Define the scores*

$$a_t^*(\alpha) := \begin{cases} 0 & n\alpha < R_{n;t} - 1 \\ R_{n;t} - n\alpha & R_{n;t} - 1 \leq n\alpha \leq R_{n;t} \\ 0 & R_{n;t} < n\alpha \end{cases}, \tag{2.8}$$

where $R_{n;t}$ denotes the rank of ε_t among $\varepsilon_1, \ldots, \varepsilon_n$. Assuming that (F1)-(F2) are satisfied, let

$$\tilde{a}_t(\alpha) := I[\varepsilon_t > F^{-1}(\alpha)], \quad 1 \le t \le n, \quad 0 \le \alpha \le 1.$$

Then,

$$\sup_{0\le\alpha\le1} n^{-1/2} \left| \sum_{t=1}^{n} \left[(X_{n;2,t} - \bar{X}_{n;2})\, \hat{a}_t^{(n)}(\alpha) - X_{n;2,t}^{\perp} \tilde{a}_t(\alpha) \right] \right| \xrightarrow{\mathrm{P}} 0 \qquad (2.9)$$

and

$$\sup_{0\le\alpha\le1} n^{-1/2} \left| \sum_{t=1}^{n} \left[(X_{n;2,t} - \bar{X}_{n;2})\, \hat{a}_t^{(n)}(\alpha) - X_{n;2,t}^{\perp} a_t^*(\alpha) \right] \right| \xrightarrow{\mathrm{P}} 0, \qquad (2.10)$$

as $n \to \infty$. Moreover, the process $\{T_n(\alpha) : 0 \le \alpha \le 1\}$ converges in distribution in $(C_{[0,1]}, \mathcal{C})$ to the Brownian bridge $\{Z(\alpha) : 0 \le \alpha \le 1\}$.

Proof of Lemma 2.1. The convergence (2.9) is a direct application of Theorem 3.2 of Hallin and Jurečková (1999). The approximation results on the rank score process given in Hájek (1965) (see also Jurečková 1999) then yield (2.10). The asymptotic behaviour of $\{T_n(\alpha) : 0 \le \alpha \le 1\}$ follows from the fact that

$$\left\{ D_n^{-1} n^{-1/2} \sum_{t=1}^{n} X_{n;2,t}^{\perp} a_{n;t}^*(\alpha), \quad 0 \le \alpha \le 1 \right\}$$

converges weakly to the Brownian bridge $\{Z(\alpha)\}$: see Theorem 1 in Section V.3.5 of Hájek and Šidák (1967). □

Proof Theorem 2.1. From Theorem V.3.3.a of Hájek and Šidák (1967), we have, for positive x and y and all $0 \le \alpha \le 1$,

$$\mathrm{P}\left[Z(\alpha) \le x(1-\alpha) + \alpha y\right] = 1 - \exp(-2xy)$$

and

$$\mathrm{P}\left[-x(1-\alpha) - \alpha y \le Z(\alpha) \le x(1-\alpha) + \alpha y\right] = 1 - 2\sum_{k=1}^{\infty} (-1)^{k+1} \exp(-2k^2 xy).$$

The asymptotic distributions (2.6) and (2.7) directly follow from letting $x = y$. □

Thus, the one-sided test based on K_n^+ rejects $\mathcal{H}_0^n$ at (asymptotic) probability level α whenever K_n^+ is larger than the critical value $\left(-\frac{1}{2}\log\alpha\right)^{1/2}$,

while the two-sided based on $K_n^{\pm}$ rejects $\mathcal{H}_0^n$ at probability level α whenever $K_n^{\pm}$ exceeds the α-quantile $q(\alpha) := Q^{-1}(\alpha)$ of the distribution function

$$Q(x) := 2\sum_{k=1}^{\infty}(-1)^{k-1}\exp(-2k^2x^2).$$

This function $Q(x)$ has been tabulated by Smirnov (1948). Note that both critical values are entirely distribution-free, as they do not depend neither on f nor on the nuisance $\boldsymbol{\theta}_{(1)}$.

We now turn to a study of the power of the tests based on K_n^+ (the case of $K_n^{\pm}$ follows along similar lines, and is left to the reader). The following result provides asymptotic powers against local alternatives of the form

$$\mathcal{H}_n : \theta_p = \theta_p^n := n^{-1/2}\tau, \quad \tau \in \mathbb{R}, \quad \boldsymbol{\theta}_{(1)} = (\theta_1, ..., \theta_{p-1}) \in \mathbb{R}^{p-1} \quad \text{unspecified.}$$

Theorem 2.2 *Assume that (F1)-(F2) are satisfied. Let*

$$p_n(\boldsymbol{\theta}_{(1)};\tau) := \mathrm{P}_{\boldsymbol{\theta}^n}\left[K_n^+ \geq \left(-\frac{1}{2}\log\alpha\right)^{1/2}\right],$$

where $\mathrm{P}_{\boldsymbol{\theta}^n}$ *is computed under* $\boldsymbol{\theta}^n := (\boldsymbol{\theta}_{(1)}, \theta_p^n)$. *Then, for any* $0 < \alpha < 1$,

$$\begin{aligned}\lim_{n\to\infty} p_n(\boldsymbol{\theta}_{(1)};\tau) &= B(\alpha, f, \tau)\\ &:= \mathrm{P}\left[\max_{0\leq u\leq 1}\left\{Z(u) + f_1(F_1^{-1}(u))\mathcal{I}^{-1/2}(f)\tau\right\} \geq \left(-\frac{1}{2}\log\alpha\right)^{1/2}\right], \qquad (2.11)\end{aligned}$$

where f_1 *and* F_1 *stand for the standardized versions of* f *and* F, *respectively. Furthermore, for* $\tau \longrightarrow 0$ *(the notation* $\xi(\tau) \sim \zeta(\tau)$ *means that the ratio* $\xi(\tau)/\zeta(\tau)$ *tends to one as* $\tau \to 0$; Φ, *as usual, stands for the standard normal distribution function),*

$$B(\alpha, f, \tau) - \alpha \sim \left(2\sigma_f\mathcal{I}^{-1/2}(f)\alpha\left(-\frac{1}{2}\log\alpha\right)^{1/2}\int_0^1 \phi(u, f)\psi(\alpha, u)\,\mathrm{d}u\right)\tau, \qquad (2.12)$$

with $\phi(u, f) := -\dfrac{f'(F^{-1}(u))}{f(F^{-1}(u))}$ *and*

$$\psi(\alpha, u) := 2\Phi\left[\left(-\frac{1}{2}\log\alpha\right)^{1/2}(2u-1)(u(1-u))^{-1/2}\right] - 1.$$

Proof. This proof, as well as the proof of Theorem 3.1, heavily relies on Sections VI.4.5 and VII.2.3 of Hájek and Šidák (1967); the various constants appearing there here are to be taken as

$$d_t = n^{-1/2}\tau Y_{t-p}, \quad c_t = n^{-1/2}X_{n;2,t}^{\perp} \text{ (hence, } \bar{c} = 0), \qquad \rho_2 = 1,$$

and

$$b^2 = \lim_{n\to\infty} \mathcal{I}(f)\tau^2 n^{-1} \sum_{t=1}^{n} (Y_{t-p} - \bar{Y})^2 = \tau^2 \mathcal{I}(f)\sigma_f^2.$$

Theorem VI.3.2 in Hájek and Šidák (1967) entails that, under $\mathcal{H}_n$, the process

$$T_n(\alpha) - f(F^{-1}(\alpha))D_n\mathcal{I}^{-1/2}(f)\tau$$

converges in distribution to the Brownian bridge $\{Z(\alpha)\}$. Since $K_n^+ = \max_{0\le\alpha\le 1} T_n(\alpha)$ and $D_n^2 = \sigma_f^2 + o_{\mathrm{P}}(1)$, equation (2.6) in turn implies that $p_n(\boldsymbol{\theta}_{(1)};\tau)$ converges, as $n \to \infty$, to

$$\mathrm{P}\left[\max_{0\le u\le 1}\left(Z(u) + f(F^{-1}(u))\sigma_f\mathcal{I}^{-1/2}(f)\tau\right) \ge \left(-\frac{1}{2}\log\alpha\right)^{1/2}\right]$$

for any α; (2.11) follows from noting that $\sigma_f f(F^{-1}(u)) = f_1(F_1^{-1}(u))$. As for the linear approximation (2.12), it is an immediate consequence of Theorem 1 in Section VI.4.5 of Hájek and Šidák (1967). □

3 Local Asymptotic Efficiency of the Test Based on K_n^+

Pitman asymptotic relative efficiencies of the tests based on K_n^+ with respect to the corresponding Gaussian Lagrange multiplier procedures, or with respect to the autoregression rank score test proposed by Hallin and Jurčková (1999) cannot be computed as easily as in the usual case of asymptotically normal or chi-square test statistics, for which AREs are obtained as ratios of noncentrality parameters. Actually, the analytical form of asymptotic powers, for given $\boldsymbol{\theta}$ and τ does not allow for any simple result, and AREs typically will depend on α, f, $\boldsymbol{\theta}$ and τ. This problem already appears in the classical case of Kolmogorov-Smirnov tests for linear models with independent observations; see Hájek and Šidák (1967), Section VII.2.3.

However, *local* ARE results can be obtained from the linear approximation (2.12) of the asymptotic power $B(\alpha, f, \tau)$ as $\tau \to 0$. More precisely, denote by $e(\alpha, f, \boldsymbol{\theta}, \tau)$ the ARE of the test based on K_n^+ with respect, for instance, to the locally optimal Gaussian test : the following result, inspired from Section VII.2.3 of Hájek and Šidák (1967), shows that the limit

$$e(\alpha, f) := \lim_{\tau\to 0} e(\alpha, f, \boldsymbol{\theta}, \tau) \tag{3.1}$$

does not depend on $\boldsymbol{\theta}$ and τ anymore, and thus allows for local or approximate comparisons of asymptotic performance.

Theorem 3.1 *Under (F1)-(F2), the local asymptotic relative efficiency of K_n^+ with respect to the locally optimal Gaussian test is*

$$e(\alpha, f) = 4\pi\alpha^2(-\log\alpha)\mathcal{I}^{-1}(f)\exp(z_\alpha^2)\left[\int_0^1 \phi(u,f)\psi(\alpha,u)\,du\right]^2, \quad (3.2)$$

where z_α denotes the $(1-\alpha)$-quantile of the normal distribution. Moreover,

$$\lim_{\alpha\to 0} e(\alpha) = \mathcal{I}^{-1}(f)\left[\int_0^1 \phi(u,f)\mathrm{sgn}(2u-1)\,du\right]^2 \quad (3.3)$$

Proof. The proof readily follows from Theorem VII.2.3 of Hájek and Šidák (1967). □

For instance, under the Laplace (double-exponential) density $f(x) = \frac{1}{2}\exp(-|x|)$, we have

$$\phi(u,f) = \mathrm{sgn}(2u-1), \quad 0 < u < 1.$$

One can easily check that

$$\int_0^1 \phi(u,f)\mathrm{sgn}(2u-1)\,du = 1 - 2\Phi\left(-(-\log\alpha)^{1/2}\right)/\alpha,$$

so that (3.2) becomes

$$e(\alpha) = 4\pi(-\log\alpha)\exp(z_{1-\alpha}^2)\left[\alpha - 2\Phi\left(-(-\log\alpha)^{1/2}\right)\right]^2.$$

In general, however, the local asymptotic efficiency $e(\alpha, f)$ still involves the actual innovation density f. If we apply the Cauchy-Schwarz inequality to (3.2), we obtain the uniform (in f) upper bound

$$\bar{e}(\alpha) := \sup_{f\in\mathcal{F}} e(\alpha, f) = 4\pi\alpha^2(-\log\alpha)\exp(z_{1-\alpha}^2)\int_0^1 \psi(\alpha,u)^2\,du.$$

Using the fact that

$$\lim_{x\longrightarrow\infty} 2\pi x^2\exp(x^2)(1-\Phi(x))^2 = 1,$$

it follows that

$$\lim_{\alpha\longrightarrow 0} 4\pi\alpha^2(-\log\alpha)\exp(z_{1-\alpha}^2) = 1,$$

hence

$$\lim_{\alpha\longrightarrow 0} \bar{e}(\alpha) = 1.$$

One can easily check by computations that $\bar{e}(0.01) = 0.864$, $\bar{e}(0.05) = 0.852$, and $\bar{e}(0.1) = 0.844$.

4 Simulation Results

The null hypothesis $\theta = 0$ has been considered in the AR(3) model

$$Y_t = 0.5Y_{t-1} - 0.2Y_{t-2} + \theta Y_{t-3} + \varepsilon_t. \tag{4.1}$$

More precisely, $N = 2000$ replications of the AR(3) series of length $n = 100$ generated by (4.1) with initial values $Y_{-2} = Y_{-1} = Y_0 = 0$, $\theta = 0$, $\theta = 0.1$, $\theta = 0.2$, $\theta = 0.3$, and $\theta = 0.4$, and three innovation densities (standard normal, standard Laplace, and standard logistic) have been subjected to the two Kolmogorov-Smirnov tests described in this paper (based on K^+ and $K^{\pm}$), respectively), and to the Gaussian Lagrange multiplier test (based on a quadratic form $Q_{\mathcal{L}}$; see Garel and Hallin 1999). For each of these tests, relative rejection frequencies (at probability levels $\alpha = 5\%$ and $\alpha = 1\%$) are reported in Table 1.

Inspection of the $\theta = 0$ column of Table 1 reveals that the three tests considered all are rather conservative, thus biased. The bias seems less severe, though, for the test based on K^+ than for the $K^{\pm}$ and $Q_{\mathcal{L}}$ ones. Since the same one-sided test based on K^+ is also significantly more powerful, under all alternatives considered here, than the two-sided one based on $K^{\pm}$, the two-sided Kolmogorov-Smirnov procedure might well be non admissible.

Under $\theta = 0.1$, the test based on K^+ is slightly more powerful than the Gaussian Lagrange multiplier test, even under Gaussian innovations, despite the local asymptotic optimality of the latter. Under larger values of θ, this advantage of K^+ over its competitors is clear under the three densities considered, but particularly marked under Laplace densities—as expected.

Though more extensive simulations should be undertaken, Table 1 nevertheless indicates that Kolmogorov-Smirnov techniques, especially the one-sided, based on nonlinear test statistics, can be expected to beat locally asymptotically optimal tests based on linear statistics.

Acknowledgement. The authors thank the Editor, Professor Ishwar Basawa, and an anonymous referee for their careful reading of the manuscript and helpful comments.

References

[1] Garel, B., and M. Hallin (1999). Rank-based AR order identification. *Journal of the American Statistical Association* **94**, 1357-1371 .

[2] Gutenbrunner, C., and J. Jurečková (1992). Regression rank scores and regression quantiles. *Annals of Statistics* **20**, 305-330.

[3] 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.

[4] Hájek, J. (1965). Extension of the Kolmogorov-Smirnov test to the regression alternatives. In *Bernoulli (1713), Bayes (1763), Laplace (1813) : Proceedings of an International Research Seminar* (L. Le Cam and J. Neyman, eds), 45-60. Berlin, Springer-Verlag.

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

[6] Hájek, J., Z. Šidák, and P.K. Sen (1999). *Theory of Rank Tests* (2nd edition). Academic Press, New York.

[7] Hallin, M., J. Jurečková, J. Kalvová, J. Picek, and T. Zahaf (1997). Nonparametric tests in *AR* models, with applications to climatic data. *Environmetrics* **8**, 651-660.

[8] Hallin, M., J. Jurečková, J. Picek, and T. Zahaf (1998). Nonparametric tests of independence of two autoregressive time series based on autoregression rank scores. *Journal of Statistical Planning and Inference* **75**, 319-330.

[9] Hallin, M., and J. Jurečková (1999). Optimal tests for autoregression models based on autoregression rank scores. *Annals of Statistics* **27**, 1385-1414.

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

[11] Hallin, M., and B.J.M. Werker (1998). Optimal testing for semiparametric autoregressive models : from Gaussian Lagrange multipliers to regression rank scores and adaptive tests. In *Asymptotics, Nonparametrics, and Time Series* (S. Ghosh, Ed.), 295-358. M. Dekker, New York.

[12] Harel, M., and M.L. Puri (1998). Autoregression quantiles and related rank score processes for generalized random coefficient autoregressive processes. *Journal of Statistical Planning and Inference* **68**, 271-294.

[13] Jurečková, J. (1991). Tests of Kolmogorov-Smirnov type based on regression rank scores. In *Transactions of the 11th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes* (J.Á. Víšek, Ed.), 41-49. Academia, Prague.

[14] Jurečková, J. (1999). Regression rank scores against heavy-tailed alternatives. *Bernoulli* **5**, 1-18.

[15] Koenker, R., and G. Bassett (1978). Regression quantiles. *Econometrica* **46**, 33-50.

[16] Koenker, R., and V. d'Orey (1987). Algorithm AS 229 : Computing regression quantiles. *Journal of the Royal Statistical Society Series C* **36**, 383-393.

[17] Koenker, R., and V. d'Orey (1994). Remark AS R92. A remark on Algorithm AS 229 : Computing dual regression quantiles and regression rank scores. *Journal of the Royal Statistical Society Series C* **43**, 410-414.

[18] Koul, H.K., and A.K.Md.E. Saleh (1995). Autoregression quantiles and related rank scores processes. *Annals of Statistics* **23**, 670-689.

[19] Smirnov, N.V. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics* **19**, 279-281.

Marc HALLIN mhallin@ulb.ac.be
Faouzi EL BANTLI elbantl@ulb.ac.be
Institut de Statistique et de Recherche Opérationnelle
Campus de la Plaine
Université Libre de Bruxelles CP210
B 1050 Brussels
Belgium

		$\theta = 0$		$\theta = 0.1$		$\theta = 0.2$		$\theta = 0.3$		$\theta = 0.4$	
	f	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
K^+	Normal	0.035	0.006	0.374	0.155	0.578	0.238	0.708	0.428	0.908	0.753
	Laplace	0.038	0.009	0.525	0.226	0.854	0.452	0.873	0.663	0.982	0.915
	Logistic	0.043	0.008	0.377	0.176	0.675	0.287	0.769	0.514	0.934	0.797
$K^\pm$	Normal	0.041	0.003	0.248	0.131	0.403	0.153	0.604	0.324	0.845	0.634
	Laplace	0.036	0.007	0.462	0.174	0.661	0.325	0.785	0.554	0.955	0.843
	Logistic	0.050	0.005	0.263	0.138	0.477	0.185	0.654	0.382	0.883	0.714
$Q_\mathcal{L}$	Normal	0.035	0.003	0.335	0.144	0.483	0.224	0.723	0.366	0.854	0.747
	Laplace	0.031	0.003	0.324	0.112	0.442	0.211	0.733	0.521	0.731	0.742
	Logistic	0.046	0.006	0.221	0.102	0.402	0.147	0.683	0.383	0.868	0.783
standard errors		0.004	0.002	0.011	0.009	0.011	0.011	0.011	0.011	.010	0.011

Table 1. Rejection frequencies of the null hypothesis $\theta = 0$ for $n = 100$, under standard Normal, Laplace, and Logistic innovation densities f, respectively, and for various values of θ, $\alpha = 5\%$ and $\alpha = 1\%$. The number of replications is $N = 2000$; maximal standard errors are provided for each column.