Abstract

We consider a (possibly) long-range dependent sequence with a shift in the mean. We estimate the location of the change-point using a cumulative sum estimator. The $1/n$ convergence rate typical of the independent case is also achieved for short-memory and long-memory sequences. To cite this article: S. Ben Hariz, J.J. Wylie, C. R. Acad. Sci. Paris, Ser. I 341 (2005).

1. Introduction

There is a wide range of important applications in which a change in the marginal distribution of a random sequence must be detected. The problem has been widely studied in the literature, and we refer the reader to the monograph of Csörgő and Horvath [2] for a comprehensive review on the subject.

In the case in which the sequence is independent, both parametric and nonparametric methods have been considered. For example, in the nonparametric setting the problem was considered by Carlstein [1]. This result was subsequently improved by Dumbgen [3] and later by Ferger [4] who obtained the exact rates of convergence in both probability and in an almost surely sense. Several works are concerned with the generalization of these results to a weakly dependent setting.

\[ E-mail \text{ addresses: samir.ben_hariz@univ-lemans.fr (S. Ben Hariz), mawylie@cityu.edu.hk (J.J. Wylie).} \]
In recent years the importance of long-memory or long-range dependent (LRD) processes has been realized in numerous applications, especially financial and telecommunication data. A second-order stationary sequence \( (X_i)_{i \geq 0} \), is generally said to be LRD if
\[
\sum_{i=1}^{\infty} |\text{Cov}(X_0, X_i)| = \infty.
\] (1)
Otherwise it is said to be short-memory or short-range dependent (SRD). The change-point test and estimation in the LRD setting was considered by Giraitis et al. [5] using a nonparametric method for the case in which the difference between the distribution functions before and after the jump tends to zero with increasing sequence length. Kokoszka and Leipus [7] considered the change in the mean for dependent observations for LRD sequences. They obtained rates in probability for the cumulative sum (CUSUM) change-point estimator and give a rate of consistency of the estimator that gets worse as the strength of the dependence increases. The problem with a jump in the mean that tends to zero was considered by Horvath and Kokoszka [6]. They proved the consistency of the estimator and gave the limiting distribution.

In this Note we will develop a unified framework for estimating a change in the mean of stationary sequences that can be either SRD or LRD. We will give rates of convergence for the widely used CUSUM estimator defined in (2). We show that the rate of convergence of the CUSUM estimator is independent of the decay rate of the correlation function and so, under very weak conditions, independent of the dependence structure. So the \( 1/n \) rate in the independent case is also achieved for both LRD and SRD sequences.

2. Main results

Let \( (X_i)_{i=1,\ldots,n} \) be a stationary sequence that can be either independent, SRD or LRD. We assume that \( \mathbb{E}(X^2) < \infty \) and without loss of generality we take \( \mathbb{E}(X) = 0 \). The correlation function of the sequence is given by: \( r(i) = \text{Cor}(X_0, X_i) \). Let \( (Y_i)_{i=1,\ldots,n} \) be a sequence with a shift in the mean, defined by \( Y_i = X_i + \delta \mathbb{1}_{i \leq n \theta_0} \) where \( 0 < \theta_0 < 1 \) is the location of the change-point. Given a sequence, \( (Y_i)_{i=1,\ldots,n} \), we aim to estimate the change-point \( \theta_0 \) using the following family of estimators:
\[
\theta_n = \frac{1}{n} \min \left\{ \text{argmax}_{1 \leq k < n} \{ |U_k| \} \right\},
\] (2)
where
\[
U_k = \left( \frac{k(n-k)}{n} \right)^{1-\gamma} \left( \frac{1}{k} \sum_{i=1}^{k} Y_i - \frac{1}{n-k} \sum_{i=k+1}^{n} Y_i \right)
\]
and \( \gamma \) is a parameter satisfying \( 0 \leq \gamma < 1 \). We now state the main result of the Note:

**Theorem 2.1.** Let \( (X_i)_{i=1,\ldots,n} \) be a stationary sequence with correlation function \( r(n) \). Assume that there exist constants \( B > 0 \) and \( \alpha > 0 \) such that \( |r(i)| \leq Bi^{-\alpha} \) for all \( i \). Then we have
\[
\lim_{x \to +\infty} \lim_{n \to +\infty} \mathbb{P}(n|\theta_n - \theta_0| > x) = 0
\] (3)
where \( \theta_n \) is defined in (2).

We note that under a very weak condition on the decay of the correlation function, we obtain the same rate for every value of \( \alpha \). We recall that if \( r(i) \sim Bi^{-\alpha} \), then \( \alpha > 1 \) corresponds to the SRD case and \( \alpha < 1 \) to the LRD case. So the \( 1/n \) rate is achieved for both SRD and LRD sequences.

The second result deals with the case in which the jump size, \( \delta = \delta_n \), depends on \( n \). As mentioned in the introduction, this has already been studied by Horváth and Kokoszka [6] where the authors obtained the exact rate in probability (see Lemma 4.5 therein), and the limiting distribution for Gaussian sequences. Here under slightly different conditions on the size of the jump we obtain the same rate in probability for more general sequences.
Corollary 2.2. Under the conditions of Theorem (3). Suppose \(|\delta_n| \to 0\) and either \(|\delta_n| n^{\alpha/2} (\ln n)^{-1} \to \infty\) if \(\gamma = 1 - \alpha/4\), \(|\delta_n| n^{\alpha/2} / (\ln n)^{-1} \to \infty\) if \(\gamma = 1 - \alpha/4\), then

\[
\lim_{x \to +\infty} \lim_{n \to +\infty} \mathbb{P}(n |\delta_n|^{2/\alpha} | \theta_n - \theta_0| > x) = 0.
\]

Before presenting the proof, we give a qualitative explanation of why the rate is independent of \(\alpha\). We define \(t_k := k/n\), then \(U_k := U_n(t_k)\) where \(U_n\) is expressed as a sum of its mean and a centered random component:

\[
U_n(t) := \frac{n^{1-\gamma}}{(t(1-t))^{\gamma}} \left( \delta g(t) + B_n(t) \right).
\]

Here, the mean component is given by \(g(t) = (1 - \theta_0) t \mathbb{1}_{t \leq \theta_0} + \theta_0(1-t) \mathbb{1}_{t > \theta_0}\) and the centered random component, \(B_n\), is given by

\[
B_n(t) := W_n(t) - t W_n(1) \quad \text{and} \quad W_n(t) := \frac{1}{n} \sum_{i=1}^{[nt]} X_i.
\]

Under appropriate conditions on the underlying sequence, the sequence of functions \(W_n(\cdot)\) (when suitably normalized) converges, in a weak sense, to either a Brownian motion, a fractional Brownian motion or more generally a Hermite process. The fact that the rate of convergence for the estimator is independent of \(\alpha\) is due to the cancellation of two opposing effects: the rate of convergence of partial sums, and the size of fluctuations of the Donsker line \(W_n(t)\). Assume for simplicity that \(r(i) \sim Ci^{-\alpha}\). Roughly speaking, in the LRD case \((\alpha < 1)\), \(W_n(t)\), and hence \(B_n(t)\), converges to zero with a rate of the order \(n^{-\alpha/2}\). So as \(\alpha\) decreases, the size of the random component of \(U_k\) becomes larger, making the estimation more difficult. However, as \(\alpha\) decreases \(W_n(t)\), and hence \(B_n(t)\), become more regular in the sense that \(n^{-\alpha/2} |W_n(t) - W_n(t')|\) has size of order \(|t - t'|^{1-\alpha/2}\). Hence, local fluctuations in the random component of \(U_k\) are reduced and so estimation becomes easier. These two effects cancel each other and the overall rate is independent of \(\alpha\).

3. Sketch of the proof

In what follows \(K_1, K_2, \ldots\) denote positive constants that are independent of \(n\). We will take \(\alpha < 1\), since any correlation function that satisfies the conditions of the theorem for \(\alpha \geq 1\) will also satisfy the conditions for any smaller value of \(\alpha\). Without loss of generality we also assume that the size of the jump is positive (\(\delta > 0\)).

We define the following sets

\[
S_{n,j} = \{ \theta: 2^j < r_n |\theta - \theta_0| \leq 2^{j+1} \},
\]

where \(r_n\) is a positive sequence to be chosen later. Let \(0 < \eta < \min(\theta_0, 1 - \theta_0)/2\) and \(J = J(n, \eta)\) be chosen such that \(2^J < r_n \eta \leq 2^{J+1}\). For the sake of brevity, we define \(w(\theta) := (\theta(1 - \theta))^{\gamma}\) and \(h(\theta) := g(\theta)/w(\theta)\). Let \(\theta_n\) be a maximum of \(\{|U_n(t)|, t \in G_n\}\), where \(G_n := [k/n, 1 \leq k < n]\). Then

\[
\mathbb{P}(r_n |\theta_n - \theta_0| > 2^M) \leq \sum_{j=M}^{J} \mathbb{P}(\theta_n \in S_{n,j}) + \mathbb{P}(|\theta_n - \theta_0| > \eta)\) \tag{6}
\]

Since \(\theta_n\) is a maximum, \(|U_n(\theta_n)| \geq |U_n(\theta_0)|\). In order to control \(\mathbb{P}(\theta_n \in S_{n,j})\) the only difficulty arises when both \(U_n(\theta_n)\) and \(U_n(\theta_0)\) are positive, in which case we have to bound

\[
\mathbb{P}\left( \frac{B_n(\theta_n)}{w(\theta_n)} - \frac{B_n(\theta_0)}{w(\theta_0)} \geq \delta (h(\theta_0) - h(\theta_n)) \right) \geq \delta C w(\theta_0) \frac{2^{j-1}}{r_n}\) \tag{7}
\]

This in turn can be reduced to considering

\[
\mathbb{P}\left( \sup_{\theta \in S_{n,j}} \left| \left( B_n(\theta) - B_n(\theta_0) \right) \right| \geq \delta C w(\theta_0) \frac{2^{j-1}}{r_n} \right).
\]

To bound the above expression, we need the following maximal inequality which is a special case of Theorem 1 in Moricz [8].
Lemma 3.1. Assume that there exists constants $B > 0$ and $\alpha < 1$ such that $|r(i)| \leq Bi^{-\alpha}$ for all $i$. Then there exists a constant $D = D(B, \alpha)$ such that

$$\mathbb{E} \left( \max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} X_j \right| \right)^2 \leq D \mathbb{E}(X^2)n^{2-\alpha}. \tag{8}$$

Using (6) and (8) yields

$$\mathbb{P}(r_n|\theta_n - \theta_0| > 2^M) \leq K_5 \delta^{-2}\sum_{j=M}^{\infty} \mathbb{E}(X^2)2^{-j\alpha} + \mathbb{P}(|\theta_n - \theta_0| > \eta). \tag{9}$$

According to the results of Kokoszka and Leipus [7], $\lim_{n \to \infty} \mathbb{P}(|\theta_n - \theta_0| > \eta) = 0$. Taking $r_n = n$ and letting $n$ go to $\infty$ yields

$$\lim_{n \to \infty} \mathbb{P}(n|\theta_n - \theta_0| > 2^M) \leq K_6 \delta^{-2}\sum_{j=M}^{\infty} 2^{-j\alpha}.$$

Letting $M$ tend to infinity completes the proof of (3). The corollary is just a simple consequence of Corollary 1.1 in Kokoszka and Leipus [7] and (9).

References