ON MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS FOR CENSORED AUTOREGRESSIVE TIME SERIES

I.A. BADZIAHIN[1], YU.S. KHARIN[2]
Belarusian State University
Minsk, Belarus
e-mail: [1]`bodiagin@cosmostv.by`, [2]`kharin@bsu.by`

Abstract

Statistical estimation of model parameters for autoregressive time series under right censoring is considered. Maximum likelihood estimators for model parameters are constructed. Comparison of the maximum likelihood estimators and some "traditional" estimators is made. Numerical results are given.

1 Introduction

Autoregressive model of order p (AR(p)) is widely used to describe stochastic processes in many fields, such as economy, finance, meteorology, medicine [1,2]. The case of "full data" for this model, where all observations are exactly known, is well studied. In practise, however, time series are usually observed under different distortions, and in this case classical estimators of model parameters are usually biased and inconsistent. That is why it is necessary to construct new robust methods for estimation of model parameters [3].

In this paper we consider distortions generated by right censoring, that are often in engineering, economics, business, etc [4]. Censoring means, that exact values of some observations are unknown and it is only known that they belong to certain intervals. In the case of right censoring it is only known that true values of some observations are greater than given cutoff values. Right censoring can appear because of detection limits of measuring devices, high costs of measurement, disorders of equipment, etc [4].

2 Mathematical model

Consider the AR(p) time series model [1]

$$x_t = \sum_{i=1}^{p} \theta_i x_{t-i} + u_t, t \in \mathbf{Z}, \tag{1}$$

where $\theta_1, \ldots, \theta_p$ are unknown coefficients of the autoregression; all roots of the characteristic polynomial $z^p - \sum_{i=1}^{p} \theta_i z^{p-i}$ are inside the unit circle; $\{u_t\}$ are i.i.d. normal random variables, $\mathrm{E}\{u_t\} = 0$, $\mathrm{D}\{u_t\} = \sigma^2 < +\infty$.

Instead of the true values $x_1, \ldots, x_T$ we observe only random events:

$$A_t^* = \{x_t \in A_t\}, \ i \in \{1, \ldots, T\}, \tag{2}$$

where $\{A_t\}$ are some known Borel sets, T is the length of the observation process. In this paper we consider two possible cases: 1) $A_t = \{x_t\}$ is a singleton, then the value of the t-th observation x_t is known; 2) $A_t = [c_t, +\infty)$ is an interval, where c_t is a known cutoff level, then the observation x_t is censored.

The right censored time series (2) can be represented as a sequence of fragments with fully observed data and fragments with fully censored data. Let τ_i be the length of the i-th fragment of censored data and t_i^* be the initial time moment of this fragment, $i \in \{1, \ldots, M\}$, M be the number of the censored fragments.

The main problem considered in this paper is to construct a robust estimators of the model parameters $\{\theta_i\}_{i=1}^{p}$, σ based on random events (2).

To simplify our results we consider only the case of AR(1) model. However, all our results can be generalized for the case of $p > 1$.

3 Log-likelihood function

Introduce the notations:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} -$$

the standard normal probability density function;

$$\mathrm{I}_\nu(y, m, s) = \int\limits_y^{+\infty} (t-m)^\nu \frac{1}{s} \varphi\left(\frac{t-m}{s}\right) dt, \ \nu \in \mathbf{N} \cup \{0\}, \ y, m, s \in \mathbf{R}; \tag{3}$$

$$\nu(l; k, i_1, \ldots, i_k) = \sum_{j=1}^{k} \left(\delta_{l,i_j} + \delta_{l,i_j+1}\right), \ l, k, i_1, \ldots, i_k \in \mathbf{N}; \quad \delta_{i,j} = \begin{cases} 1, & i = j; \\ 0, & i \neq j; \end{cases}$$

$$d(l; A, \theta, \sigma) = \begin{cases} \sigma, & l \in A; \\ \frac{\sigma}{\sqrt{1+\theta^2}}, & l \notin A, \end{cases} \quad l \in \mathbf{N}, \ A \subset \mathbf{N}, \ \theta, \sigma \in \mathbf{R};$$

$$l^*(t_1, t_2 | \theta, \sigma) = \sum_{t=t_1+1}^{t_2} \ln \frac{1}{\sigma} \varphi\left(\frac{x_t - \theta_1 x_{t-1}}{\sigma}\right), \ 0 < t_1 < t_2 \leq T;$$

$$F_1(x_{t^*-1}, x_{t^*+\tau}, \tau | \theta, \sigma) = \frac{1}{\sigma} \varphi\left(\frac{x_{t^*+\tau} - \theta^{\tau+1} x_{t^*-1}}{\sigma\sqrt{\frac{1-\theta^{2\tau+2}}{1-\theta^2}}}\right) (1+\theta^2)^{-\frac{\tau}{2}} \left(\sum_{k=1}^{+\infty} \frac{\theta^k}{k!} \times \right.$$

$$\left. \times \sum_{i_1, \ldots, i_k = 1}^{\tau-1} \prod_{l=1}^{\tau} \mathrm{I}_{\nu(l,k,i_1,\ldots,i_k)} \left(c_{t^*+l-1}, \bar{\mu}_l, d(l; \emptyset, \theta, \sigma)\right) + \prod_{l=1}^{\tau} \mathrm{I}_0 \left(c_{t^*+l-1}, \bar{\mu}_l, d(l; \emptyset, \theta, \sigma)\right) \right), \tag{4}$$

$$\bar{\mu}_l = \theta^l x_{t^*-1} + \theta^{\tau-l} \frac{1-\theta^{2l}}{1-\theta^{2\tau+2}} (x_{t^*+\tau} - \theta^{\tau+1} x_{t^*-1});$$

$$F_2(x_{t^*+\tau},\tau|\theta,\sigma)=\frac{\sqrt{1-\theta^2}}{\sigma}\varphi\left(\frac{x_{t^*+\tau}\sqrt{1-\sigma^2}}{\sigma}\right)(1+\theta^2)^{\frac{1-\tau}{2}}\times$$
$$\times\left(\sum_{k=1}^{+\infty}\frac{\theta^k}{k!}\sum_{i_1,\ldots,i_k=1}^{\tau-1}\prod_{l=1}^{\tau}\mathrm{I}_{\nu(l,k,i_1,\ldots,i_k)}\left(c_{t^*+l-1},\theta^{\tau-l+1}x_{t^*+\tau},d(l;\{1\},\theta,\sigma)\right)+\right.$$
$$\left.+\prod_{l=1}^{\tau}\mathrm{I}_0\left(c_{t^*+l-1},\theta^{\tau-l+1}x_{t^*+\tau},d(l;\{1\},\theta,\sigma)\right)\right);\quad(5)$$

$$F_3(x_{t^*-1},\tau|\theta,\sigma)=(1+\theta^2)^{\frac{1-\tau}{2}}\left(\prod_{l=1}^{\tau-1}\mathrm{I}_0\left(c_{t^*+l-1},\theta^l x_{t^*-1},d(l;\{\tau\},\theta,\sigma)\right)+\right.$$
$$\left.+\sum_{k=1}^{+\infty}\frac{\theta^k}{k!}\sum_{i_1,\ldots,i_k=1}^{\tau-1}\prod_{l=1}^{\tau}\mathrm{I}_{\nu(l,k,i_1,\ldots,i_k)}\left(c_{t^*+l-1},\theta^l x_{t^*-1},d(l;\{\tau\},\theta,\sigma)\right)\right).\quad(6)$$

Note that (3) are calculated in [5]. Also note that functions (4)–(6) cannot be explicitly calculated in practice, because they are represented as an infinite sum of terms. In practice we will use a finite number (k_{max}) of terms to approximate these functions.

Theorem. *If the autoregression order $p=1$ and the number of the censored fragments $M>1$, then the log-likelihood function for the right censored time series (1), (2) is*

$$l(\theta_1,\sigma)=\delta_{t_1^*,1}\ln F_2(x_{t_1^*+\tau_1},\tau_1|\theta_1,\sigma)+(1-\delta_{t_1^*,1})\left(\ln\varphi\left(\frac{x_1\sqrt{1-\theta_1^2}}{\sigma}\right)+\right.$$
$$\left.+\ln\frac{\sqrt{1-\theta_1^2}}{\sigma}+l^*(1,t_1^*-1|\theta_1,\sigma)+\ln F_1(x_{t_1^*-1},x_{t_1^*+\tau_1},\tau_1|\theta_1,\sigma)\right)+$$
$$+\sum_{i=2}^{M-1}\left(l^*(t_{i-1}^*+\tau_{i-1},t_i^*-1|\theta_1,\sigma)+\ln F_1(x_{t_i^*-1},x_{t_i^*+\tau_i},\tau_i)|\theta_1,\sigma\right)+$$
$$+l^*(t_{M-1}^*+\tau_{M-1},t_M^*-1|\theta_1,\sigma)+\delta_{t_M^*+\tau_M,T}\ln F_3(x_{t_M^*+\tau_M},\tau_M|\theta_1,\sigma)+$$
$$+(1-\delta_{t_M^*+\tau_M,T})\left(\ln F_1(x_{t_M^*-1},x_{t_M^*+\tau_M},\tau_M|\theta_1,\sigma)+l^*(t_M^*+\tau_M,T|\theta_1,\sigma)\right).\quad(7)$$

Maximum likelihood estimators (MLE) of (θ_1,σ) are the solutions of the extremum problem:

$$(\hat\theta_1,\hat\sigma)=\arg\max_{\theta_1,\sigma}l(\theta_1,\sigma).\quad(8)$$

4 Numerical results

To estimate the model parameters the following approach is widely used [4]: censored observations are replaced by the values of the corresponding cutoff levels c_t and then classical estimators, for example the least squares estimators (LSE), are calculated.

Computer experiments are performed to compare MLE and LSE described above for right censored times series. The experiment consist in the following: 1) autoregressive time series of the length $T = 300$ is generated with parameters $p = 1$, $\theta_1 = -0.3$, $\sigma = 1$; 2) if $x_t \geq c_t$, $c_t \equiv 0.01$, then the random event $A_t^* = \{x_t \geq c_t\}$ is observed; 3) The MLE and LSE of the parameter θ_1 are calculated in assumption that the parameter σ is known. The MLE of the θ_1 can be approximately found by the tabulation of the likelihood function (7) with certain accuracy. To calculate (7) $k_{max} = 2$ or $k_{max} = 6$ summands in (4) (6) are used.

The results of the experiments are given in the Table 1. As we can see the least square method gives unacceptable results (the average error is grater than 0.4). The maximum likelihood method gives much better results. Also the more summands are used to approximate functions (4) (6), the better results are gained.

			kmax=2		kmax=6	
№	LSE	error	MLE	error	MLE	error
1	0.17639	0.47639	-0.34586	0.045856	-0.30225	0.002252
2	0.190799	0.490799	-0.33396	0.033964	-0.26261	0.037387
3	0.155476	0.455476	-0.38946	0.089459	-0.37559	0.075586
4	0.149382	0.449382	-0.42712	0.127117	-0.40333	0.103333
5	0.195661	0.495661	-0.38351	0.083514	-0.3518	0.051802
6	0.151674	0.451674	-0.41721	0.117207	-0.3855	0.085495
7	0.178645	0.478645	-0.36568	0.065676	-0.31018	0.01018
8	0.154852	0.454852	-0.43306	0.133063	-0.40333	0.103333
9	0.164298	0.464298	-0.34169	0.041692	-0.29036	0.00964
10	0.166816	0.466816	-0.29432	0.005676	-0.24874	0.051261
mean	**0.168399**	**0.468399**	**-0.37625**	**0.077508**	**-0.33337**	**0.053027**

Table 1: Comparison of MLE and LSE

References

[1] G.E.P. Box, G.M. Jenkins. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day, San-Francisco.

[2] T.W. Anderson. (1971). *The Statistical Analysis of Time Series*. Wiley, New-York.

[3] Yu.S. Kharin. (2008). *Optimality and robustness in statistical forecasting*. BSU, Minsk (in Russian).

[4] J.W. Park, M.G. Genton, S.K. Ghosh. (2007). Censored time series analysis with autoregressive moving average models. *The Canadian Journal of Statistics*. Vol. **35**, pp. 151-168.

[5] I.M. Ryzhik, I.S. Gradshteyn (2007). *Table of Integrals, Series, and Products, seventh edition*. Academic Press.