The Annals of Statistics
1989, Vol. 17, No. 4, 1784–1794

WEAK CONVERGENCE OF THE RESIDUAL EMPIRICAL PROCESS IN EXPLOSIVE AUTOREGRESSION

By Hira L. Koul and Shlomo Levental

Michigan State University

This paper proves the weak convergence of the residual empirical process in an explosive autoregression model to the Brownian bridge. As an application the Kolmogorov–Smirnov goodness-of-fit test for testing that the errors have a specified distribution is shown to be asymptotically distribution-free.

1. Introduction. Let F be a distribution function (d.f.) on the real line $\mathbb{R}$, ε be a random variable (r.v.) with d.f. F and $\varepsilon_1, \varepsilon_2, \dots$ be independent copies of ε. In an explosive autoregression model of order 1, one observes r.v.'s $\{X_i\}$ satisfying

$$X_0 = 0, \qquad X_i = \rho X_{i-1} + \varepsilon_i, \qquad |\rho| > 1,\ i \geq 1. \tag{1.1}$$

This model arises in time series analysis and has been discussed previously in the literature. See, e.g., Basawa and Scott (1983) for an interesting discussion of the model.

This paper discusses the weak convergence of the empirical process

$$V_n(y, \hat{\rho}) = n^{-1/2} \sum_{i=1}^{n} I(X_i - \hat{\rho} X_{i-1} \leq y), \qquad y \in \mathbb{R}, \tag{1.2}$$

where $\hat{\rho}$ is an estimator of ρ based on $X_1, X_2, \dots, X_n$ and $I(A)$ is the indicator of the event A. The paper also discusses some applications of the weak convergence result to some problems of statistical inference.

The main weak convergence result for a general F is stated and proved in Section 2 as Theorem 1. As a corollary we get that if F has a uniformly bounded derivative, $E(\log^+(|\varepsilon|))$ is finite where $\log^+(x) := \log(1 \vee x)$, $x \in \mathbb{R}$, and if $|\rho^n(\hat{\rho} - \rho)| = o_p(n^{1/2})$, then

$$\sup_y |V_n(y, \hat{\rho}) - V_n(y, \rho)| \to 0 \quad \text{in probability.} \tag{1.3}$$

As an application consider the problem of testing H_0: $F = \Phi$, where Φ is the d.f. of a $N(0,1)$ r.v. Analogous to the one-sample location model, a test of H_0 could be based on

$$D_n := \sup_y |V_n(y, \hat{\rho}) - n^{1/2}\Phi(y)|.$$

From (1.3) and the well known result that

$$\{V_n(y, \rho) - n^{1/2}\Phi(y),\ y \in \mathbb{R}\} \Rightarrow \{B(\Phi(y)),\ y \in \mathbb{R}\},$$

Received March 1988; revised January 1989.

AMS 1980 *subject classifications*. Primary 60F17; secondary 62M10.

Key words and phrases. Brownian bridge, goodness-of-fit test, asymptotically distribution-free, exponential inequality.

where B is the Brownian bridge on $[0,1]$, it readily follows that if $\hat{\rho}$ is the least squares estimator or any other estimator such that $|\rho^n(\hat{\rho}-\rho)| = o_p(n^{1/2})$, then $D_n \Rightarrow \sup\{|B(u)|;\ 0 \le u \le 1\}$ under H_0. Thus, unlike in the one-sample location model, the asymptotic null distribution of D_n is completely known. This example and its suitable modifications for testing that $\{\varepsilon_i\}$ are $N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, are discussed in Section 3.

A result similar to (1.3), for the stationary autoregression model, where $|\rho| < 1$, was proved by Boldin (1982) requiring $E(\varepsilon) = 0$, $E(\varepsilon^2) < \infty$ and a uniformly bounded second derivative of F. In (1.1), $|\rho| > 1$ and $\{X_i\}$ are nonstationary. Consequently our proofs are necessarily different from those of Boldin (1982). The method of the proof of Theorem 1 uses a version of the chaining argument developed by Giné and Zinn (1984) and an exponential inequality for stopped bounded martingale-differences of Levental (1989).

In the sequel, $o(1)$, $[o_p(1)]$ stands for a sequence of numbers (r.v.'s) converging to zero (in probability): $O(1)$ $[O_p(1)]$ stands for a sequence of numbers (r.v.'s) that are bounded (in probability). For any x, $[x]$ is the greatest integer smaller than x.

2. Weak convergence. In order to state and prove the main result of this section we shall need the following assumptions.

(A.1) $E(\log^+(|\varepsilon|)) < \infty$.

(A.2) F has uniformly bounded derivative f, $f > 0$ a.e.

(A.3) The estimator $\hat{\rho}$ based on $\{X_1, \dots, X_n\}$ is such that
$$|\rho^n(\hat{\rho}-\rho)| = o_p(n^{1/2}).$$

In view of (1.1), we can rewrite

$$V_n(y, \hat{\rho}) = n^{-1/2} \sum_{i=1}^{n} I(\varepsilon_i \le y + \rho^n(\hat{\rho}-\rho)\rho^{-n} X_{i-1}), \qquad y \in \mathbb{R}. \tag{2.1}$$

Thus to study the weak convergence of the processes $V_n(\cdot, \hat{\rho})$, it suffices to investigate the processes

$$S_n(y, s) := n^{-1/2} \sum_{i=1}^{n} I(\varepsilon_i \le y + s\rho^{-n} X_{i-1}), \qquad s, y \in \mathbb{R}. \tag{2.2}$$

One of the main tools used in proving the weak convergence of the preceding processes is an exponential inequality for bounded martingale differences, given in Lemma 3. For this reason it is convenient to center the ith summand in S_n at its conditional expectation, given $\mathscr{F}_{i-1}$, where

$$\mathscr{F}_i := \sigma - \text{field}\{\varepsilon_1, \dots, \varepsilon_i\}, \qquad i \ge 1. \tag{2.3}$$

Accordingly, let

$$\begin{aligned} \mathscr{Y}_{n,i}(y, s) &= I(\varepsilon_i \le y + s\rho^{-n} X_{i-1}) - F(y + s\rho^{-n} X_{i-1}), \qquad i \ge 1, \\ \mathscr{Y}_n(y, s) &= n^{-1/2} \sum_{i=1}^{n} \mathscr{Y}_{n,i}(y, s), \qquad s, y \in \mathbb{R}. \end{aligned} \tag{2.4}$$

Observe that the term corresponding to the sum of the indicators in $\mathcal{Y}_n(y, \hat{t})$, with $\hat{t} = \rho^n(\hat{\rho} - \rho)$, equals $V_n(y, \hat{\rho})$.

In view of (A.3), to prove our main result, we need to show that for any sequence of real numbers $\{a_n\}$ satisfying

(A.4) $a_n = o(n^{1/2})$,

$$\sup_{y \in \mathbb{R},\, |s| \le |a_n|} |\mathcal{Y}_n(y, s) - \mathcal{Y}_n(y, 0)| = o_p(1). \tag{2.5}$$

It is thus convenient for us to rescale the time space and the preceding processes, for any real sequence $\{a_n\}$ and any bounded set of real numbers K, in the following way:

$$T := \{(y, t): y \in \mathbb{R}, t \in K\},$$

$$Z_{n,i}(y, t) := \mathcal{Y}_{n,i}(y, ta_n), \qquad i \ge 1, \tag{2.6}$$

$$Z_n(y, t) := n^{-1/2} \sum_{i=1}^{n} Z_{n,i}(y, t), \qquad (y, t) \in T.$$

We are now ready to state our main result.

THEOREM 1. *Assume that* (A.1), (A.2) *and* (A.4) *hold. Then*

$$\sup_T |Z_n(y, t) - Z_n(y, 0)| = o_p(1). \tag{2.7}$$

COROLLARY 1. *Under* (A.1)–(A.3),

$$\sup_{y \in \mathbb{R}} |V_n(y, \hat{\rho}) - V_n(y, \rho)| = o_p(1). \tag{2.8}$$

Consequently,

$$V_n(\cdot, \hat{\rho}) - n^{1/2}F(\cdot) \Rightarrow B(F(\cdot)).$$

Before proving the preceding results, we state some preliminary facts in the form of various lemmas. Throughout the sequel, $|\rho| > 1$ is fixed.

LEMMA 1. *Under* (A.1),

(a)
$$\sum_{i=1}^{\infty} |\rho^{-i}\varepsilon_i| < \infty \quad a.s.,$$

(b)
$$\sum_{i=1}^{n} \rho^{-n} X_{i-1} \to (\rho - 1)^{-1} Y \quad a.s.,$$

$$\sum_{i=1}^{n} |\rho^{-n} X_{i-1}| \to (|\rho| - 1)^{-1} |Y| \quad a.s.$$

where $Y = \sum_{i=1}^{\infty} \rho^{-i} \varepsilon_i$.

PROOF. The proof of (a) is straightforward, whereas that of (b) uses the Toeplitz lemma and the fact that $\rho^{-i}X_i \to Y$ a.s. □

LEMMA 2. *Assume that* (A.1), (A.2) *and* (A.4) *hold. Then*
(a)

$$\sup_{y,\, t_1,\, t_2} \sum_{i=1}^{n} |F(y + t_2 a_n \rho^{-n} X_{i-1}) - F(y + t_1 a_n \rho^{-n} X_{i-1})| = o(n^{1/2}) \quad a.s.,$$

where the supremum ranges over $-\infty < y < \infty$, $-1 \le t_1, t_2 \le 1$.
(b) *For any* $\varepsilon > 0$,

$$\limsup_{n\to\infty} \sup_{x,\, y,\, t} |n^{-1/2} \sum_{i=1}^{n} F(y + a_n t \rho^{-n} X_{i-1}) - F(x + a_n t \rho^{-n} X_{i-1})| \le \varepsilon \quad a.s.,$$

where the supremum ranges over $|t| \le 1$, x, $y \in \mathbb{R}$, $|F(x) - F(y)| \le \varepsilon n^{-1/2}$.

PROOF. (a) This follows from the mean value theorem (MVT) and Lemma 1(b).
(b) Again, by the MVT,

$$\begin{aligned} \text{L.H.S.(b)} &\le \limsup_{n\to\infty} \left\{ 2|a_n| n^{-1/2} \sum_{i=1}^{n} |\rho^{-n} X_{i-1}| \|f\|_\infty + \sup_{x,\, y} n^{1/2} |F(x) - F(y)| \right\} \\ &\le \varepsilon \quad \text{a.s.}, \end{aligned}$$

by Lemma 1(b) and (A.4). □

The proof of the next lemma appears in Levental (1989).

LEMMA 3. *Let* $(d_i)_{1\le i\le n}$ *be a real-valued martingale-difference sequence with respect to increasing* σ*-algebras* $(\mathcal{Q}_i)_{0\le i\le n}$, i.e., $E(d_i|\mathcal{Q}_{i-1}) = 0$, $i = 1,\dots,n$. *Suppose that* $\|d_i\|_\infty \le M$ *for a constant* $M < \infty$, $i = 1,\dots,n$. *Let* $\tau \le n$ *be stopping time relative to the* $(\mathcal{Q}_i)$ *that satisfies*

$$\left\| \sum_{1\le i\le \tau} E\left(d_i^2 | \mathcal{Q}_{i-1}\right) \right\|_\infty \le L \quad \textit{for a constant } L.$$

Then for each $\gamma \ge 0$,

$$P\left\{ \left| \sum_{1\le i\le \tau} d_i \right| > \gamma \right\} \le 2 \cdot \exp\{-(\gamma/2M) \cdot \operatorname{arcsinh}(M\gamma/2L)\}.$$

PROOF OF THEOREM 1. W.l.o.g. take $K = [-1, 1]$. On T define the metric d by

$$d((y_0, t_0), (y_1, t_1)) = |F(y_0) - F(y_1)|^{1/2} + |t_0 - t_1|^{1/2}. \tag{2.9}$$

Under the metric d, T is totally bounded. Thus, to prove the theorem it suffices to prove:

(a) $(Z_n(y, t) - Z_n(y, 0)) \to 0$, in probability, for every $(y, t) \in T$, and
(b) for every $0 < \varepsilon$ there is $0 < \delta$ such that

$$\limsup_n P\left\{ \sup_{(\delta)} |Z_n(y_0, t_0) - Z_n(y_1, t_1)| > \varepsilon \right\} < \varepsilon,$$

where $(\delta) = \{[(y_0, t_0), (y_1, t_1)] \in T^2\colon d((y_0, t_0), (y_1, t_1)) < \delta\}$.

PROOF OF (a). The fact that Z_n is a sum of conditionally centered Bernoulli r.v.'s yields

$$E\{Z_n(y,t) - Z_n(y,0)\}^2 = E\left[n^{-1}\sum_{i=1}^{n} E\left\{[Z_{n,i}(y,t) - Z_{n,i}(y,0)]^2 \middle| \mathscr{F}_{i-1}\right\}\right]$$
$$\le E\left[n^{-1}\sum_{i=1}^{n} |F(y + ta_n\rho^{-n}X_{i-1}) - F(y)|\right] \to 0$$

by Lemma 2(a) and the dominated convergence theorem. □

PROOF OF (b). The proof of (b) uses a chaining argument developed by Giné and Zinn (1984), which appears in Pollard (1984), pages 160–162. The chaining argument that follows uses the exponential inequality of Lemma 3 for the stopped martingales. Accordingly, fix an $\varepsilon > 0$ and define, for each $n \ge 1$, the grid points

$$\mathscr{H}_n = \{s_{i,j} = (y_i, t_j) \in T: 1 \le i \le [n^{1/2}\varepsilon^{-1}]; \quad -[n^{1/2}\varepsilon^{-1}] \le j \le [n^{1/2}\varepsilon^{-1}]\}, \tag{2.10}$$

where $y_i = F^{-1}(i\varepsilon \cdot n^{-1/2})$, $1 \le i \le [n^{1/2}\varepsilon^{-1}]$, $t_j = j\varepsilon \cdot n^{-1/2}$, $-[n^{1/2}\varepsilon^{-1}] \le j \le [n^{1/2}\varepsilon^{-1}]$ and $F^{-1}(u) := \inf\{x: F(x) \ge u\}$, $0 \le u \le 1$. We also need the following stopping time w.r.t. the filtration $\{\mathscr{F}_k\}$:

$$\tau_n = n \wedge \max\left\{1 \le k: \max_{s,h\in\mathscr{H}_n} \sum_{i=1}^{k} E\left\{[Z_{n,i}(s) - Z_{n,i}(h)]^2 \middle| \mathscr{F}_{i-1}\right\} \div d^2(s,h) < 4n\right\}. \tag{2.11}$$

To adapt our situation to that of Pollard we first prove that $P(\tau_n < n) \to 0$ [see (2.12)]. This will let us work with $n^{-1/2}(\tau_n)^{1/2}Z_{\tau_n}$ instead of with Z_n. By using Lemma 3 and the fact that arcsinh(x) is increasing and concave we can prove the following: if $\varepsilon n^{-1/2} \le d^2(s,h)/x$ for $s, h \in \mathscr{H}_n$, then

$$P\left(n^{-1/2}\tau_n^{1/2}|Z_{\tau_n}(s) - Z_{\tau_n}(h)| \ge x\right) \le 2\exp\{-(x^2/2d^2(s,h))(\varepsilon \cdot \operatorname{arcsinh}(1/8\varepsilon))\}.$$

Now we can chain between the points in $\mathscr{H}_n$ almost as in Pollard. What remains is to connect between each point in T and a point in $\mathscr{H}_n$. This will be done in (2.16). The first step in our program is to show

$$P(\tau_n < n) \to \quad \text{as } n \to \infty. \tag{2.12}$$

PROOF OF (2.12). For $s = (y_{i_0}, t_{j_0})$ and $h = (y_{i_1}, t_{j_1})$, $i_0 < i_1$, we have

$$\sum_{i=1}^{n} E\left\{[Z_{n,i}(s) - Z_{n,i}(h)]^2 \middle| \mathscr{F}_{i-1}\right\}$$
$$\le 2\sum_{i=1}^{n} E\left\{[Z_{n,i}(y_{i_0}, t_{j_0}) - Z_{n,i}(y_{i_1}, t_{j_0})]^2 \middle| \mathscr{F}_{i-1}\right\} + 2\sum_{i=1}^{n} E\left\{[Z_{n,i}(y_{i_1}, t_{j_0}) - Z_{n,i}(y_{i_1}, t_{j_1})]^2 \middle| \mathscr{F}_{i-1}\right\}. \tag{2.13}$$

Consider the first term on the r.h.s. of (2.13): The fact that $d^2(s,h) \geq (i_1 - i_0)\varepsilon n^{-1/2}$ and direct calculation yield

$$\begin{aligned}2\sum_{i=1}^{n} E\Big\{\big[Z_{n,i}(y_{i_0}, t_{j_0}) - Z_{n,i}(y_{i_1}, t_{j_0})\big]^2\Big|\mathscr{F}_{i-1}\Big\}/d^2(s,h)\\ \leq 2[(i_1 - i_0)\varepsilon]^{-1} n^{1/2} \sum_{i=1}^{n} \big\{F(y_{i_1} + a_n t_{j_0}\rho^{-n}X_{i-1})\\ - F(y_{i_0} + a_n t_{j_0}\rho^{-n}X_{i-1})\big\}\\ \leq 2\varepsilon^{-1} n^{1/2} \max_k \Big\{\sum_{i=1}^{n}\big[F(y_{k+1} + a_n t_{j_0}\rho^{-n}X_{i-1})\\ - F(y_k + a_n t_{j_0}\rho^{-n}X_{i-1})\big]\Big\},\end{aligned}$$

where the max is taken over $1 \leq k \leq [n^{1/2}\varepsilon^{-1}]$. By using Lemma 2(b) we get

$$P\Big\{2\max_{i_0, j_0, i_1}\sum_{i=1}^{n} E\Big\{\big[Z_{n,i}(y_{i_0}, t_{j_0}) - Z_{n,i}(y_{i_1}, t_{j_0})\big]^2\Big|\mathscr{F}_{i-1}\Big\} \div d^2(s,h) < 3n\Big\} \to 1. \tag{2.14}$$

By similar arguments, the second term on the r.h.s. of (2.13) is at most

$$2\varepsilon^{-1} n^{1/2} \max_k \Big\{\sum_{i=1}^{n}\big|F(y_{i_1} + a_n t_{k+1}\rho^{-n}X_{i-1}) - F(y_{i_1} + a_n t_k\rho^{-n}X_{i-1})\big|\Big\},$$

where the max is taken over $-[n^{1/2}\varepsilon^{-1}] \leq k \leq [n^{1/2}\varepsilon^{-1}]$. By using Lemma 2(a) we see that

$$P\Big\{2\max_{i_0, j_0, j_1}\sum_{i=1}^{n} E\Big\{\big[Z_{n,i}(y_{i_0}, t_{j_1}) - Z_{n,i}(y_{i_0}, t_{j_0})\big]^2\Big|\mathscr{F}_{i-1}\Big\} \div d^2(s,h) < n\Big\} \to 1. \tag{2.15}$$

This completes the proof of (2.12). □

For each $s = (y_s, t_s) \in T$ denote by $s_{\mathscr{H}_n} = (y_{k_s}, t_{j_s}) \in \mathscr{H}_n$ the point in $\mathscr{H}_n$ that is the closest to s w.r.t. the d-metric from the points of $\mathscr{H}_n$ that satisfy $y_{k_s} \leq y_s$ and $t_{j_s} \leq t_s$. Next we prove the second step in our plan:

$$P\Big\{\sup_{s\in T}\big|Z_n(s) - Z_n(s_{\mathscr{H}_n})\big| \geq 21\varepsilon\Big\} \to 0. \tag{2.16}$$

PROOF OF (2.16). We need to show that

$$P\left\{\sup_{s\in T}\left|\sum_{i=1}^{n}\left[I(\varepsilon_i \le y_s + a_n t_s \rho^{-n} X_{i-1}) - I\left(\varepsilon_i \le y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right) - F(y_s + a_n t_s \rho^{-n} X_{i-1}) + F\left(y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)\right]\right| \ge 21\varepsilon n^{1/2}\right\} \to 0. \tag{2.17}$$

First observe that

$$\begin{aligned}|F(y_s + a_n t_s \rho^{-n} X_{i-1}) - F\left(y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)| &\le |F\left(y_{k_s+1} + a_n t_s \rho^{-n} X_{i-1}\right) - F\left(y_{k_s} + a_n t_s \rho^{-n} X_{i-1}\right)| \\ &\quad + |F\left(y_{k_s} + a_n t_s \rho^{-n} X_{i-1}\right) - F\left(y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)|.\end{aligned}$$

By using Lemma 2 we get

$$P\left\{\sup_{s\in T}\left|\sum_{i=1}^{n}\left[F(y_s + a_n t_s \rho^{-n} X_{i-1}) - F\left(y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)\right]\right| \ge 2\varepsilon n^{1/2}\right\} \to 0. \tag{2.18}$$

So it suffices to show that

$$P\left\{\sup_{s\in T}\left|\sum_{i=1}^{n}\left[I(\varepsilon_i \le y_s + a_n t_s \rho^{-n} X_{i-1}) - I\left(\varepsilon_i \le y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)\right]\right| \ge 19\varepsilon n^{1/2}\right\} \to 0. \tag{2.19}$$

Next, with $\eta_i = a_n \rho^{-n} X_{i-1}$,

$$\begin{aligned}&|I(\varepsilon_i \le y_s + a_n t_s \rho^{-n} X_{i-1}) - I\left(\varepsilon_i \le y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)| \\ &\quad \le I(\eta_i \ge 0)\left\{I\left(y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1} \le \varepsilon_i \le y_{k_s+1} + a_n t_{j_s+1} \rho^{-n} X_{i-1}\right)\right\} \\ &\qquad + I(\eta_i < 0)\left\{I\left(y_{k_s} + a_n t_{j_s+1} \rho^{-n} X_{i-1} \le \varepsilon_i \le y_{k_s+1} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)\right\} \\ &\qquad + I(\eta_i < 0)\left\{I\left(y_{k_s} + a_n t_{j_s+1} \rho^{-n} X_{i-1} \le \varepsilon_i \le y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1}\right)\right\}.\end{aligned} \tag{2.20}$$

We will show that

$$P\left\{\sup_{s\in T}\sum_{i=1}^{n}\left[I(\eta_i \ge 0)\left\{I\left(y_{k_s} + a_n t_{j_s} \rho^{-n} X_{i-1} \le \varepsilon_i \le y_{k_s+1} + a_n t_{j_s+1\rho^{-n}Xi-1}\right)\right\}\right] \ge 6\varepsilon n^{1/2}\right\} \to 0. \tag{2.21}$$

By Lemma 2 we see that

$$P\left\{\sup_{s\in T}\sum_{i=1}^{n}\left[I(\eta_i\geq 0)\left\{F\left(y_{k_s+1}+a_n t_{j_s+1}\rho^{-n}X_{i-1}\right)\right.\right.\right.$$
$$\left.\left.\left.-F\left(y_{k_s}+a_n t_{j_s}\rho^{-n}X_{i-1}\right)\right\}\right]\geq 2\varepsilon n^{1/2}\right\}\to 0.$$

Now (2.21) will follow from

$$P\left\{\max_{i_o,j_0}\left|\sum_{i=1}^{n}I(\eta_i\geq 0)\left[Z_{n,i}\left(y_{i_0+1},t_{j_0+1}\right)-Z_{n,i}\left(y_{i_0},t_{j_0}\right)\right]\right|\geq 3\varepsilon n^{1/2}\right\}\to 0. \tag{2.22}$$

Because of (2.12) it suffices to show that

$$P\left\{\max_{i_o,j_0}\left|\sum_{1\leq i\leq\tau_n}I(\eta_i\geq 0)\left[Z_{n,i}\left(y_{i_0+1},t_{j_0+1}\right)-Z_{n,i}\left(y_{i_0},t_{j_0}\right)\right]\right|\geq 3\varepsilon n^{1/2}\right\}\to 0. \tag{2.23}$$

But (2.23) follows from Lemma 3: Take $\tau := \tau_n$ as defined in (2.11), $L := 8\varepsilon n^{1/2}$, $\gamma := 3\varepsilon n^{1/2}$ and $M = 1$. For each (i_0, j_0) we get

$$P\left\{\sum_{1\leq i\leq\tau_n}I(\eta_i\geq 0)\left[Z_{n,i}\left(y_{i_0+1},t_{j_0+1}\right)-Z_{n,i}\left(y_{i_0},t_{j_0}\right)\right]\geq 3\varepsilon n^{1/2}\right\}$$
$$\leq 2\exp\left\{-(3\varepsilon/4)n^{1/2}\operatorname{arcsinh}(3/16)\right\}.$$

Since the last bound does not depend on (i_0, j_0) we can bound (2.23) by $(2n/\varepsilon^2)\cdot 2\exp\{-(3\varepsilon/4)n^{1/2}\operatorname{arcsinh}(3/16)\}\to 0$, which proves (2.21). We can prove results similar to (2.21) for the rest of the terms on the r.h.s. of (2.20). This completes the proof of (2.19) thereby establishing (2.16) and hence Theorem 1. □

PROOF OF COROLLARY 1. By (A.3), there exists a positive sequence $b_n \to 0$ such that $P(|n^{-1/2}\rho^n(\hat\rho-\rho)|\leq b_n)\to 1$ as $n\to\infty$. Now apply Lemma 2(a) and Theorem 1 with $a_n = n^{1/2}b_n$ and

$$t = n^{-1/2}\rho^n(\hat\rho-\rho)b_n^{-1}I\left(|n^{-1/2}\rho^n(\hat\rho-\rho)|\leq b_n\right).$$ □

3. Applications. First consider an extension of the model (1.1) where

$$(3.1)\quad Y_0 = 0, \qquad Y_i = \mu + \rho Y_{i-1} + \sigma\varepsilon_i, \qquad \mu \in \mathbb{R},\ \sigma > 0,\ |\rho| > 1,\ i \geq 1.$$

Let $\hat{\theta} = (\hat{\mu}, \hat{\rho}, \hat{\delta})$ be estimators of $\theta = (\mu, \rho, \sigma)$, respectively. The process of interest here is

$$\begin{aligned} W(y, \hat{\theta}) &:= n^{-1/2} \sum_{i=1}^{n} I\{Y_i \leq y\hat{\sigma} + \hat{\mu} + \hat{\rho}Y_{i-1}\} \\ &= n^{-1/2} \sum_{i=1}^{n} I\{(y_i - \mu)/\sigma \leq y\hat{\sigma}/\sigma + (\hat{\mu} - \mu)/\sigma + (\hat{\rho}/\sigma)Y_{i-1}\} \\ &= n^{-1/2} \sum_{i=1}^{n} I\{(Y_i - \mu)/\sigma \leq y\hat{\sigma}/\sigma + (\hat{\mu} - \mu)/\sigma + \hat{\rho}\mu/\sigma \\ &\qquad\qquad + \hat{\rho}(Y_{i-1} - \mu)/\sigma\}. \end{aligned}$$

Thus if we identify $(Y_i - \mu)/\sigma$ with X_i of (1.1), then

$$W(y, \hat{\theta}) \equiv V_n(y\hat{\sigma}/\sigma + (\hat{\mu} - \mu)/\sigma + \hat{\rho}\mu/\sigma, \hat{\rho}).$$

Consequently we have the following.

COROLLARY 2. *Assume that* (3.1) *holds and that* $\hat{\mu}$, $\hat{\rho}$ *and* $\hat{\delta}$ *are such that*

$$(\text{A.5})\qquad n^{1/2}\{|\hat{\mu} - \mu| + |\hat{\sigma}\sigma^{-1} - 1|\} + |\rho^n(\hat{\rho} - \rho)| = O_p(1).$$

Furthermore assume that

(A.6) *F, the d.f. of ε_1 in* (3.1), *is strictly increasing and has a bounded derivative f.*

Then

$$\sup_y |W(y, \hat{\theta}) - W(y, \theta) - n^{1/2}\{(\hat{\mu} - \mu) + (\hat{\sigma} - \sigma)y\}\sigma^{-1}f(y)| = o_p(1).$$

PROOF. This follows from Corollary 1 applied to $X_i \equiv (Y_i - \mu)/\sigma$, the tightness of the $Z_n(\cdot, 0)$processes and the fact that $n^{1/2}(\hat{\rho} - \rho) = o_p(1)$, in a routine fashion. □

Now consider the model (1.1) and the problem of testing H_0: $F = F_0$, where F_0 is a known d.f. and the statistic

$$T_n := \sup_y |V_n(y, \hat{\rho}) - n^{1/2}F_0(y)|.$$

Corollary 1 implies that F_0 and $\hat{\rho}$ satisfy (A.1)–(A.3). Then $T_n \Rightarrow \sup_{0 \leq u \leq 1}|B(u)|$, under H_0. Hence the test of H_0 based on T_n is asymptotically distribution-free.

Next, consider (1.1) and H_{01}: $F = N(\mu, 1)$, $\mu \in \mathbb{R}$. This is equivalent to assuming (3.1) with $\sigma = 1$ and testing that $\{\varepsilon_i\}$ of (3.1) are i.i.d. $N(0,1)$. In this case the test statistic is

$$T_{n1} := \sup_y |W(y, \hat{\theta}) - n^{1/2}\Phi(y)|,$$

where now $\hat{\theta} = (\hat{\mu}, \hat{\rho}, 1)$ and

$$\hat{\mu} = \overline{Y} - \hat{\rho}\overline{Y}_1, \qquad \overline{Y} = n^{-1}\sum_{i=1}^{n} Y_i, \qquad \overline{Y}_1 = n^{-1}\sum_{i=1}^{n} Y_{i-1},$$

(3.2)

$$\hat{\rho} = \left[\sum_{i=1}^{n}(Y_{i-1} - \overline{Y}_1)(Y_i - \overline{Y})\right] \Big/ \left[\sum_{i=1}^{n}(Y_{i-1} - \overline{Y}_1)^2\right].$$

From Corollary 2, under H_{01},

$$T_{n1} = \sup_y \left| n^{-1/2}\sum_{i=1}^{n}\{I(\varepsilon_i \le y) - \Phi(y) + \varepsilon_i\varphi(y)\} - n^{-1/2}(\hat{\rho} - \rho)\sum_{i=1}^{n} Y_{i-1}\varphi(y)\right| + o_p(1),$$

where φ is the density of Φ. But under H_{01} and (3.1),

$$n^{-1/2}(\hat{\rho}_n - \rho)\sum_{i=1}^{n} Y_{i-1} = n^{-1/2}\rho^n(\hat{\rho}_n - \rho)\sum_{i=1}^{n}\rho^{-n}Y_{i-1} = o_p(1).$$

From Durbin (1973) it follows that $T_{n1} \Rightarrow \sup_y|G(y)|$, where G is a mean zero Gaussian process with covariance function $\Phi(x)(1 - \Phi(y)) - \varphi(x)\varphi(y)$, $x \le y$. The distribution of $\sup_y|G(y)|$ is available in Durbin (1973).

Finally consider the model (1.1) and the problem of testing H_{02}: $F = N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$. Consider the test statistic

$$T_{n2} := \sup_y |W(y, \hat{\theta}) - n^{1/2}\Phi(y)|,$$

where now $\hat{\theta} = (\hat{\mu}, \hat{\rho}, \hat{\sigma})$, with $\hat{\mu}$, $\hat{\rho}$ as in (3.2) and $\hat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}(Y_i - \hat{\mu} - \hat{\rho}Y_{i-1})^2$. Now arguing as for T_{n1}, one finds the limiting null distribution of T_{n2} to be similar to its analogue in the one-sample location-scale model.

It goes without saying that analogous results remain valid for any other goodness-of-fit test statistics based on the empirical process of the residuals. Finally we point out that the results of Section 2 are general enough to allow the investigation of the asymptotic power of the preceding tests, although we do not discuss it here.

REFERENCES

BASAWA, I. V and SCOTT, D. J. (1983). *Asymptotic Optimal Inference for Nonergodic Models. Lecture Notes in Statist.* **17**. Springer, New York.

BOLDIN, M. V. (1982). Estimation of the distribution of noise in an autoregression scheme. *Theory Probab. Appl.* **27** 866–871.

DURBIN, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function.* SIAM, Philadelphia.

GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–989.

LEVENTAL, S. (1989). A uniform CLT for uniformly bounded families of martingale-differences. *J. Theoret. Probab.* **2** 271–287.

POLLARD, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

DEPARTMENT OF STATISTICS
AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824