Statistics and Probability Letters 112 (2016) 79–91

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Epidemic change tests for the mean of innovations of an $AR(1)$ process

J. Markevičiūtė

Faculty of Mathematics and Informatics, Vilnius University, Naugarduko str. 24, LT-03225 Vilnius, Lithuania

ARTICLE INFO

Article history:
Received 22 July 2015
Received in revised form 2 February 2016
Accepted 3 February 2016
Available online 10 February 2016

Keywords:
Autoregressive process
Epidemic change
Uniform increments statistics
Dyadic increments statistics

ABSTRACT

We study a first order autoregressive process with the autoregressive coefficient $\phi = 1$ or $|\phi| < 1$. Our aim is to test whether there is an epidemic type change in the mean of innovations with the statistics based on the observations. We use two equivalent Hölderian test statistics: uniform and dyadic increments statistics. We find the limit under null hypothesis of no change, then we establish consistency conditions under alternative.

1. Introduction

The autoregressive processes are one of the most popular processes in modeling economic time series. Nevertheless time series might have changes in the distribution over time. So an important question in time series analysis is a change point detection in the innovations of an autoregressive process. Such problems were investigated in Neumeyer and Selk (2013), who studied nonparametric autoregression model under conditional heteroscedasticity with the aim to test whether the innovation distribution changes in time, while Hlávka et al. (2012) developed a procedure for monitoring changes in the error distribution of autoregressive time series, controlling the overall size of the sequential test. Also, the errors of regression models were studied, for example, Neumeyer and van Keilegom (2009) investigated change point tests for detecting change distribution of errors in non-parametric regression.

In this paper we investigate two test statistics for detecting the so called epidemic change in the innovations. Testing an epidemic type change in the mean of innovations means to test null hypothesis that $\mathbb{E}\varepsilon_i = 0, i = 1, \ldots, n$, against alternative $\mathbb{E}\varepsilon_1 = \cdots = \mathbb{E}\varepsilon_{k^*} = \mathbb{E}\varepsilon_{m^*+1} = \cdots = \mathbb{E}\varepsilon_n = 0$ and $\mathbb{E}\varepsilon_{k^*+1} = \cdots = \mathbb{E}\varepsilon_{m^*} = a$, for some unknown $1 < k^* < m^* \leq n$ and $a \neq 0$. Here k^* is the beginning and m^* is the end of epidemic change state.

To be more precise, suppose we are given observations $y_1, y_2, \ldots, y_n$ generated by the first order autoregressive process

$$y_k = \phi y_{k-1} + \varepsilon_k + a_k, \quad k = 1, \ldots, n, \ y_0 = 0, \tag{1}$$

where (ε_k) are i.i.d., centered and at least square integrable random variables, ϕ is a constant in $(-1, 1]$ and a_k is a sequence of numbers. The aim of this paper is to propose test statistics for the null hypothesis

$$H_0 : a_1 = \cdots = a_n = 0$$

E-mail address: jurgita.markeviciute@mif.vu.lt.

http://dx.doi.org/10.1016/j.spl.2016.02.001
0167-7152/

against the alternative:

$$H_A: \text{ there exist } 1 \le k^* < m^* \le n \quad \text{such that } a_k = a\mathbf{1}_{\mathbb{I}^*}(k),\ a \neq 0,\ 1 \le k \le n,$$

where $\mathbb{I}^*$ is an interval $\mathbb{I}^* = \{k^*+1, \dots, m^*\}$ and $\mathbf{1}_{\mathbb{I}^*}$ denotes its indicator function. If a_k is not zero for some interval $\mathbb{I}^*$, it might be interpreted as the change in the mean of the innovations. We also denote $\ell^* = m^* - k^*$ as the length or duration of the epidemic state and we assume that $\ell^* \to \infty$, as $n \to \infty$.

Though the epidemic change model can be formulated as multiple change-point model, Gombay (1994) pointed out that tests constructed with account of a particular form of changes may have bigger power. An epidemic type change problem in the context of abortion epidemiology was formulated by Levin and Kline (1985) (we also refer to Csörgő and Horváth, 1997, Section 1.4). Yao (1993) studied various test statistics in order to detect an epidemic change in the mean values of a sequence of independent normally distributed random variables, while Gombay (1994) suggested statistic based on ranks and signs of observations. For more information on this subject see, for example, Chen and Gupta (2012, Sections 9.3, 9.4), Aston and Kirch (2012), Jarušková and Piterbarg (2011), Račkauskas and Suquet (2006).

In this paper we use a classical methodology to build the test statistics, that is using continuous functionals of a partial sums process. Set for any random sequence (X_k)

$$S_n(0) = 0, \qquad S_n(t) = \sum_{1 \le k \le nt} X_k, \quad 0 < t \le 1.$$

A natural way to construct test statistic for detecting the epidemic change in the mean of random variables is to construct the uniform increments statistic:

$$T_{0,n}(X_1, \dots, X_n) = \max_{1 \le \ell \le n-1} \max_{1 \le k \le n-\ell} \left| S_n\left(\frac{k+\ell}{n}\right) - S_n\left(\frac{k}{n}\right) - \frac{\ell}{n} S_n(1) \right|.$$

Such statistic can detect epidemic state whose length ℓ^* is such that $n^{1/2} = o(\ell^*)$ (see Račkauskas and Suquet, 2004a). For the shorter durations, Račkauskas and Suquet (2004a) proposed the uniform increments statistics with an additional normalization. For $\alpha \in [0, 1)$ the class of statistics is defined by

$$T_{\alpha,n}(X_1, \dots, X_n) = \max_{1 \le \ell \le n-1} \ell^{-\alpha} \max_{1 \le k \le n-\ell} \left| S_n\left(\frac{k+\ell}{n}\right) - S_n\left(\frac{k}{n}\right) - \frac{\ell}{n} S_n(1) \right|.$$

Račkauskas and Suquet (2004a) showed that statistics $T_{\alpha,n}(X_1, \dots, X_n)$, where (X_k) are i.i.d. random variables, for any $0 < \alpha < 1/2$, are able to detect epidemics with duration $n^\delta = o(\ell^*)$, where $\delta = (1-2\alpha)/(2-2\alpha)$. Mikosch and Račkauskas (2010) have studied $T_{\alpha,n}$ for regularly varying random variables (ε_i) and $0 < \alpha < 1$. The uniform increments statistics for the first order nearly nonstationary autoregressive processes were studied by Markevičiūtė et al. (2014).

Together with the uniform increments statistics we investigate dyadic increments statistics defined by (see Račkauskas and Suquet, 2004a)

$$D_{\alpha,n}(X_1, \dots, X_n) = \max_{1 \le j \le \log_2(n)} 2^{j\alpha} \max_{r \in D_j} \left| S_n(r) - \frac{1}{2} S_n(r^+) - \frac{1}{2} S_n(r^-) \right|,$$

where D_j is the set of dyadic numbers of the level j, $\log_2(n)$ denotes the logarithm with basis 2, $r^+ = r + 2^{-j}$ and $r^- = r - 2^{-j}$. It is worth mentioning that uniform increments and dyadic increments statistics are equivalent, but the dyadic increments statistics are easier to implement in numerical computations, while uniform increments statistics is more convenient in theoretical investigations.

Since our data generating process is a first order autoregressive process, the innovations are not observed, so we construct our test statistics on observations $y_1, y_2, \dots, y_n$. Set for $0 \le \alpha < 1/2$

$$\widetilde{T}_{\alpha,n} := T_{\alpha,n}(y_1, \dots, y_n), \tag{2}$$

$$\widetilde{D}_{\alpha,n} := D_{\alpha,n}(y_1, \dots, y_n). \tag{3}$$

We are interested in how the use of extra weighting (the case when $\alpha > 0$) improves the detection of (relatively) short epidemics. Of course the range of detection will be smaller here than that in the case of independent samples.

The choice of two statistics (with $\alpha = 0$ and $\alpha \neq 0$) requires different assumptions on integrability of innovations. If $\alpha = 0$, then the innovations are assumed to have finite second moment. Whereas for $\alpha \neq 0$ the innovations should satisfy the stronger integrability condition

$$\lim_{t \to \infty} t^p P(|\varepsilon_0| > t) = 0, \quad p \ge 2,\ p = \frac{1}{1/2 - \alpha}. \tag{4}$$

The paper is organized as follows. Section 2 is devoted to some preliminaries and Hölderian framework. Section 3 presents the limit behavior of statistics under null hypothesis. In Section 4 we show the consistency of test statistics. Section 5 is devoted to the functional limit theorems that are used as tools in the proofs of main results.