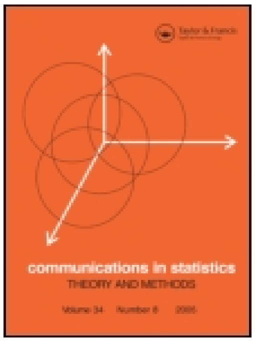

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

A METHOD FOR ESTIMATING PARAMETER IN NONNEGATIVE MA(1) MODELS

Jiří Anděl & Jitka Zichová

COMMUNICATIONS IN STATISTICS
Theory and Methods
Vol. 31, No. 11, pp. 2101–2111, 2002

REGRESSION AND TIME SERIES

A METHOD FOR ESTIMATING PARAMETER IN NONNEGATIVE MA(1) MODELS

Jiří Anděl* and Jitka Zichová†

Department of Statistics, Charles University,
Sokolovská 83, 186 75 Prague 8, Czech Republic

ABSTRACT

A method for estimating parameter in nonnegative MA(1) models is proposed and investigated in the paper. The method also gives nontrivial confidence sets on confidence level 1. Small sample properties of new estimator are demonstrated in a simulation study.

Key Words: Nonnegative time series; Moving-average models; Estimating parameters

1. INTRODUCTION

The aim of this paper is to derive a new method of estimating the parameter in nonnegative MA(1) models. In the statistical analysis of time

*Corresponding author. E-mail: andel@karlin.mff.cuni.cz
†E-mail: zichova@karlin.mff.cuni.cz

DOI: 10.1081/STA-120015019 0361-0926 (Print); 1532-415X (Online)

series it is sometimes known in advance that the values of the investigated process must be nonnegative. Therefore some interest has been focused on estimation procedures in nonnegative time series. Bell and Smith[1] proposed a simple strongly consistent estimate of the parameter of an AR(1) model which was applied in water quality analysis. Their method was generalized to AR(2) models by Anděl.[2] Later on, some attempts were made to find estimates of parameters in nonnegative MA models based on the idea of Bell and Smith (Anděl[3]) but the bias of these estimates was extremely large.

In this paper we propose a new estimate with reduced bias for the parameter of a nonnegative MA(1) model. In MA(1) series of length $n = 100$ with exponentially distributed white noise and with parameter $b = 0.5$ simulations show that the bias of the method presented in Anděl[3] is 0.170 whereas the bias of our new estimate is 0.012. A nontrivial confidence set on confidence level 1 for the parameter of the MA(1) model is also available. The finite sample behavior of the estimate was investigated by means of a simulation study which confirmed its satisfactory convergence properties. A proposition concerning the convergence of some convenient functions of data is proved and it is demonstrated how these functions can be used in the construction of the estimate and the confidence set. Some results of simulations are also presented here.

2. MA(1) MODEL

Let $\{e_t\}$ be positive i.i.d. random variables with a distribution function F. Assume that

$$F(d) - F(c) > 0 \quad \text{for all} \quad 0 < c < d < \infty \tag{2.1}$$

Define MA(1) process $\{X_t\}$ by

$$X_t = e_t + be_{t-1} \tag{2.2}$$

where $b \in (0, 1)$. Then $\{X_t\}$ is a positive invertible process. The problem is to estimate parameter b when variables $X_1, \dots, X_n$ are given. Since

$$\frac{X_t + X_{t-2}}{X_{t-1}} = b + \frac{e_t + (1 - b^2)e_{t-2} + be_{t-3}}{e_{t-1} + be_{t-2}},$$

a method for estimating b proposed by Anděl[3] was based on

$$b_n^* = \min_{3 \le t \le n} \frac{X_t + X_{t-2}}{X_{t-1}}.$$

It was proved that under Eq. (2.1) the estimate b_n^* is strongly consistent but it has a positive bias since $b_n^* > b$. The bias is very serious and tends to zero very slowly as $n \to \infty$.

Let $c \in [0, 1]$. Define

$$Z_t^{(m)}(c) = \sum_{k=0}^{2m} (-c)^k X_{t-k}, \qquad m \geq 1,$$

$$W_n^{(m)}(c) = \min_{2m+1 \leq t \leq n} Z_t^{(m)}(c).$$

It is clear that $W_n^{(m)}(0) > 0$. Since

$$Z_t^{(m)}(c) = e_t + (b - c) \sum_{k=0}^{2m-1} (-c)^k e_{t-1-k} + b\, c^{2m} e_{t-1-2m}$$

we have $W_n^{(m)}(b) > 0$. We shall study asymptotic properties of $W_n^{(m)}(c)$ and of related functions. It will be shown that these functions enable to construct a strongly consistent estimator of the parameter b. For convenience, theoretical results are summarized in the following propositions. Their proofs can be found in the Appendix.

Proposition 1. *Under condition Eq.* (2.1), *we have*

$$W_n^{(m)}(c) \to 0 \quad \textit{for } c = 0,\ c = b,$$

$$W_n^{(m)}(c) \to -\infty \quad \textit{for } c \in (0, b) \cup (b, 1]$$

a.s. as $n \to \infty$.

Proposition 2. *Let* $\lfloor x \rfloor$ *be the integer part of the number* x. *Define*

$$w_n(c) = \min_{1 \leq m \leq \lfloor (n-1)/2 \rfloor} W_n^{(m)}(c), \quad 0 \leq c \leq 1.$$

Then we have $w_n(0) > 0$, $w_n(b) > 0$, *and*

$$w_n(c) \to 0 \quad \textit{for } c = 0,\ c = b, \tag{2.3}$$

$$w_n(c) \to -\infty \quad \textit{for } c \in (0, b) \cup (b, 1] \tag{2.4}$$

a.s. as $n \to \infty$. *Further,* $\{w_n(c)\}_{n=1}^{\infty}$ *is non-increasing sequence for every* $c \in [0, 1]$.

Let ε be a fixed arbitrary small positive number such that $b + \varepsilon < 1$. Define $J = [b + \varepsilon, 1]$, $M_n = \sup_{c \in J} w_n(c)$.

Proposition 3. *We have* $\mathsf{P}\{M_n \leq -1\} \to 1$ *as* $n \to \infty$.

Since $w_n(b) > 0$ and $\mathsf{P}(M_n \leq -1) \to 1$, we propose

$$\hat{b}_n = \sup_{0 \leq c \leq 1} \{c : w_n(c) \geq 0\}$$

as an estimator for b. It is clear that $\hat{b}_n > b$ and thus $\hat{b}_n$ has a positive bias. The behavior of $\hat{b}_n$ was investigated in a simulation study. The results show that the bias is not serious. However, in some short realizations with parameter b near 1 it happened that $w_n(c) > 0$ for all $c \in [b, 1]$. In this case $\hat{b}_n = 1$ and we can say that the method fails. Some details are introduced in Sec. 3.

It would be desirable to derive formulas for $\mathsf{E}\hat{b}_n$, var $\hat{b}_n$, etc., but this is very hard task and the authors were not able to do so. However, it is possible to prove that $\hat{b}_n$ is a strongly consistent estimator for b.

Proposition 4. *We have* $\hat{b}_n \to b$ *a.s. as* $n \to \infty$.

Proposition 2 gives that $w_n(b) > 0$. If we get for some $c \in [0, 1]$ that $w_n(c) \leq 0$ then we are sure that $c \neq b$. Then it is sure that the set $Q = \{c : w_n(c) > 0\}$ covers the unknown value of b. Since $\mathsf{P}\{b \in Q\} = 1$, we can say that Q is a confidence set for b on level 1. Typically, Q is union of two disjoint intervals (see Fig. 6). Nevertheless, the length of Q may be not very large and so Q can give nontrivial information about b. For example, in MA(1) series of length $n = 100$ with exponentially distributed white noise and with parameter $b = 0.5$ a simulation study gives that the average length of Q is 0.31.

If the condition (2.1) is not satisfied, the method may not work. We present only a simple example. Let $a > 0$ be a constant. If $e_t = a$ for all t then $X_t = a(1 + b)$ and

$$Z_t^{(m)}(c) = a(1 + b)\frac{1 + c^{2m+1}}{1 + c}.$$

Since $Z_t^{(m)}(c) > 0$ for all $m \geq 1$ and all $c \in [0, 1]$, we have also $w_n(c) > 0$ for all n and all $c \in [0, 1]$.

3. SIMULATIONS

A simulation of model (2.2) with $b = 0.5$, $n = 100$, and with exponentially distributed white noise $e_t \sim \mathsf{Ex}(1)$ was performed. Functions $W_{100}^{(m)}(c)$ obtained from the same realization for $m = 1$ and $m = 2$ are plotted in Figs. 1 and 2, respectively. All the functions $W_{100}^{(m)}(c)$ for $m = 1, \ldots, 49$ are plotted in Fig. 3. Since we are interested in the values of the functions $W_n^{(m)}(c)$

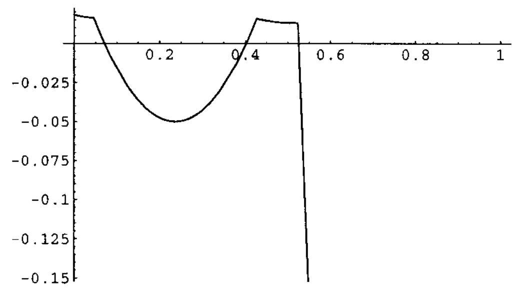

Figure 1. Graph of $W_{100}^{(1)}(c)$.

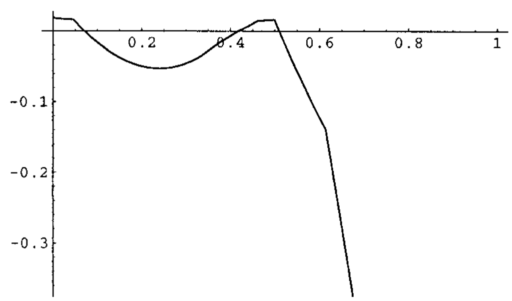

Figure 2. Graph of $W_{100}^{(2)}(c)$.

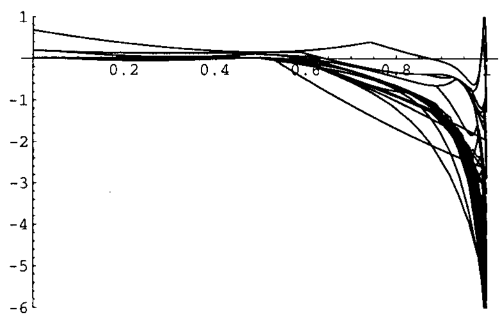

Figure 3. Graph of functions $W_{100}^{(m)}(c)$.

near to their roots, we can see in Fig. 4 behavior of these functions in the vicinity of the real axis. The largest root of the equation $W_{100}^{(m)}(c) = 0$ for $m = 1, \ldots, 7$ is given in Table 1. The function $w_{100}(c)$ is plotted in Figs. 5 and 6 where we can also see the confidence set $\{c : w_n(c) > 0\}$ for the

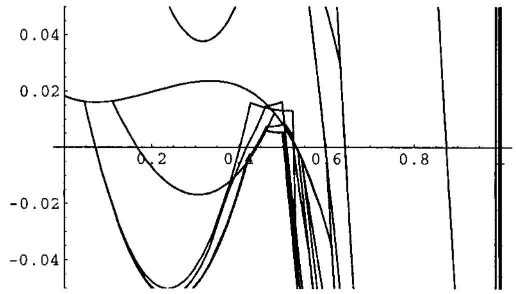

Figure 4. Graph of $W_{100}^{(m)}(c)$, detail.

Table 1. The Largest Root of $W_{100}^{(m)}\ (c) = 0$

m	1	2	3	4	5	6	7
c	0.525	0.510	0.525	0.506	0.503	0.503	0.503

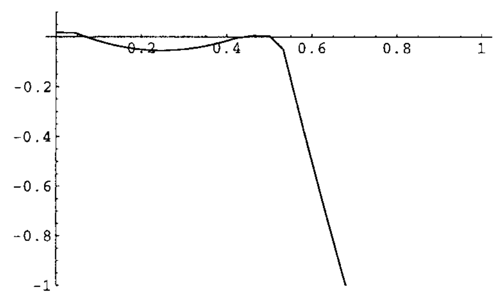

Figure 5. Graph of $W_{100}(c)$.

parameter b. The largest root of the equation $w_n(c) = 0$ is 0.503 which is our estimate for b.

We investigated some sample properties of the estimate $\hat{b}_n$ by calculating averages and standard deviations from 100 realizations of $\{X_t\}$ in the case of exponential white noise $e_t \sim \mathsf{Ex}(1)$. The length n of the series $\{X_t\}$ was chosen $n = 20$, 50 and 100. The parameter of the model (2.2) was chosen $b = 0.1$, 0.5 and 0.9. Our method failed in some realizations with $b = 0.9$ when the value of $\hat{b}_n$ was 1. Averages and sample standard deviations are introduced in Table 2 separately for the case when we used all realizations including those giving $\hat{b}_n = 1$ and separately for the case when the results

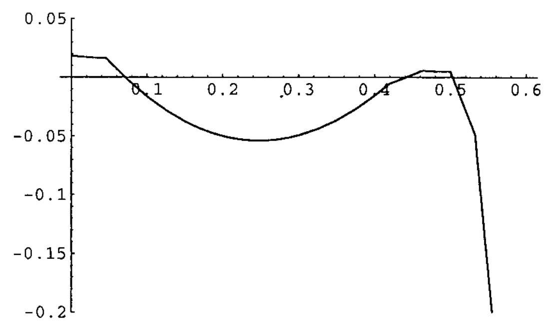

Figure 6. Graph of $W_{100}^{(c)}$, detail.

Table 2. Results of a Simulation with Exponential White Noise Averages, Sample Standard Deviations (in Parentheses), and Number of Failed Cases (in Brackets) Calculated from 100 Simulations

			$b = 0.9$	
n	$b = 0.1$	$b = 0.5$	All Cases	Without Failed Cases
20	0.182	0.597	0.976	0.946
	(0.069)	(0.088)	(0.001)	(0.001)
	[0]	[0]	[56]	
50	0.123	0.527	0.929	0.928
	(0.023)	(0.023)	(0.023)	(0.021)
	[0]	[0]	[2]	
100	0.111	0.512	0.912	0.911
	(0.012)	(0.013)	(0.013)	(0.009)
	[0]	[0]	[1]	

with $\hat{b}_n = 1$ were excluded. The results show that the biases of the estimates are quite large when n is small.

APPENDIX

Proof of Proposition 1. If $c = 0$ or $c = b$ it suffices to show that for every $\varepsilon > 0$

$$\sum_{n=2m+1}^{\infty} \mathsf{P}\left\{ W_n^{(m)}(c) > \varepsilon \right\} < \infty.$$

Then, according to Borel-Cantelli lemma [see Rao,[4] Chap. 2, Appendix 2B], with probability 1 the events $\{W_n^{(m)}(c) > \varepsilon\}$ occur only for a finite number of sample sizes n from which we immediately get $W_n^{(m)}(c) \to 0$ a.s. as $n \to \infty$.

Let $c = 0$. In this case $Z_t^{(m)}(0) = X_t$. Assumption (2.1) implies that

$$\mathsf{P}(X_t \le \varepsilon) \ge \mathsf{P}\Big(e_t \le \frac{\varepsilon}{2},\ be_{t-1} \le \frac{\varepsilon}{2}\Big) > 0$$

for arbitrary $\varepsilon > 0$ and consequently

$$\pi_0 = \mathsf{P}(X_t > \varepsilon) < 1.$$

We have

$$\begin{aligned}\mathsf{P}\Big\{W_n^{(m)}(0) > \varepsilon\Big\} &= \mathsf{P}\{X_t > \varepsilon, t = 2m+1, \dots, n\} \\ &\le \mathsf{P}\Big\{X_t > \varepsilon, t = 2m+1, 2m+3, \dots, 2\Big\lfloor \frac{n-1}{2} \Big\rfloor + 1\Big\} \\ &= \pi_0^{2\lfloor (n-1)/2 \rfloor + 1 - 2m}\end{aligned}$$

and thus

$$\sum_{n=2m+1}^{\infty} \mathsf{P}\Big\{W_n^{(m)}(0) > \varepsilon\Big\} \le \pi_0^{1-2m} \sum_{n=2m+1}^{\infty} \pi_0^{2\lfloor (n-1)/2 \rfloor} < \infty.$$

Let $c = b$. We have $Z_t^{(m)}(b) = e_t + b^{2m+1} e_{t-2m-1}$. In view of Eq. (2.1),

$$0 < F\Big(\frac{\varepsilon}{2}\Big) F\bigg(\frac{\varepsilon}{2\, b^{2m+1}}\bigg) = \mathsf{P}\Big(e_t \le \frac{\varepsilon}{2}, b^{2m+1} e_{t-2m-1} \le \frac{\varepsilon}{2}\Big) \le \mathsf{P}(Z_t^{(m)}(b) \le \varepsilon).$$

Therefore we get

$$\pi_b = \mathsf{P}\Big(Z_t^{(m)}(b) > \varepsilon\Big) < 1.$$

Define

$$u = (2m+2)\left\lfloor \frac{n-2m-1}{2m+2} \right\rfloor + 2m+1, \qquad R = \left\lfloor \frac{n+1}{2m+2} \right\rfloor. \tag{A.1}$$

This leads to

$$\begin{aligned}\mathsf{P}\Big\{W_n^{(m)}(b) > \varepsilon\Big\} &= \mathsf{P}(Z_t^{(m)}(b) > \varepsilon,\ t = 2m+1, \dots, n) \\ &\le \mathsf{P}\big(e_t + b^{2m+1} e_{t-2m-1} > \varepsilon,\ t = 2m+1, 4m+3, \dots, u\big) \\ &= \mathsf{P}\big(e_{2m+1} + b^{2m+1} e_0 > \varepsilon\big)^R = \pi_b^R < 1.\end{aligned}$$

Further,

$$\sum_{n=2m+1}^{\infty} \mathsf{P}\Big\{W_n^{(m)}(b) > \varepsilon\Big\} \leq \sum_{n=2m+1}^{\infty} \pi_b^R = (2m+2)\sum_{k=1}^{\infty}\pi_b^k = (2m+2)\frac{\pi_b}{1-\pi_b} < \infty.$$

If $0 < c < b$ or $b < c \leq 1$ we need to prove that for every $\varepsilon > 0$

$$\sum_{n=2m+1}^{\infty} \mathsf{P}\Big\{W_n^{(m)}(c) > -\varepsilon\Big\} < \infty$$

and then to apply the Borel-Cantelli lemma. We have

$$Z_t^{(m)}(c) = e_t + (b-c)\sum_{k=0}^{2m-1}(-c)^k e_{t-k-1} + b\,c^{2m} e_{t-2m-1}.$$

Suppose first $0 < c < b$. The coefficients $1, (-c)^2, \ldots, (-c)^{2m-2}$ by the variables $e_{t-1}, e_{t-3}, \ldots, e_{t-2m+1}$ are positive but the coefficients $-c$, $(-c)^3, \ldots, (-c)^{2m-1}$ by e_{t-2}, $e_{t-4}, \ldots, e_{t-2m}$ are negative. Let $\varepsilon > 0$ be arbitrary. Under Assumption (2.1) we get

$$\begin{aligned}
0 &< F\Big(\frac{\varepsilon}{m+2}\Big)F\Big(\frac{\varepsilon}{(b-c)(m+2)}\Big)F\Big(\frac{\varepsilon}{(b-c)c^2(m+2)}\Big) \times \cdots \\
&\quad \times F\Big(\frac{\varepsilon}{(b-c)c^{2m-2}(m+2)}\Big)F\Big(\frac{\varepsilon}{b\,c^{2m}(m+2)}\Big)\Big[1 - F\Big(\frac{2\varepsilon}{(b-c)cm}\Big)\Big] \\
&\quad \times \Big[1 - F\Big(\frac{2\varepsilon}{(b-c)c^3 m}\Big)\Big] \times \cdots \times \Big[1 - F\Big(\frac{2\varepsilon}{(b-c)c^{2m-1}m}\Big)\Big] \\
&= \mathsf{P}\Big(e_t \leq \frac{\varepsilon}{m+2}\Big)\mathsf{P}\Big(e_{t-1} \leq \frac{\varepsilon}{(b-c)(m+2)}\Big) \\
&\quad \times \mathsf{P}\Big(e_{t-3} \leq \frac{\varepsilon}{(b-c)c^2(m+2)}\Big) \times \cdots \\
&\quad \times \mathsf{P}\Big(e_{t-2m+1} \leq \frac{\varepsilon}{(b-c)c^{2m-2}(m+2)}\Big) \\
&\quad \times \mathsf{P}\Big(e_{t-2m-1} \leq \frac{\varepsilon}{bc^{2m}(m+2)}\Big) \times \mathsf{P}\Big(e_{t-2} > \frac{2\varepsilon}{(b-c)cm}\Big) \\
&\quad \times \mathsf{P}\Big(e_{t-4} > \frac{2\varepsilon}{(b-c)c^3 m}\Big) \times \cdots \times \mathsf{P}\Big(e_{t-2m} > \frac{2\varepsilon}{(b-c)c^{2m-1}m}\Big) \\
&\leq \mathsf{P}\Big(Z_t^{(m)}(c) \leq -\varepsilon\Big)
\end{aligned}$$

and

$$\pi = \mathsf{P}\Big(Z_t^{(m)}(c) > -\varepsilon\Big) < 1.$$

Since $Z_t^{(m)}(c)$ and $Z_s^{(m)}(c)$ are independent for $t - s > 2m + 1$ we get

$$\begin{aligned}\mathsf{P}\Big\{W_n^{(m)}(c) > -\varepsilon\Big\} &= \mathsf{P}\Big(Z_t^{(m)}(c) > -\varepsilon,\ t = 2m+1, \dots, n\Big)\\ &\leq \mathsf{P}\Big(Z_t^{(m)}(c) > -\varepsilon,\ t = 2m+1, 4m+3, \dots, u\Big)\\ &= \mathsf{P}\Big(Z_{2m+1}^{(m)}(c) > -\varepsilon\Big)^R = \pi^R\end{aligned}$$

where u and R are given by Eq. (A.1). Summing these probabilities we get

$$\sum_{n=2m+1}^{\infty} \mathsf{P}\Big\{W_n^{(m)}(c) > -\varepsilon\Big\} \leq (2m+2)\frac{\pi}{1-\pi} < \infty.$$

If $b < c \leq 1$ the proof is analogous.

Proof of Proposition 2. The inequalities $w_n(0) > 0$, $w_n(b) > 0$ follow directly from the definition of $w_n(c)$. Formulas (2.3) and (2.4) are easy consequences of Proposition 1. It remains to prove that $\{w_n(c)\}_{n=1}^{\infty}$ is non-increasing. Let $1 \leq n < n^*$. Then $W_{n^*}^{(m)}(c) \leq W_n^{(m)}(c)$ and thus

$$\begin{aligned}w_{n^*}(c) &= \min_{1\leq m\leq\lfloor(n^*-1)/2\rfloor} W_{n^*}^{(m)}(c)\\ &\leq \min_{1\leq m\leq\lfloor(n-1)/2\rfloor} W_{n^*}^{(m)}(c) \leq \min_{1\leq m\leq\lfloor(n-1)/2\rfloor} W_n^{(m)}(c) = w_n(c).\end{aligned}$$

Proof of Proposition 3. We have

$$Z_t^{(1)}(c) = X_t - cX_{t-1} + c^2X_{t-2} = e_t + (b-c)(e_{t-1} - ce_{t-2}) + bc^2e_{t-3}.$$

Consider $c \in J$. Introduce events

$$A_t = \left\{e_t \leq \frac{1}{2},\ e_{t-1} \geq e_{t-2} + \frac{2}{\varepsilon},\ e_{t-3} \leq \frac{1}{2}\right\}.$$

Since $\mathsf{P}(A_t) > 0$ does not depend on t, Borel-Cantelli lemma implies that infinitely many events A_t occur in the sequence (e_1, e_2, e_3, e_4), $(e_5, e_6, e_7, e_8), \dots$ with probability 1. Thereby, infinitely many events A_t occur in the whole sequence $e_1, e_2, \dots$ with probability 1. Whenever A_t occurs, we have

$$\sup_{c\in J} Z_t^{(1)}(c) \leq U_t = e_t - \varepsilon(e_{t-1} - e_{t-2}) + e_{t-3} \leq -1.$$

Thus we obtain

$$\mathsf{P}\left\{\sup_{c\in J} w_n(c) \leq -1\right\} \geq \mathsf{P}\left\{\sup_{c\in J}\min_{3\leq t\leq n} Z_t^{(1)}(c) \leq -1\right\} \geq \mathsf{P}(A_3 \cup \cdots \cup A_n) \to 1.$$

Proof of Proposition 4. It suffices to prove that for each sufficiently small $\eta > 0$ we have

$$\lim_{N\to\infty} \mathsf{P}\left\{\sup_{n\geq N}\left|\hat{b}_n - b\right| > \eta\right\} = 0$$

[see Rao,[4] Chap. 2, Formula (2c.2.3)]. We already know that $\hat{b}_n > b$ and that $\{w_n(c)\}_{n=1}^{\infty}$ is non-increasing. Thus

$$\begin{aligned}\lim_{N\to\infty} \mathsf{P}\left\{\sup_{n\geq N}\left|\hat{b}_n - b\right| > \eta\right\} &= \lim_{N\to\infty} \mathsf{P}\left\{\sup_{n\geq N}(\hat{b}_n - b) > \eta\right\}\\ &= \lim_{N\to\infty} \mathsf{P}\left\{\hat{b}_N - b > \eta\right\}\\ &= \lim_{N\to\infty} \mathsf{P}\left\{\sup_{c\in(b+\eta,\,1]} w_N(c) \geq 0\right\} = 0\end{aligned}$$

since it follows from Proposition 3 that $\mathsf{P}\{\sup_{c\in[b+\eta,\,1]} w_N(c) > -1\} \to 0$.

ACKNOWLEDGMENTS

The authors are grateful to two anonymous referees. Their comments substantially improved the paper. Financial support from grants GAČCR 201/00/0770 and CEZ: J13/98: 113200008 is gratefully acknowledged.

REFERENCES

1. Bell, C.B.; Smith, E.P. Inference for Non-negative Autoregressive Schemes. Commun. Statist.—Theor. Meth. **1086**, *15*(8), 2267–2293.
2. Anděl, J. Non-negative Autoregressive Processes. J. Time Ser. Anal. **1989**, *10*(1), 1–11.
3. Anděl, J. Nonnegative Moving-average Models. In *Asymptotic Statistics*; Proceedings of the Fifth Prague Symposium: Prague, Czech Republic, Sept 4–9, 1993; Mandl, P., Hušková, M., Eds.; Physica Verlag: Heidelberg, 1994, 163–171.
4. Rao, C.R. *Linear Statistical Inference and its Applications*, 2nd Ed.; Wiley: New York, 1973.