A Bootstrap View on Dickey-Fuller Control Charts for AR(1) Series

Ansgar Steland

RWTH Aachen University, Germany

Abstract: Dickey-Fuller control charts aim at monitoring a time series until a given time horizon to detect stationarity as early as possible. That problem appears in many fields, especially in econometrics and the analysis of economic equilibria. To improve upon asymptotic control limits (critical values), we study the bootstrap and establish its a.s. consistency for fixed alternatives. Simulations indicate that the bootstrap control chart works very well.

Keywords: Autoregressive Time Series, Invariance Principle, Least Squares, Resampling, Unit Root.

1   Introduction

We study the bootstrap for a sequential detection procedure with finite time horizon (truncated control chart) which aims at detecting whether the in-control model of a random walk (unit root) is violated in favor of stationarity. The bootstrap is known to provide easy-to-use and accurate procedures, and its application to the design of monitoring procedures (control charts) is promising.

The problem to study truncated control charts for random walks has several motivations. Often, it is not clear whether an econometric time series has a unit root or is stationary. In many practical applications monitoring stops latest when a known pre-specified time horizon is reached. For instance, in portfolio management continuous or near-continuous trading is often not feasible, because time- and cost-intensive (econometric) analyses are required to obtain valid distributional model specifications for the underlying time series, in particular with regard to their degree of integration. Instead, the portfolio is updated, e.g., only on a quarterly basis. Between these updates it is advisable to apply monitoring procedures to influential time series. Additional trades, hedges, or further analyses can then be initiated if a monitoring rule (control chart) provides a signal. A further interesting application is to monitor equilibrium errors of a (known) co-integration relationship in order to confirm co-integration in the sense of Granger (1981), i.e. stationarity of equilibrium errors. Finally, the question 'random walk or stationary process' is also crucial to choose a valid method when analyzing a series of observations to detect trends. These methods usually assume stationarity as in Steland (2004), Steland (2005b), Pawlak et al. (2004), Hušková (1999), Hušková and Slabý (2001), or Ferger (1995), among others. When detecting trends in random walks by, e.g., a kernel smoother, the results change drastically as shown in Steland (2005c).

Detection rules using (modifications of) the Dickey-Fuller test statistic have been studied in (Steland, 2005a). To improve the detection quality it is advisable to introduce a weighting mechanism which assigns smaller weights to past contributions than to more recent ones. Simulations revealed that the bandwidth parameter h controlling the weights

and the start of monitoring k have to be chosen carefully when using asymptotic critical values. The bootstrap should avoid this problem, since it resamples the procedure as it is. Whereas for fixed sample Dickey-Fuller type unit root tests the bootstrap has been studied, see Park (2003) and the references given there, the present article establishes a first consistency result for a bootstrap scheme of the underlying (weighted) Dickey-Fuller process and a related control chart.

2 Model and Dickey-Fuller Type Detection Procedures

Let us consider an autoregressive time series model of order 1, AR(1), given by

$$Y_{t+1} = \rho Y_t + u_t, \qquad t \geq 1, \tag{1}$$

where $\{u_t\}$ are independent and identically distributed error terms with common variance $0 < \sigma^2 < \infty$. We assume $Y_0 = 0$. The autoregressive parameter, ρ, determines whether the time series Y_t is stationary ($|\rho| < 1$) or a random walk ($\rho = 1$). The well known least-squares based Dickey-Fuller test for testing the unit root null hypothesis $H_0 : \rho = 1$ against the alternative of stationarity $H_1 : |\rho| < 1$ based on a sample $Y_1, \ldots, Y_T$ of (fixed) size T is the Dickey-Fuller test (DF test) which is given by

$$\widehat{\rho}_T = \sum_{t=1}^{T} Y_{t-1} Y_t \left/ \sum_{t=1}^{T} Y_{t-1}^2 \right., \qquad D_T = T(\widehat{\rho}_T - 1) = \frac{\sum_{t=1}^{T} Y_{t-1} \Delta Y_t}{\frac{1}{T} \sum_{t=1}^{T} Y_{t-1}^2},$$

where $\Delta Y_t = Y_t - Y_{t-1}$. D_T has a nonstandard limit distribution under H_0, if $T \to \infty$. H_0 is rejected for small (negative) values of D_T. Usually, one relies on asymptotic theory, i.e., one approximates the appropriate critical value c by the corresponding quantile of the the asymptotic distribution for $T \to \infty$. The DF test and its asymptotic properties, particularly its non-standard limit distribution have been studied by White (1958), Fuller (1976), Rao (1978), Dickey and Fuller (1979), Chan and Wei (1988), among others.

The weighted Dickey-Fuller control chart as introduced in Steland (2005a) is defined as follows. For a control limit c define for $t = k, k+1, \ldots, T$

$$S_T = \inf\{k \leq t \leq T : D_t < c\}, \qquad D_t = \frac{\frac{1}{t} \sum_{s=1}^{t} K[(t-s)/h] Y_{s-1}(Y_s - Y_{s-1})}{\frac{1}{t^2} \sum_{s=1}^{t} Y_{s-1}^2},$$

for some non-negative kernel function $K : [0, \infty) \to \mathbb{R}_0^+$. The detection procedure S_T gives a signal at time t, if the weighted Dickey-Fuller test statistic D_t is smaller than the control limit c for the first time. Monitoring starts at the k-th observation and ends at the latest at the time horizon T. Here and in the sequel we assume that c is chosen to attain a given nominal type I error that a false signal is given, at least asymptotically, i.e., $\lim_{T \to \infty} P_0(S_T < T) = \alpha$. However, the modifications to control other characteristics as the average run length or the median run length are straightforward.

The kernel function K is used to downweight past contributions to the sum, which improves the detection properties in change-point models, see Steland (2005a). Usually one takes a decreasing density function with unique maximum at 0. We assume the following regularity assumptions for K:

(K) $0 < \int_0^\infty K(z)dz < \infty$, $K \in C^2$, $\|K''\|_\infty < \infty$, and $\int |dK| < \infty$.
The bandwidth parameter h appearing in the definition of the kernel weights $K[(t-s)/h]$ may be absorbed into the kernel, but it is more common to rescale a given (standardized) kernel by a bandwidth. Notice that the procedure uses the most recent h observations, if K has support $[0, 1]$. In this sense h defines the memory of the control chart. We assume that $h = h_T$ satisfies

$$T/h_T \to \zeta \in (1, \infty),$$

as $T \to \infty$. The memory parameter ζ is chosen in advance.

In the sequel let $\lfloor x \rfloor$, $x \in \mathbb{R}$, denote the floor function. Introduce the stochastic process related to D_T obtained by rescaling time,

$$D_T(s) = D_{\lfloor Ts \rfloor}, \qquad s \in [0, 1].$$

Note that the trajectories of D_T are constant on intervals of the form (a, b), $a < b$, thus being elements of the function space $D[0, 1]$ of right-continuous functions with existing left-hand limits. We have the representation $S_T = T \inf\{s \in [\kappa, 1] : D_T(s) < c\}$, which allows to base the stochastic analysis of S_T on the distributional properties of S_T from the properties of D_T. Further, we assume that the start of monitoring, k, is given by

$$k = \lfloor T\kappa \rfloor, \qquad \text{for some } \kappa \in (0, 1).$$

Equip $D[\kappa, 1]$ with the Skorokhod topology induced by the Skorokhod metric d. Weak convergence of a sequence $X, X_T : (\Omega, \mathcal{A}, P) \to D[0, 1]$ under the probability P is denoted by $X_T \overset{w,P}{\to} X$, $T \to \infty$. For details we refer to Billingsley (1968) or Jacod and Shiryaev (2003). Further, $\{B(s) : s \in [0, 1]\}$ is a standard Brownian motion with $B(0) = 0$.

Assuming the same AR(1) model the following functional central limit theorem providing the asymptotic distribution of the Dickey-Fuller process was shown in Steland (2005a), where it was used to choose the control limit based on the asymptotic distribution.

Theorem 2.1 *Assume either $K(z) = z$ for all z or (K). Then, as $T \to \infty$,*

$$D_T(s) \overset{w,P}{\to} \mathcal{D}(s) = \int_0^s K[\zeta(r-s)]B(r)\, dB(r) \Big/ \int_0^s B(r)^2\, dr, \qquad s \in [0, 1],$$

and, as $T \to \infty$,

$$S_T/T \overset{d}{\to} \inf\{s \in [\kappa, 1] : \mathcal{D}(s) < c\}.$$

Remark 2.1 *Assume $K(z) = z$ for all $z \in \mathbb{R}$. Then the limit process is given by*

$$\mathcal{D}(s) = (2s)^{-1}[B(s)^2 - s] \Big/ \int_0^s B(r)^2\, dr, \qquad s \in [0, 1].$$

3   The Bootstrap

The basic idea of the bootstrap is to invest in computing resources to obtain a resampling based estimate of the control limit satisfying certain statistical properties, instead of relying on the asymptotic distribution as an approximation. In the sequential (monitoring) setting considered in this article the data $Y_1, \ldots, Y_T$ are observed sequentially, and at each time point $t \in \{1, \ldots, T\}$ the current data $Y_1, \ldots, Y_t$ are used to decide whether the in-control model holds or not. Thus, it seems natural to calculate at each time point a $\sigma(Y_1, \ldots, Y_t)$-measurable bootstrap estimate $\widehat{c}_t$ of the control limit c, based on all available observations Y_s, $s \leq t$, which implicitly have been classified as in-control data. Then, at each time point we consider the stopping rule

$$\widehat{S}_T = \inf\{k \leq t \leq T : D_t < \widehat{c}_t\}.$$

At each monitoring step there are at least k observations from which one can bootstrap.

An appropriate bootstrap scheme has to generate bootstrap samples $Y_1^*, \ldots, Y_T^*$ which mimic the distributional behavior of $Y_1, \ldots, Y_{T^*}$ under the null hypothesis (in-control model). Allowing for a maximum bootstrap sample size T^* different from T will simplify the proofs and sometimes increases the accuracy of the bootstrap approximation, but usually one takes $T^* = T$. At each time point t perform the following steps.

1. Draw a bootstrap sample $u_1^*, \ldots, u_{T^*}^*$ (with replacement) from the observed centered differences, $\widehat{u}_i = \Delta Y_i - \overline{\Delta Y}$, $i = 1, \ldots, t$, $\overline{\Delta Y} = t^{-1} \sum_{s=1}^{t} \Delta Y_s$.

2. Put $Y_r^* = Y_{r-1}^* + u_r^*$, $r = 1, \ldots, T^*$, $Y_0^* = 0$.

3. Calculate the trajectory of the associated bootstrap process

$$D_{T^*}^*(s) = \frac{\lfloor T^* s \rfloor^{-1} \sum_{r=1}^{\lfloor T^* s \rfloor} K[(\lfloor T^* s \rfloor - r)/h] Y_{r-1}^* \Delta Y_r^*}{\lfloor T^* s \rfloor^{-2} \sum_{r=1}^{\lfloor T^* s \rfloor} Y_{r-1}^{*2}}, \qquad s \in (0, 1], \quad (2)$$

$D_{T^*}^*(0) = 0$, at the time points $s = t/T$, $t = 1, \ldots, T$.

4. Calculate $m^* = \min_{k \leq r \leq T} D_{T^*}^*(r/T)$.

5. Repeat the resampling step B times to obtain B replicates $m^*(1), \ldots, m^*(B)$ and estimate the control limit by $\widehat{c}_{tB}$, the empirical α-quantile of $m^*(1), \ldots, m^*(B)$.

The bootstrap scheme defines a bootstrap probability $P^* = P_t^*$ for given $Y_1, \ldots, Y_t$. Under P^*, the common distribution function of the u_i^* is given by the empirical distribution function, $\widehat{F}_t$, of the values $\widehat{u}_r = \Delta Y_r - \overline{\Delta Y}$, $r = 1, \ldots, t$.

Remark 3.1 *Note that the bootstrap is based on the H_0-residuals $\widehat{u}_t$ calculated under H_0. Usually, one has to bootstrap from residuals being consistent for the error terms also under H_1, i.e., bootstrap from the $AR(1)$-residuals $\widehat{u}_t = Y_t - \widehat{\rho} Y_{t-1}$ after centering, for some consistent estimator $\widehat{\rho}$. Nevertheless, our main result shows that the simplified bootstrap is consistent, which is a special feature of the statistic D_t.*

For a bootstrap sample $Y_1^*, \ldots, Y_{T^*}^*$ the bootstrapped detection rule associated to the process defined in (2) is given by

$$S_{T^*}^* = \inf\{k \leq t \leq T^* : D_{T^*}^*(t/T^*) < c\}.$$

The following main result yields P-a.s. consistency of the proposed bootstrap scheme. It asserts that under the bootstrap probability P_t^* the process $D_{T^*}^*$ converges weakly to $\mathcal{D}$, as $T^* \to \infty$. For a given specification of the conditional probability P_t^* the result holds true for every sample $Y_1, \ldots, Y_t$, thus P-a.s. for any conditional probability.

Theorem 3.1 *Fix t. Assume $0 < \rho \le 1$. Under P^* (given $Y_1, \ldots, Y_t$) we have P-a.s.*

$$D_{T^*}^* \overset{w,P^*}{\to} \mathcal{D}$$

and

$$S_{T^*}/T^* \overset{w,*}{\to} \inf\{s \in [\kappa, 1] : \mathcal{D}(s) < c\} \,,$$

as T^* tends to ∞ such that $T^*/h \to \zeta$.

The consistency of the detection rule using the bootstrapped control limit appears now as a Corollary.

Corollary 3.1 *Under the assumptions of Theorem 3.1 the bootstrap estimate of the control limit provides an a.s. consistent procedure in the sense that under the bootstrap probability P^* we have P-a.s.*

$$\widehat{S}_T/T \overset{w,P^*}{\to} \inf\{s \in [\kappa, 1] : \mathcal{D}(s) < c\} \,, \qquad T \to \infty \,.$$

4 Simulations

In our simulation study we focused on a small (maximum) sample size to investigate the type I error of a false detection if the series is a random walk and the conditional average run length (CARL), i.e., the conditional expected run length, $\mathrm{E}(\widehat{S}_T | \widehat{S}_T < T)$, given that the procedure provides a signal, calculated under the alternative. The CARL says after how many observations, on average, we get the signal, if the procedure rejects the unit root hypothesis, and it is very informative under the alternative. We study both bootstrapping from differences and $AR(1)$-residuals.

The most recent bootstrap replicates are almost as informative as the current ones, since they are drawn from a very similar set of observations. Thus, we used the following simulation approach. When monitoring starts, a bootstrap sample of size 10000 using the available k observations is drawn and stored. At time $t = k + l \cdot 20$, $l \in \mathbb{N}$, i.e., at every $20th$ observation, the 10% oldest bootstrap replicates are replaced by 1000 new ones using the available t observations $Y_1, \ldots, Y_t$. Using the $AR(1)$ model (1) time series of length $T = 150$ and two settings with $T = 250$ with $N(0, 1)$-distributed error terms u_t were simulated for $\rho = 1$ (null hypothesis) and $\rho = 0.9$ (alternative), respectively. The parameters h (bandwidth) and k (start of monitoring) were taken from $\{25, 50\}$. Each table entry is based on 10000 repetitions.

The results indicate that the bootstrap works very well for a wide range of combinations of the parameters T, h, and k. Even for small k accuracy is good. Bootstrapping from $AR(1)$-residuals seems to provide slightly more accurate tests in terms of the type I error, but our simulations do not indicate an advantage in terms of power. The CARL

Table 1: Type I error, power, and conditional average run length given a signal.

| | | | H_0-residuals | | | $AR(1)$-residuals | | |
T	h	k	type I error	power	CARL	type I error	power	CARL
Epanechnikov kernel:								
150	25	25	0.0462	0.1878	17.0	0.0467	0.1849	17.0
150	50	25	0.0480	0.224	21.8	0.0452	0.2326	20.7
150	25	50	0.0452	0.5235	32.6	0.0487	0.5288	32.5
150	50	50	0.0463	0.4461	23.5	0.0486	0.4514	22.8
Gaussian kernel:								
150	25	25	0.0506	0.2327	22.3	0.0514	0.2367	23.7
150	50	25	0.0528	0.3943	43.7	0.0504	0.4037	43.8
150	25	50	0.0463	0.4821	25.1	0.0451	0.4739	25.5
150	50	50	0.0471	0.5748	32.0	0.0496	0.5756	32.4
250	50	25	0.0525	0.4291	57.1	0.0491	0.4312	55.7
250	50	50	0.0468	0.6411	48.5	0.0483	0.6407	47.6

values indicate a trade-off: The CARL is smaller for smaller h, but the power decreases. Using fewer recent observations yields a loss of power, but if we reject, we get this decision very early. For the same reason, the Gaussian kernel provides more power but higher CARLs than the Epanechnikov kernel which has bounded support.

Acknowledgements

Appendix: Proofs

We apply the following result due to Kurtz and Protter (1996 resp. 2004, Sec. 7).
Theorem (Kurtz and Protter). Suppose X_n is a semimartingale for each n, and H_n is predictable for each n. If $(H_n, X_n) \overset{w}{\to} (H, X)$ in the Skorokhod space $D_{\mathbb{R}^2}[0,1]$, and if $\sup_n \mathrm{var}(X_n) < \infty$, then, as $n \to \infty$,

$$\left(H_n, X_n, \int H_n \, dX_n \right) \overset{w}{\to} \left(H, X, \int H \, dX \right).$$

Proof (of Theorem 3.1). We first consider the case $\rho = 1$. The bootstrap defines an array $\{ Y_{ti}^* : i = 1, \ldots, T^*, \, t = 1, \ldots, T \}$, of bootstrap observations $Y_{ti}^* = \sum_{j=1}^{i} u_j^*$. In the sequel we will suppress the t in the notation. Donsker's theorem yields

$$T^{-1/2} \sum_{i=1}^{\lfloor Ts \rfloor} u_i \overset{w,P}{\to} \sigma B(s), \quad T \to \infty, \qquad (T^*)^{-1/2} \sum_{i=1}^{\lfloor T^* s \rfloor} u_i^* \overset{w,P^*}{\to} s_t B^*(s), \quad T^* \to \infty,$$

since under P^* the r.v.s. u_i^* are iid with mean $\overline{\overline{u}} = 0$ and common variance $s_t^2 = \frac{1}{t} \sum_{i=1}^{t} \widehat{u}_i^2$. Here B and B^* are two Brownian motions with $B(0) = 0$. Define the bootstrap process

$$Z_{T*}^*(s) = (T^*)^{-1/2} Y_{\lfloor T^* s \rfloor}^*, \qquad s \in [0, 1],$$

and the filtration $\mathcal{F}_r^* = \sigma(u_1, \ldots, u_{\lfloor Tr \rfloor})$, $r \in [0, 1]$. By definition of $Y_i^* = \sum_{j=0}^{i} u_j^*$, $E^*[Z_{T^*}^*(s)|\mathcal{F}_r^*] = Z_{T^*}^*(r)$, and $E^*[(T^*)^{-1/2} Z_{T^*}^*(s)]^2 = E^*(u_1^{*2})\lfloor T^* s \rfloor / T$, i.e., Z_{T^*} and $K[(\lfloor T^* \circ \rfloor - \lfloor T^* s \rfloor)/h] Z_{T^*}^*(\circ)$ are L_2-martingales wrt P^*. Thus, for each T^* the numerator $(\lfloor T^* s \rfloor)^{-1} \sum_{i=1}^{\lfloor T^* s \rfloor} K[(\lfloor T^* s \rfloor - i)/h] Y_{i-1}^* \Delta Y_i^*$ of $D_{T^*}^*$ can be represented as Ito integral, $\lfloor T^* s \rfloor^{-1} T \int_0^s K[(\lfloor T^* s \rfloor - \lfloor T^* r \rfloor)/h] Z_{T^*}^*(r) \, dZ_{T^*}^*(r)$. Let us verify fidi-convergence of

$$s \mapsto N_{T^*}^*(s) = \int_0^s K[(\lfloor T^* s \rfloor - \lfloor T^* r \rfloor)/h] Z_{T^*}^*(r) \, dZ_{T^*}^*(r), \qquad s \in [\kappa, 1].$$

Consider for $s \in [\kappa, 1]$ the random function

$$r \mapsto I_{T^*}^*(r; s) = K[(\lfloor T^* s \rfloor - \lfloor T^* r \rfloor)/h] Z_{T^*}^*(r). \qquad (3)$$

By the Skorokhod/Dudley/Wichura representation theorem, we may assume that $\|Z_{T^*}^* - s_t B^*\|_\infty \to 0$, $T^* \to \infty$, P^*-a.s. Then, since K is Lipschitz continuous,

$$\sup_{s \in [\kappa,1], \, r \in [0,1]} \left| K[(\lfloor T^* s \rfloor - \lfloor T^* r \rfloor)/h] Z_{T^*}^*(r) - s_t K[\zeta(s-r)] B^*(r) \right| \to 0,$$

if $T^* \to \infty$, P^*-a.s. Hence, $(Z_{T^*}^*(\circ), I_{T^*}^*(\circ, s)) \to (s_t B^*(\circ), K[\zeta(s-\circ)] s_t B^*(\circ))$, in the supnorm, and even uniformly in $s \in [\kappa, 1]$. It follows that for each fixed $s \in [\kappa, 1]$ the triple $(Z_{T^*}^*(\circ), I_{T^*}^*(\circ, s), N_{T^*}^*(s))$ converges in distribution to

$$\left(s_t B^*(\circ), K[\zeta(s-\circ)] s_t B^*(\circ), s_t^2 \int_0^s K[\zeta(s-r)] B(r) \, dB(r) \right),$$

under P^*, as $T^* \to \infty$. By taking linear combinations, this can be extended to convergence of all fidis, yielding fidi convergence of $N_{T^*}^*$. Since $\{I_{T^*}^*\}$ and $\{Z_{T^*}^*\}$ are tight, $(I_{T^*}^*, Z_{T^*}^*)$ is also tight implying tightness of $N_{T^*}^* = \int I_{T^*}^* \, dZ_{T^*}^*$. Thus, the numerator converges weakly to the random element $s_t^2 \int_0^s K[\zeta(s-r)] B^*(r) \, dB^*(r)$. Further, the denominator of $D_{T^*}^*$ is a continuous functional of $Z_{T^*}^*$, $\lfloor T^* s \rfloor^{-2} \sum_{i=1}^{\lfloor T^* s \rfloor} Y_{i-1}^{*2} = \int_0^s (Z_{T^*}^*(u))^2 \, du \overset{w,P^*}{\to} s_t^2 \int_0^s B(u)^{*2} du$, as $T^* \to \infty$. The continuous mapping theorem yields that $(\int_0^\circ (Z_{T^*}^*(u))^2 \, du, N_{T^*}^*(\circ))$ converges weakly. Thus, we can conclude that

$$D_{T^*}^*(\circ) = \frac{N_{T^*}^*(\circ)}{\int_0^\circ (Z_{T^*}^*(r))^2 \, dr} \overset{w,P^*}{\to} \frac{s_t^2 \int_0^\circ K[\zeta(\circ - r)] B(r) \, dB(r)}{s_t^2 \int_0^\circ B(r)^2 \, dr},$$

as $T^* \to \infty$. Note that s_t cancels in the limit. Hence the result holds for all t. Noting that the limit process has a.s. continuous sample paths, convergence in distribution of S_T follows, for details we refer to Steland (2004).

It remains to verify the weak limit for the case $0 < \rho < 1$. First note that $\Delta Y_t = (\rho - 1) Y_{t-1} + u_t = \sum_{i \geq 0} (\rho - 1)^i u_{t-i}$ is a stationary AR(1) process if $0 < \rho < 1$, implying $\overline{\Delta Y} \overset{P}{\to} 0$. Hence, $\{\Delta Y_i - \overline{\Delta Y} : i = 1, \ldots, t\}$ is a stationary series. Thus, the iid bootstrap values $u_t^* \sim F_t$ satisfy an invariance principle, $(T^*)^{-1/2} \sum_{i=1}^{\lfloor T^* s \rfloor} u_i^* \overset{w,P^*}{\to} s_t' B^*(s)$, $T^* \to \infty$, with $s_t' = \text{var}(\Delta Y_1 - \overline{\Delta Y})$. Since s_t' cancels in the limit process, for $0 < \rho < 1$ we obtain the same limit process for $D_{T^*}^*$ under P^* as for $\rho = 1$. $\quad\square$

Proof of Corollary 3.1. The corollary is shown as follows. Define $\widehat{c}_T(s) = \widehat{c}_{\lfloor Ts \rfloor}$, $s \in [\kappa, 1]$, and note that $S_T/T = \inf\{s \in [\kappa, 1] : D_T(s) - \widehat{c}_T(s) < 0\}$. By a.s. consistency of the bootstrap as shown above, the process $\widehat{c}_T(s)$ converges weakly to constant, $\widehat{c}_T(s) \overset{P^*}{\to} c$, as $T \to \infty$, P-a.s. Hence, by Slutzky's theorem, $D_T - \widehat{c}_T \overset{w,P^*}{\to} \mathcal{D} - c \in C[\kappa, 1]$, as $T \to \infty$, P-a.s., yielding $S_T/T \overset{w,P^*}{\to} \inf\{s \in [\kappa, 1] : \mathcal{D}(s) - c < 0\}$, if $T \to \infty$, a.s. $\quad\square$

References

Billingsley, P. (1968). *Convergence of probability measures*. New York: Wiley and Sons.

Chan, N. H., and Wei, C. Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics*, *16*, 367-401.

Dickey, D. A., and Fuller, W. A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427-431.

Ferger, D. (1995). Nonparametric tests for nonstandard change-point problems. *Annals of Statistics*, *23*, 1848-1861.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*. Wiley.

Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 121-130.

Hušková, M. (1999). Gradual change versus abrupt change. *Journal of Statistical Planning and Inference*, *76*, 109-125.

Hušková, M., and Slabý, A. (2001). Permutation tests for multiple changes. *Kybernetika*, *37*, 605-622.

Jacod, J., and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*. Springer.

Park, J. Y. (2003). Bootstrap unit root tests. *Econometrica*, *71*, 1845-1895.

Pawlak, M., Rafajlowicz, E., and Steland, A. (2004). Detecting jumps in time series - Nonparametric setting. *Journal of Nonparametric Statistics*, *16*, 329-347.

Rao, M. M. (1978). Asymptotic distribution of an estimator of the boundary parameter of an unstable process. *Annals of Statistics*, *6*, 185-190.

Steland, A. (2004). Sequential control of time series by functionals of kernel-weighted empirical processes under local alternatives. *Metrika*, *60*, 229-249.

Steland, A. (2005a). On detection of unit roots generalising the classic Dickey-Fuller approach. In *Proceedings of the 5th st. petersburg workshop on simulations* (p. 653-658).

Steland, A. (2005b). Optimal sequential kernel smoothers under local nonparametric alternatives for dependent processes. *Journal of Statistical Planning and Inference*, *132*, 131-147.

Steland, A. (2005c). Random walks with drift - A sequential view. *Journal of Time Series Analysis*, *26*, 917-942.

White, J. S. (1958). The limiting distribution of the serial coefficient in the explosive case. *Annals of Mathematical Statistics*, *29*, 1188-1197.

Author's address:

Ansgar Steland
Faculty of Mathematics
Ruhr/University Bochum
Universitätstrasse 150
D-44780 Bochum
Germany

E-mail: ansgar.steland@ruhr-uni-bochum.de