STATISTICAL ANALYSIS OF COUNT CONDITIONALLY NONLINEAR AUTOREGRESSIVE TIME SERIES BY FREQUENCIES-BASED ESTIMATORS

Yu.S. Kharin, M.I. Kislach
Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University
Minsk, BELARUS
e-mail: kharin@bsu.by, kislachm@gmail.com

Abstract

Models of count time series with denumerable states space with conditional probability distributios generated by Bernoulli trial scheme (Poisson (model M_1), Geometric (model M_2), Negative binomial (model M_3), Borel-Tanner (model M_4)) conditionally nonlinear autoregressive time series are developed. Consistent estimators for parameters of proposed models based on Markov properties are constructed. Algorithms for statistical forecasting of count time series are developed. Results of computer experiments are given.
Keywords: data science, count data, nonlinear autoregression, frequencies-based estimator

1 Mathematical models of count time series with denumerable state space and their probability properties

Count time series are widely used in different applications: genetics, economics, information protection [1-4]. The case of finite states space is considered in [3]. In this paper we develop our results from [3] to the case of denumerable states space.

Let in probability space $(\Omega, \mathcal{F}, P)$ a time series $x_t \in A = \{0, 1, \dots\}$ be defined. We call it the conditionally nonlinear autoregressive time series if the conditional probability distribution of the random value x_t under its prehistory $\{x_{t-1}, x_{t-2}, \dots\}$ depends only on s-prehistory $X_{t-s}^{t-1} = (x_{t-1}, x_{t-2}, \dots, x_{t-s})' \in A^s$ for some depth $s \in N$:

$$P\{x_t = j | x_{t-1} = j_{t-1}, x_{t-2} = j_{t-2}, \dots\} = P\{x_t = j | X_{t-s}^{t-1} = J_1^s\} = \mathbf{Q}(j; J_1^s), \quad (1)$$

where $j \in A, J_1^s \in A^s$, $\mathbf{Q}(\cdot; J_1^s)$ is some discrete probability distribution on A for each $J_1^s \in A^s$. We will assume that this function is parameterized in the following way:

$$\mathbf{Q}(j; J_1^s) ::= \mathbf{q}(j; \theta(J_1^s)), j \in A, J_1^s \in A^s. \quad (2)$$

Here $q(\cdot; \theta)$ is some fixed (standard) discrete probability distribution with some parameter $\theta \in R^1$, and $\theta = \theta(J_1^s)$ is some function describing dependence of this parameter on the s - prehistory in the form:

$$\theta = \theta(J_1^s) ::= F\left(\sum_{i=1}^{m} a_i \psi_i(J_1^s)\right), \tag{3}$$

where $F(\cdot) : R^1 \to R^1$ is some khown function; $\Psi(u) = (\psi_1(u), \ldots, \psi_m(u))' : A^s \to R^m$ are some base functions, $\psi_i(\cdot) : A^s \to R^1$; $a = (a_i) \in R^m$ is some unknown column-vector of parameters.

Consider four special cases of proposed model (1)–(3) for count time series [2]:

$$\mathbf{q}(j;\theta) = \begin{cases} \theta^j e^{-\theta}/j! & for \;\; model \;\; M_1 \\ \theta(1-\theta)^j, j \in A & for \;\; model \;\; M_2, \\ C_{r+j-1}^r \theta^r (1-\theta)^j, j \geq r & for \;\; model \;\; M_3, \\ e^{-j\theta} r \theta^{j-r} j^{j-r-1}/(j-r)!, j \geq r & for \;\; model \;\; M_4, \end{cases} \tag{4}$$

where $r \in N$ is some fixed value.

We use in (3) the function $F(z) = e^z$ for the model M_1, and the logistic cumulative distribution function for the models $M_2 - M_4$:

$$F(z) = e^z/(1+e^z), z \in R^1. \tag{5}$$

Lemma 1. *Count time series determined by model (1)–(3) is the denumerable homogeneous Markov chain of order s with the states space A and the one-step transition probabilities:*

$$P\{x_t = j | X_{t-s}^{t-1} = J_1^s\} = \mathbf{q}(j; \theta(J_1^s)), j \in A, J_1^s \in A^s, \tag{6}$$

where $\mathbf{q}(\cdot)$ is determined by (4) for the considered special cases.

2 Statistical estimation of model parameters

Give two auxiliary results.

Lemma 2. *For model (1)–(3) the conditional mean is ($J_1^s \in A^s$):*

$$\mu(J_1^s) ::= E\{x_t | X_{t-s}^{t-1} = J_1^s\} = M(\theta(J_1^s)) = \begin{cases} \theta(J_1^s) & for \;\; M_1, \\ (1-\theta(J_1^s))/\theta(J_1^s) & for \;\; M_2, \\ r(1-\theta(J_1^s))/\theta(J_1^s) & for \;\; M_3, \\ r/(1-\theta(J_1^s)) & for \;\; M_4. \end{cases} \tag{7}$$

Lemma 3. *For model (1)–(3) the following equations hold:*

$$a'\Psi(J_1^s) = F^{-1}(M^{-1}(\mu(J_1^s))) = \begin{cases} \ln \mu(J_1^s) & for \;\; model \;\; M_1, \\ -\ln \mu(J_1^s) & for \;\; model \;\; M_2, \\ -\ln(\mu(J_1^s)/r) & for \;\; model \;\; M_3, \\ \ln((\mu(J_1^s) - r)/r) & for \;\; model \;\; M_4, \end{cases} \tag{8}$$

where $\mu(\cdot)$ is determined by (7).

To construct statistical estimator for the unknown vector of parameters $a=(a_i)\in R^m$ in the models (1)–(5) we will use the approach based on the frequencies-based estimators proposed in [3].

Introduce the notation: $I\{C\}$ is the indicator function of the event C;

$$\nu(J_1^s)=\sum_{t=s+1}^{T} I\left(X_{t-s}^{t-s}=J_1^s\right);$$

$$B\left(X_1^T\right)=\{J_1^s{\in}A^s\colon \nu(J_1^s)>0\}=\left\{J_1^{s,(1)},\ldots,J_1^{s,(K)}\right\},$$

where $K\le T-s$ and $\nu\left(J_1^{s,(i)}\right)\ge\nu\left(J_1^{s,(j)}\right)$ for all $i<j,(i,j=1,\ldots,K);K_0=K_0\left(m,T,s\right):N^3\to N,m\le K_0\left(m,T,s\right)\le K$, is a function nondecreasing w.r.t. m; $B_0=\left\{J_1^{s,(1)},J_1^{s,(2)},\ldots,J_1^{s,(K_0)}\right\}\subset B\left(X_1^T\right)$ with the cardinality $|B_0|=K_0$.

Theorem 1. *For model (1)–(3) under the observed realization* $X_1^T=$ $=(x_1,x_2,\ldots,x_T)'\in A^T$ *the statistical estimator*

$$\hat{\mu}(J_1^s)=\sum_{t=s+1}^{T} x_t I(X_{t-s}^{t-s}=J_1^s)/\nu(J_1^s)$$

is a consistent estimator of $\mu(J_1^s)$ *for* $T\to+\infty$.

Theorem 2. *For model (1)–(5) under the observed realization* $X_1^T=$ $=(x_1,x_2,\ldots,x_T)'\in A^T$ *the statistical estimator*

$$\hat{a}=H^{-1}C, \tag{9}$$

is a consistent estimator of vector parameter a *, where* $H=\sum_{J_1^s\in B_0}\Psi(J_1^s)\Psi^T(J_1^s)$,

$$C=\sum_{J_1^s\in B_0}F^{-1}(M^{-1}(\mu(J_1^s)))\Psi(J_1^s)=\sum_{J_1^s\in B_0}\begin{cases}\ln(\hat{\mu}(J_1^s))\Psi(J_1^s) & for\ \ M_1,\\ -\ln(\hat{\mu}(J_1^s))\Psi(J_1^s) & for\ \ M_2,\\ -\ln(\hat{\mu}(J_1^s)/r)\Psi(J_1^s) & for\ \ M_3,\\ \ln\left(\left(\hat{\mu}(J_1^s)-r\right)/r\right)\Psi(J_1^s) & for\ \ M_4,\end{cases}$$

and $\{\hat{\mu}(J_1^s)\}$ *are from Theorem 1.*

3 Statistical forecasting of count time series

Theorem 3. *For the model (1)–(3) under the observed realization* $X_1^T=$ $=(x_1,x_2,\ldots,x_T)'\in A^T$ *and* $|H|\neq0$ *the optimal forecasting statistic for the future state* $x_{T+1}\in A$ *that minimizes the mean square error of forecasting [2] is:*

$$\hat{x}_{T+1}=\begin{cases}\left\lfloor\hat{\theta}\right\rfloor & for\ \ model\ \ M_1,\\ \left\lfloor(1-\hat{\theta})/\hat{\theta}\right\rfloor & for\ \ model\ \ M_2,\\ \left\lfloor r(1-\hat{\theta})/\hat{\theta}\right\rfloor & for\ \ model\ \ M_3,\\ \left\lfloor r/(1-\hat{\theta})\right\rfloor & for\ \ model\ \ M_4,\end{cases} \tag{10}$$

where $\hat{\theta} = F\left(\hat{a}'\Psi(X_{T-s+1}^{T})\right)$, $\lfloor y \rfloor$ *means the floor function of* y.

4 Results of computer experiments

Experiments were performed in R computer language. Figure 1 for model M_1 illustrates dependence of the Monte-Carlo estimate of the mean square error (MSE) for estimator (9) from log_2T with $M = 100$ replications.

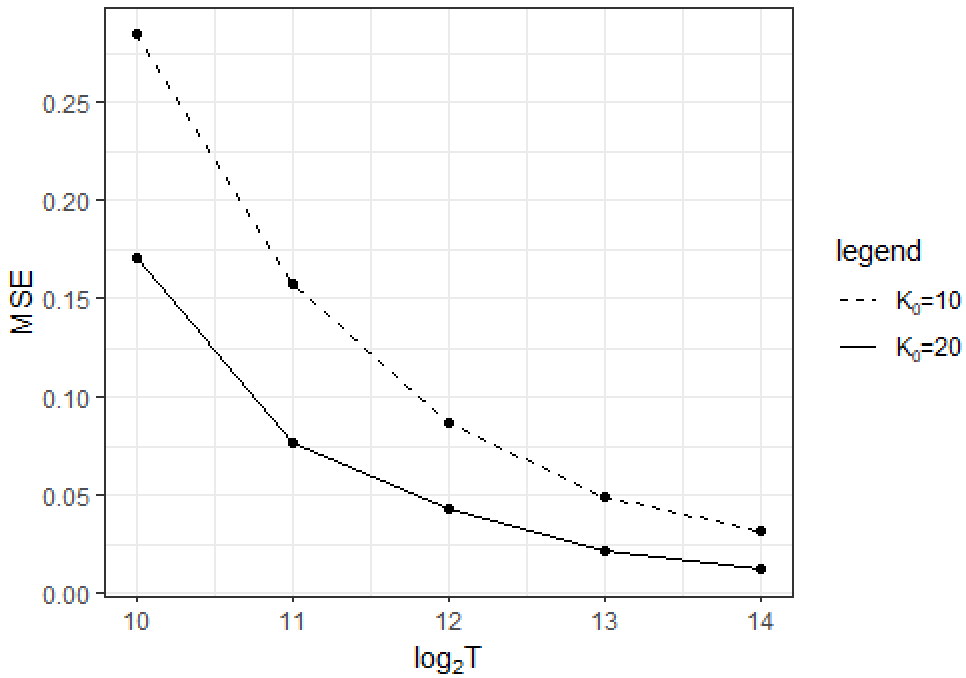

Figure 1: Dependence of the mean square error from log_2T

References

[1] Frank A.H., Melvin A.B. (1960). The Borel-Tanner distribution. *Biometrika* Vol. **47 (1-2)**, pp. 143-150.

[2] Kharin Y.S. (2013). *Robustness in Statistical Forecasting*. Springer, New York.

[3] Kharin Y.S., Voloshko V.A., Medved E.A. (2018). Statistical estimation of parameters for binary conditionally nonlinear autoregressive time series. *Math Met Stat*. Vol. **17 (2)**, pp. 103-118.

[4] Hastings N.A, Peacock J.B. (1979). *Statistical distributions*. John Wiley & Sons Inc, New Jersey.