ISSN 1066-5307, Mathematical Methods of Statistics, 2019, Vol. 28, No. 2, pp. 144–154.

On the Asymptotic Power of Tests of Fit under Local Alternatives in Autoregression

M. V. Boldin[1*]

[1] *Dept. Mech. Math., Moscow State Lomonosov Univ., Moscow, Russia*

Received February 1, 2019; in final form, May 22, 2019

Abstract—We consider a stationary $AR(p)$ model. The autoregression parameters are unknown as well as the distribution of innovations. Based on the residuals from the parameter estimates, an analog of empirical distribution function is defined and the tests of Kolmogorov's and ω^2 type are constructed for testing hypotheses on the distribution of innovations. We obtain the asymptotic power of these tests under local alternatives.

Keywords: autoregression, residuals, empirical distribution function, Kolmogorov's and omega-square tests, local alternatives.

AMS 2010 Subject Classification: primary 62G10; secondary 62M10, 62G30.

DOI: 10.3103/S1066530719020042

1. INTRODUCTION AND PROBLEM SETTING

In this paper we deal with asymptotic study of the local power of goodness-of-fit tests of Kolmogorov–Smirnov and Cramér–von Mises–Smirnov (omega-square) type (henceforth, briefly, Kolmogorov and omega-square type) as applied to autoregression models. These tests are based on the empirical distribution function (d.f.) of residuals. These functions and tests of fit based on them in linear and nonlinear models have been studied for a long time. In particular, the paper [3] dealt with a stationary $AR(p)$ model

$$u_t = \beta_1 u_{t-1} + \cdots + \beta_p u_{t-p} + \varepsilon_t, \quad t \in \mathbb{Z}. \tag{1}$$

Here $\{\varepsilon_t\}$ are independent identically distributed (i.i.d.) random variables with unknown d.f. $G(x)$ such that $\mathsf{E}\varepsilon_1 = 0$, $0 < \mathsf{E}\varepsilon_1^2 < \infty$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ is the vector of unknown parameters such that the roots of the corresponding characteristic equation lie in the unit circle.

Let observations $u_{1-p}, \ldots, u_n$ form a sample from a stationary solution to the equation (1), and let $\hat{\boldsymbol{\beta}}_{\boldsymbol{n}} = (\hat\beta_{1n}, \ldots, \hat\beta_{pn})^T$ be any $n^{1/2}$-consistent estimate of $\boldsymbol{\beta}$ based on these observations. For example, the least squares estimate (LSE) is suitable, since it is asymptotically normal under our assumptions, see, e.g., [1], Chapter 5. The quantities

$$\hat\varepsilon_t = u_t - \hat\beta_{1n} u_{t-1} - \cdots - \hat\beta_{pn} u_{t-p}, \quad t = 1, \ldots, n,$$

are called *residuals* and the function

$$\hat G_n(x) = n^{-1} \sum_{t=1}^{n} I(\hat\varepsilon_t \le x), \quad x \in \mathbb{R}^1,$$

is called the *residual empirical d.f.* Here and henceforth $I(\cdot)$ denotes the indicator of an event.

The function $\hat G_n(x)$ is a counterpart of the empirical d.f.

$$G_n(x) = n^{-1} \sum_{t=1}^{n} I(\varepsilon_t \le x)$$

of the unobservable *innovations* $\varepsilon_1, \ldots, \varepsilon_n$.

[*] E-mail: boldin_m@hotmail.com

It was shown in [3] that if $G(x)$ is twice differentiable with $g(x) = G'(x)$ and $\sup_x |g'(x)| < \infty$, then

$$\sup_x |n^{1/2}[\hat{G}_n(x) - G_n(x)]| \xrightarrow{P} 0, \quad n \to \infty. \tag{2}$$

This enables us to test the hypothesis

$$H_0\colon G(x) = G_0(x), \quad G_0(x) \text{ is completely specified,}$$

by means of Kolmogorov and omega-square type tests. Namely, let $G_0^{-1}(t)$, $t \in [0,1]$, be the inverse function to $G_0(x)$ and let

$$\hat{v}_n(t) = n^{1/2}[\hat{G}_n(G_0^{-1}(t)) - t]$$

be the residual empirical process. This is a counterpart of the empirical process

$$v_n(t) = n^{1/2}[G_n(G_0^{-1}(t)) - t].$$

Due to (2)

$$\sup_t |\hat{v}_n(t) - v_n(t)| \xrightarrow{P} 0, \quad n \to \infty, \tag{3}$$

under H_0. It follows from (3) and well-known properties of $v_n(t)$ (see, e.g., [2], Chapter 3) that the process $\hat{v}_n(t)$ under H_0 weakly converges in the Skorohod space $D[0,1]$ to the Brownian bridge $v(t)$:

$$\hat{v}_n(t) \xrightarrow{D[0,1]} v(t), \quad n \to \infty. \tag{4}$$

The Kolmogorov and omega-square statistics for testing H_0 based on $\hat{v}_n(t)$ are

$$\hat{D}_n := \sup_t |\hat{v}_n(t)|, \quad \hat{\omega}_n^2 := \int_0^1 [\hat{v}_n(t)]^2\, dt.$$

According to (4) under H_0

$$P(\hat{D}_n \le \lambda) \to P(\sup_t |v(t)| \le \lambda) = K(\lambda),$$
$$P(\hat{\omega}_n^2 \le \lambda) \to P\left(\int_0^1 [v(t)]^2 dt \le \lambda\right) = S(\lambda), \quad n \to \infty,$$

where $K(\lambda)$ and $S(\lambda)$ are well-known tabulated Kolmogorov's and Smirnov's distribution functions. Therefore the statistics $\hat{D}_n$ and $\hat{\omega}_n^2$ can be used for testing H_0 with large n in the same way as the usual statistics based on $v_n(t)$.

The attractive results stated above were established only under the hypothesis H_0, while the behavior of the residual empirical process $\hat{v}_n(t)$ under local alternatives has not been studied so far. This matter will be treated in this paper.

The aim of this paper is to establish the weak limits in $D[0,1]$ of $\hat{v}_n(t)$ and the statistics $\hat{D}_n$ and $\hat{\omega}_n^2$ based on it under local alternatives. We will carry over the results valid for $v_n(t)$ under local alternatives to the residual empirical process. A systematic treatment of weak convergence of $v_n(t)$ in various metric spaces was given in [5]. Let us state some results of that paper to be used here.

Since testing the hypothesis $H_0\colon G(x) = G_0(x)$ with continuous $G_0(x)$ is equivalent to testing $G_0(\varepsilon_1), \dots, G_0(\varepsilon_n)$ for uniformity on $[0,1]$, we will discuss testing the hypothesis $F_0(t) = t$, $t \in [0,1]$. We consider alternatives to $F_0(t)$ of the form

$$F_n(t) = F_0(t) + n^{-1/2}\delta_n(t), \tag{5}$$

where $\delta_n(t)$ converges to a function $\delta(t)$ as $n \to \infty$. The mode of convergence depends on the metric space in which the weak convergence of $v_n(t)$ is studied. There may be various metric spaces, but the general results of [5] imply that whenever $\delta_n(t)$ uniformly converges to a continuous function $\delta(t)$, we have

$$v_n(t) \xrightarrow{D[0,1]} v(t) + \delta(t), \quad n \to \infty,$$

under alternatives (5). Of course, for Kolmogorov's statistic D_n and omega-square statistic ω_n^2 based on $v_n(t)$ this implies that

$$D_n \xrightarrow{d} \sup_t |v(t) + \delta(t)|, \quad \omega_n^2 \xrightarrow{d} \int_0^1 [v(t) + \delta(t)]^2 \, dt, \quad n \to \infty.$$

Now we turn back to our problem. In Subsection 2.2 below we assume that $\{\varepsilon_t\}$ in (1) are i.i.d. random variables with d.f. in the form of a mixture

$$A_n(x) := (1 - n^{-1/2}) G_0(x) + n^{-1/2} H_n(x) \quad \text{with } H_n(x) \text{ being a d.f.} \tag{6}$$

Assumption (6) will be regarded as a local alternative to H_0 to be denoted by H_{1n}. The variables $G_0(\varepsilon_1), \ldots, G_0(\varepsilon_n)$ under this alternative have the following d.f. of the form (5):

$$F_n(t) = A_n(G_0^{-1}(t)) = t + n^{-1//2} \delta_n(t), \quad \delta_n(t) = H_n(G_0^{-1}(t)) - t.$$

The representation (6) of the alternative in the form of a mixture will be convenient for formulating conditions on $G_0(x)$ and $H_n(x)$ providing the properties of the estimates of $\boldsymbol{\beta}$ under H_{1n} needed for our results.

The results about the asymptotic behavior of $\hat{v}_n(t)$ and the statistics $\hat{D}_n$ and $\hat{\omega}_n^2$ under H_{1n} will be given in Theorem 2.2 and Corollary 2.2 (Subsection 2.2). To obtain them, we first prove an analog of relation (2) for the case when the distribution of innovations depends on n. This will be done under more general assumptions than (6) in Theorem 2.1 and Corollary 2.1 (Subsection 2.1).

By now relations of type (2) (i.e., uniform stochastic expansions of the residual empirical d.f.) have been established for various autoregression models: $ARMA$, explosive and unstable autoregression, $AR(\infty)$, $ARCH$, $GARCH$, and some others, see [4] and references therein. Therefore in these models one can also test hypotheses on the distribution of innovations by means of Kolmogorov and omega-square type tests. The present paper is a step towards the study of the power of these tests.

The main results are stated in Section 2, the proofs are collected in Section 3.

2. MAIN RESULTS

2.1. Stochastic Expansion for the Residual Empirical d.f.

In this subsection we do not deal with hypothesis testing, but focus on obtaining a stochastic expansion for the residual empirical d.f. when the distribution of innovations may depend on n.

Namely, we will assume that $\{\varepsilon_t\}$ in (1) are i.i.d. r.v.'s with d.f. $A_n(x)$. But we stress once more that this d.f. need not satisfy (6), we impose on it the following very general conditions. (In what follows, in order to emphasize dependence on n, we write E_n for the expectation with respect to a d.f. depending on n.)

Condition (i). $\mathsf{E}_n \varepsilon_1 = 0, \quad \sup_n \mathsf{E}_n \varepsilon_1^2 < \infty.$

Condition (ii). The d.f. $A_n(x)$ is differentiable with derivative $a_n(x)$ satisfying the Lipschitz condition:

$$|a_n(x_1) - a_n(x_2)| < L|x_1 - x_2| \quad \text{for all } x_1, x_2 \in \mathbb{R}^1,$$

where the constant $L > 0$ does not depend on n.

Similarly to the residuals $\{\hat{\varepsilon}_t\}$ from the estimate $\hat{\boldsymbol{\beta}}_n$ defined in the Introduction, let us define the residuals from a nonrandom vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^T \in \mathbb{R}^p$ by

$$\varepsilon_t(\boldsymbol{\theta}) := u_t - \theta_1 u_{t-1} - \ldots - \theta_p u_{t-p}, \quad t = 1, \ldots, n.$$

Let us define the corresponding residual empirical d.f.

$$G_n(x, \boldsymbol{\theta}) = n^{-1} \sum_{t=1}^n I(\varepsilon_t(\boldsymbol{\theta}) \le x), \quad x \in \mathbb{R}^1.$$

When $\boldsymbol{\theta} = \boldsymbol{\beta}$, the function $G_n(x, \boldsymbol{\beta})$ coincides with the empirical d.f. $G_n(x)$ of $\varepsilon_1, \ldots, \varepsilon_n$.

In what follows $|\cdot|$ denotes the Euclidean norm of a vector.

Theorem 2.1. *Let $\{\varepsilon_t\}$ be i.i.d. r.v.'s with d.f. $A_n(x)$ satisfying Conditions (i) and (ii). Then, for any $0 \le \Theta < \infty$ and $\delta > 0$,*

$$\mathsf{P}(\sup_{x,|\boldsymbol{\tau}|\le\Theta} |n^{1/2}[G_n(x, \boldsymbol{\beta} + n^{-1/2}\boldsymbol{\tau}) - G_n(x)]| > \delta) \to 0, \quad n \to \infty.$$

Let $\hat{\boldsymbol{\beta}}_n$ be an estimate for $\boldsymbol{\beta}$. Put

$$\hat{G}_n(x) := G_n(x, \hat{\boldsymbol{\beta}}_n).$$

Theorem 2.1 implies the following Corollary.

Corollary 2.1. *Let the conditions of Theorem 2.1 hold. Let $\hat{\boldsymbol{\beta}}_n$ be an $n^{1/2}$-consistent estimate of $\boldsymbol{\beta}$. Then for any $\delta > 0$*

$$\mathsf{P}(\sup_x |n^{1/2}[\hat{G}_n(x) - G_n(x)]| > \delta) \to 0, \quad n \to \infty.$$

The proof of this Corollary is carried out in a standard manner (cf., e.g., the proof of Corollary 2.1 in [4]) and hence is omitted.

2.2. Residual Empirical Process and Test Statistics under Local Alternatives

Theorem 2.1 is valid for any sequence of d.f.'s $A_n(x)$ satisfying Conditions (i), (ii), and Corollary 2.1 holds under an additional assumption about $n^{1/2}$-consistency of the estimate $\hat{\boldsymbol{\beta}}_n$. Note that $A_n(x)$ need not converge to $G_0(x)$ as $n \to \infty$. Moreover, Conditions (i), (ii) do not ensure the existence of a $n^{1/2}$-consistent estimate of $\hat{\boldsymbol{\beta}}_n$.

We now turn to the sequence of local alternatives H_{1n} as in (6) that converge to $G_0(x)$. It will be convenient to state the requirements on the specific d.f. $A_n(x)$ defined by (6) directly in terms of $G_0(x)$ and $H_n(x)$. We will impose on them Conditions (iii), (iv), which will imply Conditions (i), (ii) and moreover ensure that the LSE is $n^{1/2}$-consistent. Hence this will enable us to use the results of Subsection 2.1 under the alternatives H_{1n}.

Of course, Conditions (iii), (iv) are one of possible versions of requiremehts on $A_n(x)$ as in (6) under which the results stated below hold.

Condition (iii). The d.f.'s $G_0(x)$ and $H_n(x)$ have zero means and variances σ_0^2 and σ_{nH}^2 such that

$$0 < \sigma_0^2 < \infty, \quad \sigma_{nH}^2 = o(n^{1/2}), \quad n \to \infty.$$

The intuitive meaning of Condition (iii) is that the tails of $H_n(x)$ may become heavier and the variance σ_{nH}^2 may grow with growing n. However the variance of $A_n(x)$ as in (6) tends to σ_0^2 as $n \to \infty$, i.e. Condition (iii) implies Condition (i). Moreover Condition (iii) enables us to construct an $n^{1/2}$-consistent estimate for the unknown parameter $\boldsymbol{\beta}$, specifically the LSE, which will be discussed later.

Condition (iv). The d.f.'s $G_0(x)$ and $H_n(x)$ are differentiable with derivatives satisfying the Lipschitz condition; the Lipschitz constant for $H_n'(x)$ is L_{nH} and

$$L_{nH} = O(n^{1/2}), \quad n \to \infty.$$

Intuitively Condition (iv) means that the density $H_n'(x)$ may oscillate with growing frequency of oscillation when n grows. Nevertheless the d.f.'s $A_n(x)$ as in (6) will satisfy the Lipschitz condition with a constant independent of n, i.e. Condition (iv) implies Condition (ii).

Let $\hat{\boldsymbol{\beta}}_n$ be any estimate of $\boldsymbol{\beta}$, which is $n^{1/2}$-consistent under Conditions (iii) and (iv). For example, the least squares estimate (LSE) $\hat{\boldsymbol{\beta}}_{n,LS}$ is well suited, which under the sole Condition (iii) is asymptotically normal:

$$n^{1/2}(\hat{\boldsymbol{\beta}}_{n,LS} - \boldsymbol{\beta}) \xrightarrow{d} \mathbb{N}(\mathbf{0}, \sigma_0^2\mathbb{K}^{-1}), \quad n \to \infty. \tag{7}$$

Here $\mathbb{K}$ is the $p \times p$ matrix,

$$\mathbb{K} = (k_{ij}) > 0, \quad k_{ij} = \mathsf{E} u_0^0 u_{i-j}^0, \quad i, j = 1, \dots, p,$$

where $\{u_t^0\}$ is a stationary solution of (1) with innovations having the d.f. $G_0(x)$.

Relation (7) (to be proved in Section 3) means that, subject to Condition (iii), the LSE remains asymptotically normal under H_{1n}, $n \to \infty$, with the same parameters as under H_0.

Now, by Corollary 2.1, we obtain that under H_{1n}, subject to Conditions (iii) and (iv),

$$\sup_t |\hat{v}_n(t) - v_n(t)| \xrightarrow{P} 0, \quad n \to \infty. \tag{8}$$

Relation (8) and the well-known properties of the process $v_n(t)$ under local alternatives, see [5], Theorem 4.1, imply the following theorem.

Theorem 2.2. *Let the alternative H_{1n} as in (6) hold and Conditions (iii), (iv) be satisfied. Let the functions*

$$\delta_n(t) := H_n(G_0^{-1}(t)) - t, \quad t \in [0,1], \tag{9}$$

uniformly converge to a continuous function $\delta(t)$ as $n \to \infty$. Then

$$\hat{v}_n(t) = n^{1/2}[\hat{G}_n(G_0^{-1}(t)) - t] \xrightarrow{D[0,1]} v(t) + \delta(t), \quad n \to \infty,$$

where $v(t)$ is a Brownian bridge.

Theorem 2.2 immediately implies:

Corollary 2.2. *Under the conditions of Theorem 2.2 the following convergence in distribution holds:*

$$\hat{D}_n \xrightarrow{d} \sup_t |v(t) + \delta(t)|, \quad \hat{\omega}_n^2 \xrightarrow{d} \int_0^1 [v(t) + \delta(t)]^2 dt, \quad n \to \infty. \tag{10}$$

Remark 2.1. Let $H[0,1]$ be the Hilbert space of functions on $[0,1]$ with the norm $|x(t)|_H^2 = \int_0^1 x^2(t)\,dt$. Let the functions $\delta_n(t)$ as in (9) converge to $\delta(t)$ in $H[0,1]$ as $n \to \infty$. By Theorem 5.1 of [5], under H_{1n},

$$v_n(t) = n^{1/2}[G_n(G_0^{-1}(t)) - t] \xrightarrow{H[0,1]} v(t) + \delta(t), \quad n \to \infty.$$

Therefore, if (8) holds (i.e., under Conditions (iii) and (iv)), then the residual empirical process also converges:

$$\hat{v}_n(t) = n^{1/2}[\hat{G}_n(G_0^{-1}(t)) - t] \xrightarrow{H[0,1]} v(t) + \delta(t), \quad n \to \infty.$$

This relation implies (10) for $\hat{\omega}_n^2$. The condition $|\delta_n(t) - \delta(t)|_H \to 0$, $n \to \infty$, is weaker than the assumption about uniform convergence of $\delta_n(t)$ in Theorem 2.2.

3. PROOF OF THEOREM 2.1.

We will present the proof for $p = 1$. The proof for an arbitrary p is more cumbersome, but does not offer principal difficulties. So, setting $\beta_1 = \beta$, we will consider the $AR(1)$ equation

$$u_t = \beta u_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}. \tag{11}$$

In (11), $\{\varepsilon_t\}$ are i.i.d. r.v.'s with d.f. $A_n(x)$ satisfying Conditions (i) and (ii); $|\beta| < 1$. Then the stationary solution of (11) has the form

$$u_t = \sum_{j \geq 0} \beta^j \varepsilon_{t-j},$$

where the series converges in L^2,

$$\sup_n \mathsf{E}_n u_t^2 = \sup_n \mathsf{E}_n \varepsilon_1^2 \sum_{j\ge 0} \beta^{2j} < \infty.$$

Henceforth we will write E instead of E_n.

By (11),

$$\varepsilon_t(\beta + n^{-1/2}\tau) = \varepsilon_t - n^{-1/2}\tau u_{t-1}, \quad t = 1,\dots,n,$$

$$G_n(x, \beta + n^{-1/2}\tau) = n^{-1}\sum_{t=1}^{n} I(\varepsilon \le x + n^{-1/2}\tau u_{t-1}).$$

We will need the process

$$u_n(x,\tau) := n^{-1/2}\sum_{t=1}^{n}[I(\varepsilon_t \le x + n^{-1/2}\tau u_{t-1}) - A_n(x + n^{-1/2}\tau u_{t-1})].$$

Consider the sigma-algebra $\mathcal{F}_t = \sigma\{\varepsilon_s, s \le t\}$, then the summands in the definition of $u_n(x,\tau)$ form a martingale-difference with respect to $\{\mathcal{F}_t\}$. Since

$$\begin{aligned} &n^{1/2}[G_n(x, \beta + n^{-1/2}\tau) - G_n(x,\beta)] \\ &\quad = u_n(x,\tau) - u_n(x,0) + n^{-1/2}\sum_{t=1}^{n}[A_n(x + n^{-1/2}\tau u_{t-1}) - A_n(x)], \end{aligned}$$

for the proof of the theorem it suffices to prove the following two statements:

$$\sup_{x\in\mathbb{R}^1, |\tau|\le\Theta} |u_n(x,\tau) - u_n(x,0)| = o_P(1), \tag{12}$$

$$\sup_{x\in\mathbb{R}^1, |\tau|\le\Theta} \left| n^{-1/2}\sum_{t=1}^{n}[A_n(x + n^{-1/2}\tau u_{t-1}) - A_n(x)]\right| = o_P(1), \quad n\to\infty, \tag{13}$$

where $0 < \Theta < \infty$ is fixed.

Let us prove (12). First of all we need a discrete approximation for

$$\sup_{x\in\mathbb{R}^1, |\tau|\le\Theta} |u_n(x,\tau) - u_n(x,0)|.$$

Split the interval $[-\Theta n^{-1/2}, \Theta n^{-1/2}]$ into 3^{m_n} subintervals (m_n are positive integers such that $3^{m_n} \sim \log n$ as $n\to\infty$) by the points

$$\eta_s = -\Theta n^{-1/2} + 2\Theta n^{-1/2} 3^{-m_n} s, \quad s = 0,1,\dots,3^{m_n}.$$

Let

$$\hat{u}_{ts} = u_t[1 - 2\Theta n^{-1/2}3^{-m_n}\eta_s^{-1} I(u_t \le 0)],$$
$$\tilde{u}_{ts} = u_t[1 - 2\Theta n^{-1/2}3^{-m_n}\eta_s^{-1} I(u_t \ge 0)].$$

From among the points $\{\eta_s\}$ select the point η_j, which is nearest on the right to $n^{-1/2}\tau$, then

$$0 \le \eta_j - n^{-1/2}\tau \le 2\Theta n^{-1/2} 3^{-m_n}.$$

These definitions immediately imply

$$\begin{aligned} &\eta_j \tilde{u}_{t-1,j} \le n^{-1/2}\tau u_{t-1} \le \eta_j \hat{u}_{t-1,j}, \\ &|\hat{u}_{ts}| \le 3|u_t|, \ |\tilde{u}_{ts}| \le 3|u_t|, \quad t = 1,\dots,n. \end{aligned} \tag{14}$$

Take the points

$$-\infty = x_0 < x_1 < \ldots < x_{N_n} = +\infty$$

so that

$$A_n(x_i) = i/N_n, \quad N_n \sim n^{1/2}\log n \quad \text{as } n \to \infty.$$

If $x \in [x_i, x_{i+1})$, then

$$x_i + \eta_j \tilde{u}_{t-1,j} \le x + n^{-1/2}\tau u_{t-1} \le x_{i+1} + \eta_j \hat{u}_{t-1,j}. \tag{15}$$

Define the vectors

$$\hat{U}_j = (\hat{u}_{0,j}, \dots, \hat{u}_{n-1,j}), \quad \tilde{U}_j = (\tilde{u}_{0,j}, \dots, \tilde{u}_{n-1,j}),$$

and let

$$p_n(x, \hat{U}_j) := n^{-1/2}\sum_{t=1}^{n}[I(\varepsilon_t \le x + \eta_j \hat{u}_{t-1,j}) - A_n(x + \eta_j \hat{u}_{t-1,j}) - I(\varepsilon_t \le x) + A_n(x)].$$

Monotonicity of $I(\varepsilon_t \le y)$ and $A_n(y)$ and inequalities (15) imply the following two inequalities:

$$\begin{aligned}
u_n(x,\tau) - u_n(x,0) \le p_n(x_{i+1}, \hat{U}_j) &\\
&+ n^{-1/2}\sum_{t=1}^{n}[I(\varepsilon_t \le x_{i+1}) - A_n(x_{i+1}) - I(\varepsilon_t \le x_i) + A_n(x_i)]\\
&+ n^{-1/2}\sum_{t=1}^{n}[A_n(x_{i+1}) - A_n(x_i)]\\
&+ n^{-1/2}\sum_{t=1}^{n}[A_n(x_{i+1} + \eta_j \hat{u}_{t-1,j}) - A_n(x_i + \eta_j \tilde{u}_{t-1,j})]
\end{aligned}$$

and, similarly,

$$\begin{aligned}
u_n(x,\tau) - u_n(x,0) \ge p_n(x_i, \tilde{U}_j) &\\
&- n^{-1/2}\sum_{t=1}^{n}[I(\varepsilon_t \le x_{i+1}) - A_n(x_{i+1}) - I(\varepsilon_t \le x_i) + A_n(x_i)]\\
&- n^{-1/2}\sum_{t=1}^{n}[A_n(x_{i+1}) - A_n(x_i)]\\
&- n^{-1/2}\sum_{t=1}^{n}[A_n(x_{i+1} + \eta_j \hat{u}_{t-1,j}) - A_n(x_i + \eta_j \tilde{u}_{t-1,j})].
\end{aligned}$$

The last two inequalities imply

$$\sup_{x\in\mathbb{R}^1, |\tau|\le\Theta} |u_n(x,\tau) - u_n(x,0)|$$

$$\le \max_{i,j}\{|p_n(x_{i+1}, \hat{U}_j)| + |p_n(x_i, \tilde{U}_j)|\} \tag{16}$$

$$+ \max_{i} |n^{-1/2}\sum_{t=1}^{n}[I(\varepsilon_t \le x_{i+1}) - A_n(x_{i+1}) - I(\varepsilon_t \le x_i) + A_n(x_i)]| \tag{17}$$

$$+ \max_{i} |n^{-1/2}\sum_{t=1}^{n}[A_n(x_{i+1}) - A_n(x_i)]| \tag{18}$$

$$+ \max_{i,j} n^{-1/2}\sum_{t=1}^{n}[A_n(x_{i+1} + \eta_j \hat{u}_{t-1,j}) - A_n(x_i + \eta_j \tilde{u}_{t-1,j})]. \tag{19}$$

The discrete approximation is completed.

We will show that expressions (16) to (19) tend to zero in probability as $n \to \infty$.

Lemma 3.1. *Under Conditions* (i) *and* (ii) *the expression* (16) *is* $o_p(1)$ *as* $n \to \infty$.

Proof. In the proof of this lemma and subsequent statements we will use boundedness of the density $a_n(x)$. Namely, by Taylor's formula and Condition (i) we have for an intermediate point $\tilde{x} \in (x, x+1)$ that

$$A_n(x+1) = A_n(x) + a_n(\tilde{x}) = A_n(x) + a_n(x) + a_n(\tilde{x}) - a_n(x),$$
$$|a_n(\tilde{x}) - a_n(x)| < L,$$

whence

$$\sup_{x,n} a_n(x) < L + 1.$$

Now we turn to the proof of Lemma 3.1 per se. We will show that

$$\max_{i,j}\{|p_n(x_{i+1}, \hat{U}_j)|\} = o_p(1), \quad n \to \infty,$$

the reasoning for the second term in (16) is quite similar. Put

$$\begin{aligned} V_t(i,j) &:= I(\varepsilon_t \le x_{i+1} + \eta_j \hat{u}_{t-1,j}) \\ &\qquad - A_n(x_{i+1} + \eta_j \hat{u}_{t-1,j}) - I(\varepsilon_t \le x_{i+1}) + A_n(x_{i+1}), \\ S_n(i,j) &:= \sum_{t=1}^{n} V_t(i,j). \end{aligned}$$

Then

$$p_n(x_{i+1}, \eta_j, \hat{U}_j) = n^{-1/2} \sum_{t=1}^{n} V_t(i,j) = n^{-1/2} S_n(i,j).$$

Let $\mathcal{F}_t = \sigma\{\varepsilon_s,\ s \le t\}$ be the sigma-algebra introduced before. Obviously, the sequence $\{V_t(i,j), \mathcal{F}_t\}$ is a martingale-difference. Therefore the sequence $\{S_n(i,j), \mathcal{F}_n\}$, $n \ge 1$, is a martingale. By Rosenthal's inequality (see [6], p. 23)

$$\mathsf{E} S_n^4(i,j) \le c\Big\{\mathsf{E}\Big[\sum_{t=1}^{n} \mathsf{E}(V_t^2(i,j) \mid \mathcal{F}_{t-1})\Big]^2 + \sum_{t=1}^{n} \mathsf{E} V_t^4(i,j)\Big\}.$$

The constants c here and $c_1, c_2, \dots$ henceforth do not depend on n, t, i, j. Obviously,

$$\mathsf{E} V_t^4(i,j) \le c_1, \quad \sum_{t=1}^{n} \mathsf{E} V_t^4(i,j) \le c_1 n.$$

Next we use the well-known inequality: for $x_1, x_2 \in \mathbb{R}^1$,

$$\mathsf{E}|I(\varepsilon_1 \le x_1) - A_n(x_1) - I(\varepsilon_1 \le x_2) + A_n(x_2)|^2 \le |A_n(x_1) - A_n(x_2)|.$$

In view of this inequality and (14), we have almost sure

$$\begin{aligned} \mathsf{E}(V_t^2(i,j) \mid \mathcal{F}_{t-1}) &\le |A_n(x_{i+1} + \eta_j \hat{u}_{t-1,j}) - A_n(x_{i+1})| \\ &\le \sup_{x,n} a_n(x) \mid \eta_j \hat{u}_{t-1,j}| \le c_2 n^{-1/2} |u_{t-1}|. \end{aligned}$$

Hence by the Cauchy–Bunyakovskii inequality

$$\begin{aligned} \mathsf{E}\Big[\sum_{t=1}^{n} \mathsf{E}(V_t^2(i,j)) \mid \mathcal{F}_{t-1}\Big]^2 &\le c_3 \mathsf{E}\Big(\sum_{t=1}^{n} n^{-1/2} |u_{t-1}|\Big)^2 \\ &\le c_3 \sum_{t=1}^{n} \mathsf{E} u_{t-1}^2 \le c_4 n. \end{aligned}$$

Therefore

$$\mathsf{E}S_n^4(i,j) \le c_5 n.$$

Using this and Chebyshev's inequalities we obtain

$$\begin{aligned}\mathsf{P}(\max_{i,j}\{|p_n(x_{i+1},\eta_j,\hat{U}_j)|\} > \delta) &\le \sum_{i,j}\mathsf{P}(|p_n(x_{i+1},\eta_j,\hat{U}_j)| > \delta)\\ &\le \sum_{i,j}\delta^{-4}n^{-2}\mathsf{E}S_n^4(i,j) \le c_5\delta^{-4}(N_n+1)(3^{m_n}+1)n^{-1}\\ &= O(n^{-1/2}\log^2 n) = o(1), \quad n\to\infty,\end{aligned}$$

due to the choice of the sequences $\{N_n\}, \{m_n\}$. The proof of Lemma 3.1 is completed. □

Lemma 3.2. *Under Condition* (ii) *the expression* (17) *is* $o_p(1)$ *as* $n \to \infty$.

Proof. Let

$$\nu_t(i) := I(\varepsilon_t \le x_{i+1}) - A_n(x_{i+1}) - I(\varepsilon_t \le x_i) + A_n(x_i),$$

$$q_n(i) := \sum_{t=1}^n \nu_t(i).$$

Obviously, $\{\nu_t(i)\}$ are a sequence of i.i.d. r.v.'s with $\mathsf{E}\nu_1(i) = 0$, $|\nu_t(i)| \le 2$, and

$$\mathsf{E}\nu_1^2(i) \le |A_n(x_{i+1}) - A_n(x_i)| \le N_n^{-1} \le cn^{-1/2}.$$

Then

$$\mathsf{E}q_n^4(i) = n\mathsf{E}\nu_1^4(i) + \sum_{t\neq s}\mathsf{E}\nu_t^2(i)\mathsf{E}\nu_s^2(i) \le c_1 n.$$

According to this and Chebyshev's inequalities,

$$\begin{aligned}\mathsf{P}(\max_i\{|n^{-1/2}q_n(i)|\} > \delta) &\le \sum_i \mathsf{P}(|n^{-1/2}q_n(i)| > \delta)\\ &\le \sum_i \delta^{-4}n^{-2}\mathsf{E}q_n^4(i) \le c_1\delta^{-4}(N_n+1)n^{-1} = o(1), \quad n\to\infty,\end{aligned}$$

due to the choice of the sequence $\{N_n\}$. This proves Lemma 3.2. □

The expression (18) is $o(1)$ as $n \to \infty$ due to the choice of the sequence $\{N_n\}$.

Lemma 3.3. *Under Conditions* (i) *and* (ii) *expression* (19) *is* $o_p(1)$ *as* $n \to \infty$.

Proof. By means of Taylor's formula we rewrite the expression under the maximum sign in (19) as follows:

$$n^{-1/2}\sum_{t=1}^n[A_n(x_{i+1}) - A_n(x_i)] \tag{20}$$

$$+\, n^{-1/2}\eta_j\sum_{t=1}^n a_n(\hat{x}_t)\hat{u}_{t-1,j} - n^{-1/2}\eta_j\sum_{t=1}^n a_n(\tilde{x}_t)\tilde{u}_{t-1,j}, \tag{21}$$

where $\hat{x}_t, \tilde{x}_t$ are intermediate points.

The expression (20) is $n^{1/2}N_n^{-1} = o(1)$ as $n \to \infty$.

Rewrite the first sum in (21) as follows:

$$n^{-1/2}\eta_j\sum_{t=1}^n a_n(\hat{x}_t)\hat{u}_{t-1,j}$$

$$= n^{-1/2}\eta_j \sum_{t=1}^{n} a_n(x_{i+1})\hat{u}_{t-1,j} + n^{-1/2}\eta_j \sum_{t=1}^{n}[a_n(\hat{x}_t) - a_n(x_{i+1})]\hat{u}_{t-1,j}. \qquad (22)$$

In view of Condition (i) and inequality (14) the second sum in (22) is no greater than

$$9L\Theta^2 n^{-3/2}\sum_{t=1}^{n} u_{t-1}^2 = o_P(1), \quad n \to \infty$$

in absolute value. The absolute value of the first sum in (22) is no greater than

$$\Theta \sup_{x,n} a_n(x)\Big|n^{-1}\sum_{t=1}^{n} u_{t-1}\Big| + 2\Theta \sup_{x,n} a_n(x) 3^{-m_n} n^{-1}\sum_{t=1}^{n} |u_{t-1}| = o_P(1), \quad n \to \infty.$$

For obtaining the last bound we used the following two facts: first, it follows from the autoregression equation that

$$n^{-1}\sum_{t=1}^{n} u_{t-1} = (1-\beta)n^{-1}\sum_{t=1}^{n}\varepsilon_t = o_P(1), \quad n \to \infty,$$

and secondly, we employed the definition of $\{\hat{u}_{t-1,j}\}$.

The arguments for the second sum in (21) are quite similar.

Since all the above bounds are uniform in i, j, the proof of Lemma 3.3 is completed. □

It remains to justify (13). We apply the Taylor expansion to the expression under the supremum sign in (13):

$$n^{-1/2}\sum_{t=1}^{n}[A_n(x + n^{-1/2}\tau u_{t-1}) - A_n(x)] = \tau n^{-1}\sum_{t=1}^{n} a_n(x_t)u_{t-1}$$

$$= \tau n^{-1}\sum_{t=1}^{n} a_n(x)u_{t-1} + \tau n^{-1}\sum_{t=1}^{n}[a_n(x_t) - a_n(x)]u_{t-1}, \qquad (23)$$

where x_t is an intermediate point between x and $x + n^{-1/2}\tau u_{t-1}$.

The first sum in (23) is no greater in absolute value than

$$\Theta \sup_{x.n} a_n(x)|n^{-1}\sum_{t=1}^{n} u_{t-1}| = o_p(1), \quad n \to \infty,$$

and the absolute value of the second sum is no greater than

$$L\Theta^2 n^{-3/2}\sum_{t=1}^{n} u_{t-1}^2 = o_p(1), \quad n \to \infty.$$

This proves Theorem 2.1. □

Proof of (7). We will prove (7) for $p = 1$. The proof for an arbitrary p is cumbersome, but presents no principal differences. So, we will consider the $AR(1)$ equation (11).

The LSE $\hat{\beta}_{n,LS}$ in autoregression (11), properly normalized, is representable as

$$n^{1/2}(\hat{\beta}_{n,LS} - \beta) = \frac{n^{-1/2}\sum_{t=1}^{n}\varepsilon_t u_{t-1}}{n^{-1}\sum_{t=1}^{n} u_{t-1}^2}.$$

Let us introduce three mutually independent sequences of i.i.d. r.v.'s: $\{\varepsilon_t^0\}$ with d.f. $G_0(x)$, $\{f_t\}$ with Bernoulli distribution Bern$(n^{-1/2})$, and $\{h_t\}$ with d.f. $H_n(x)$. Then the r.v.'s

$$\tilde{\varepsilon}_t = (1 - f_t)\varepsilon_t^0 + f_t h_t, \quad t \in \mathbb{Z},$$

have the d.f. $A_n(x)$ as in (6), i.e., the sequences $\{\varepsilon_t\}$ under $H_{!n}$ and $\{\tilde\varepsilon_t\}$ have the same distribution. Therefore the sequences

$$\Big\{u_t = \sum_{j\ge 0}\beta^j\varepsilon_{t-j}\Big\} \quad \text{and} \quad \Big\{\tilde u_t := \sum_{j\ge 0}\beta^j\tilde\varepsilon_{t-j}\Big\}.$$

are also equally distributed.

Let $\tilde\beta_{n,LS}$ be the LSE based on $\tilde u_0,\dots,\tilde u_n$. In view of the above statement the r.v.'s $\hat\beta_{n,LS}$ under H_{1n} and $\tilde\beta_{n,LS}$ have the same distribution for any $n\ge 2$.

Now it is not difficult to find the limiting distribution of

$$n^{1/2}(\tilde\beta_{n,LS}-\beta) = \frac{n^{-1/2}\sum_{t=1}^n \tilde\varepsilon_t\tilde u_{t-1}}{n^{-1}\sum_{t=1}^n \tilde u_{t-1}^2}.$$

To this end, consider the LSE $\beta^0_{n,LS}$ based on $u^0_0,\dots,u^0_n$, where $u^0_t=\sum_{j\ge 0}\beta^j\varepsilon^0_{t-j}$. Then, see [1], Chap. 5,

$$n^{1/2}(\beta^0_{n,LS}-\beta) = \frac{n^{-1/2}\sum_{t=1}^n \varepsilon^0_t u^0_{t-1}}{n^{-1}\sum_{t=1}^n (u^0_{t-1})^2} \xrightarrow{d} \mathbb{N}(0,1-\beta^2), \quad n\to\infty,$$

i.e., the asymptotic relation (7) holds. But

$$\begin{aligned} &n^{1/2}(\tilde\beta_{n,LS}-\beta) - n^{1/2}(\beta^0_{n,LS}-\beta) \\ &\quad = \frac{n^{-1/2}\sum_{t=1}^n \tilde\varepsilon_t\tilde u_{t-1}}{n^{-1}\sum_{t=1}^n \tilde u_{t-1}^2} - \frac{n^{-1/2}\sum_{t=1}^n \varepsilon^0_t u^0_{t-1}}{n^{-1}\sum_{t=1}^n (u^0_{t-1})^2} = o_P(1), \quad n\to\infty, \end{aligned}$$

since

$$\mathsf{E}\Big| n^{-1/2}\sum_{t=1}^n \tilde\varepsilon_t\tilde u_{t-1} - n^{-1/2}\sum_{t=1}^n \varepsilon^0_t u^0_{t-1}\Big|^2 \to 0, \tag{24}$$

$$\mathsf{E}\Big| n^{-1}\sum_{t=1}^n \tilde u_{t-1}^2 - n^{-1}\sum_{t=1}^n (u^0_{t-1})^2\Big| \to 0, \quad n\to\infty. \tag{25}$$

We omit the elementary proofs of (24), (25), which follow directly from the definitions of the r.v.'s $\tilde\varepsilon_t$, $\tilde u_{t-1}$, ε^0_t, u^0_{t-1}. Thus the proof of (7) is completed. ☐

ACKNOWLEDGMENTS

The author is sincerely grateful to Prof. D. M. Chibisov for useful discussions.

REFERENCES

1. T. V. Anderson, *The Statistical Analysis of Time Series* (Wiley and Sons, New York, 1971).
2. P. Billingsley, *Convergence of Probability Measures* (Wiley and Sons, New York, 1968).
3. M. V. Boldin, "Estimation of the Distribution of Noise in an Autoregression Scheme", Theory Probab. Appl. **27** (4), 866–871 (1983).
4. M. V. Boldin and M. N. Petriev, "On the Empirical Distribution Function of Residuals in Autoregression with Outliers and Pearson's Chi-Square Type Tests", Math. Methods Statist **27** (4), 1–17 (2018).
5. D. M. Chibisov, "An Investigation of the Asymptotic Power of the Tests of Fit", Theory Probab. Appl. **10** (3), 421–437 (1965).
6. P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Applications* (Academic Press, New York, 1980).