Application of Self-normalized Method in Long-memory Multi-Means Change-Point Test

Liu Yi

{Wuoshiwzm@hotmail.com}

Xi'an University of Science and Technology Xi'an, Shaanxi 710699

Abstract. In this paper, we have developed a novel attempt to be sensitive to multiple means of long-memory time series in an unsupervised manner using our own legitimate method.Self-regular, can avoid estimating the gradual variance variance and use the regular method at the same time. The method can be conveniently and conveniently applied to the first-order stationary data with long memory (stationary with long memory) dependency), no change points, statistics are collected and summarized in non-exit distribution. We describe this statistic and evaluate its effects. At the same time, the feasibility of the method is illustrated by real data.

Keywords: Long memory, time series, change points, self-normalized.

1 Introduction

The research on the change point of time series first started decades ago, GOMBAY and HORVATH (1990) [1], discussed the use of maximum likelihood estimation mean.

Long-term dependent data, also known as long-memory data or long-persistence data, is a phenomenon that has recently emerged frequently in time-series data analysis and is usually observed in long-term series data. In fact, there are many definitions of long memory. In this article, we use the following definitions. Let:

$$(1-B)^{\mathrm{m}}(X_t-\mu)=u_t, t\in Z \quad (1)$$

where B is the post-position operator, μ is the unknown mathematical expectation, and, $\{u_t\}_{t\in Z}$ is a short dependent sequence with 0 mean and finite variance.。 When the parameter m = 0, the time series X_t is called a " short-memory time series ", when m ∈ (−1/2, 0) ∪ (0, 1/2), we call $\{X\}_{t=1}^{n}$ is the " long-memory time series ".

When we try to construct first-order statistics such as mean change point, cumulative sum statistics (CUSUM) is often a more commonly used method. The proof of its limit theory is relatively simple, and the calculation of numerical simulation is relatively easy.

Its form is similar to the following:

$$G_n(\theta)=\frac{\sqrt{n}(\theta_n-\theta)}{\hat{\sigma}_F} \quad (2)$$

However, this kind of statistical structure often has an additional condition, that is, the asymptotic variance needs to be estimated, and these methods without exception refer to more preset parameters. For the drawbacks of estimating asymptotic variance for CUSUM statistics,

MSEA 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2334401

Kiefer et al. (2000)[2] and Lobato (2001)[3] first proposed similar methods, which can successfully avoid estimating asymptotic variance.

2 Model and test statistics

At first we want to discuss the mean change point. With the series $X_1, \ldots, X_n$ given, to test the H_0 , of no change in the mean, namely:

Let $\{X\}_{t=1}^{n}$ be the long memory time series. The null hypothesis is that no change point.

$$H_0 : E(X_1) = E(X_1) = \cdots = E(X_n) = \mu \tag{3}$$

The alternative is that there being more than one change point, that is, there are m ≥ 1 change points, $k_0 = 1 < k_1 < \cdots < k_m < n = k_{m+1}$ make:

$$H_1 : E(X_i) \neq E(X_{i+1}), i \in \{k_1, \ldots, k_m\} \tag{4}$$

For other positions, $E(X_i) = E(X_{i+1})$。 The number of change points m, and the position $k_1 \ldots k_m$ are unknown yet. Let $\bar{X}_{j,k} = (k - j + 1)^{-1} \sum_{i=j}^{k} X_i$ be the sample mean of X , $S(j, k) = \sum_{i=j}^{k} X_i$ be the corresponding partial sum, with 1≤j≤k≤n.

To test the null hypothesis (3), a common method is to compare recursive means: $\bar{X}_{1,k} = S_{1,k}/k$, the global mean $\bar{X}_{1,n} = S_{1,n}/n$, And check whether there is an obvious deviation between the two, and the corresponding research object becomes the called " cumulative sum process " :

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} (X_i - \bar{X}_{1,n}), \ t \in [0,1]$$

⌊x⌋ means the largest integer that less than x. As the H_0 (3) of no change points, stochastic process $\{Z_n(t)\}_{t\in[0,1]}$ independent of the unknown μ. A suitable critical value can be obtained by the central limit theory of cumulative sum process as follows.

$$S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} (X_i - \mu), \ t \in [0,1]$$

Let B( ) be the standard Brownian motion, ⇒ means weak convergence in Skorokhod space, and the following assumptions are made:

(IP) There exists σ > 0, such that {Sn(t), 0 ≤ t ≤ 1} ⇒ {σB(t), 0 ≤ t ≤ 1} , let:

$$S_n(r) = n^{-(\frac{1}{2}+d)} \sum_{t=1}^{\lfloor nt \rfloor} (X_i - \mu) \quad \rightarrow \quad C_d B_d(r), r \in [0,1]$$

$$Z_n(t) = n^{-(\frac{1}{2}+d)} \sum_{t=1}^{\lfloor nr \rfloor} (X_t - \bar{X}_{1,n})$$, then there is:

$$Z_n(t) \rightarrow C_d B_d(t) - tB_d \tag{5}$$

where Cd is a positive constant, Bd(·) s called fractional Brownian motion.The assumption (IP) is often referred to as the principle of invariance.

For example the mixing coefficients of Hannan (1979)[4] and Herrndorf (1984)[5], the functional dependence measures of Berkes et al. and other reference contributions therein. By postulate (IP) and the continuous mapping theorem, weak convergence can be obtained as:

$$Z_n(t) \to_D \sigma\{B(t) - tB(1)\}$$

As reviewed by Shao and Zhang (2010) [6], both theoretical and empirical studies in the literature have found that using data-dependent bandwidth leads to testing with non-monotonic powers. To solve this problem, Shao and Zhang (2010) propose to adopt the idea of self-regularization and generalize it to the change point problem.

Proof.

$$Z_n(t) = n^{-\left(\frac{1}{2}+d\right)} \sum_{1}^{(nt)} \left[x_i - \mu - \frac{1}{n} \sum_{1}^{n} (x_j - \mu) \right]$$

$$= n^{-\left(\frac{1}{2}+d\right)} \sum_{1}^{\lfloor nt \rfloor} (x_i - \mu) - n^{-\left(\frac{1}{2}+d\right)} \frac{\lfloor nt \rfloor}{n} \sum_{1}^{n} (x_j - \mu)$$

$$= n^{-\left(\frac{1}{2}+d\right)} \textstyle\sum_1^{\lfloor nt \rfloor} (X_i - \mu) - t \cdot n^{-\left(\frac{1}{2}+d\right)} \sum_1^{\lfloor n \times 1 \rfloor} (x_j - \mu) \to C_d B_d(t) - t \cdot C_d B_d \quad (6)$$

The postulate (IP) generally refers to the principle of invariance and has been shown to hold for most short dependent processes, e.g. Hannan (1979) and Herrndorf (1984) for mixed parameters. From the hypothesis (IP) and the principle of continuous mapping, we get the following weak convergence conclusion:

$Z_n(t) \to \sigma_D\{B(t) - tB(1)\}$, This asymptotic distribution depends on the parameter σ。

Statistics T_n:

In the statistics, n corresponds to the sample population, and $\Xi_{n,f}(j_1, j_2, j_3)$ and $\Xi_{n,b}(j_1, j_2, j_3)$ correspond to the test statistics of positions $j_1, j_2, j_3$to determine whether there is change point. The overall statistics traverse all the combinations of j_1, j_2, j_3 as a judgment on whether there is a change point as a whole.

$$D_{n,f}(j_1, j_2, j_3) = \frac{j_2 - j_1 + 1}{\sqrt{j_3 - j_1 + 1}} \times \left(\bar{X}_{j_1, j_2} - \bar{X}_{j_1, j_3} \right) \quad (7)$$

$$L_{n,f}(j_1, j_2, j_3) = \textstyle\sum_{i=j_1}^{j_2} \left(\frac{i - j_1 + 1}{j_3 - j_1 + 1} \right)^2 \times \left(x_{j_1, i} - x_{j_1, j_2} \right)^2 \quad (8)$$

$$R_{n,f}(j_1, j_2, j_3) = \textstyle\sum_{i=j_2+1}^{j_3} \left(\frac{j_3 - i + 1}{j_3 - j_1 + 1} \right)^2 \times \left(\bar{X}_{i, j_3} - \bar{X}_{j_2+1, j_3} \right)^2 \quad (9)$$

$$\Xi_{n,f}(j_1, j_2, j_3) = L_{n,f}(j_1, j_2, j_3) + R_{n,f}(j_1, j_2, j_3) \quad (10)$$

$$T_{n,f} = \max_{0 \le x \le 1} \frac{D_{n,f}(1, l_1, l_2)^2}{\Xi_{n,f}(1, l_1, l_2)} \quad (11)$$

$$D_{n,b}(j_1, j_2, j_3) = \frac{j_3 - j_2 + 1}{\sqrt{j_3 - j_1 + 1}} \times \left(\bar{X}_{j_2, j_3} - \bar{X}_{j_1, j_3} \right) \quad (12)$$

$$L_{n,b}(j_1, j_2, j_3) = \textstyle\sum_{i=j_1}^{j_2-1} \left(\frac{i - j_1 + 1}{j_3 - j_1 + 1} \right)^2 \times \left(\bar{X}_{j_1, i} - \bar{X}_{j_1, j_2 - 1} \right)^2 \quad (13)$$

$$R_{n,b}(j_1, j_2, j_3) = \sum_{i=j_2}^{j_3} \left(\frac{j_3 - i + 1}{j_3 - j_1 + 1}\right)^2 \times \left(\bar{X}_{i,j_3} - \bar{X}_{j_2,j_3}\right)^2 \quad (14)$$

$$\Xi_{n,b}(j_1, j_2, j_3) = L_{n,\mathrm{b}}(j_1, j_2, j_3) + R_{n,b}(j_1, j_2, j_3) \quad (15)$$

$$T_{n,b} = \max_{(m_1, m_2) \in \Omega_n(\varepsilon)} \frac{D_{n,b}(m_1, m_2, n)^2}{\Xi_{n,b}(m_1, m_2, n)} \quad (16)$$

$$\Omega(\varepsilon) = \{(t_1, t_2) : \varepsilon \le t_1 \le t_2 \le 1 - \varepsilon\},$$

$$\Omega_n(\varepsilon) = \{(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor) : (t_1, t_2) \in \Omega(\varepsilon)\}$$

Our test statistic is:

$$T_n = T_{n,b} + T_{n,f} = \max_{(l_1, l_2) \in \Omega(\varepsilon)} \frac{D_{n,f}(1, l_1, l_2)^2}{\Xi_{n,f}(1, l_1, l_2)} + \max_{(m_1, m_2) \in \Omega_n(\varepsilon)} \frac{D_{n,b}(m_1, m_2, n)^2}{\Xi_{n,b}(m_1, m_2, n)} \quad (17)$$

Unlike Shao's supervised test, which fixed the partitions by formulating segments according to a pre-specified change points number, the provided test is free of supervised, with change points detected. Following theorem tells the asymptotic properties of Tn under H_0 and H_A .

Theorem 2.1. H_0 Limit Theory

$$T_n \to T(B) \quad (18)$$

And we have

$$T(B) = \sup_{(r_1, r_2) \in \Omega(\varepsilon)} \frac{D(B, 0, \gamma_1, r_2)^2}{\Xi(B, 0, \gamma_1, r_2)} + \sup_{(s_1, s_2) \in \Omega(\varepsilon)} \frac{D(B, S_1, s_2, 1)^2}{\Xi(B, S_1, S_2, 1)}$$

$$D(B, t_1, t_2, t_3) = \frac{1}{\sqrt{t_3 - t_1}} \times \left[B(t_2) - B(t_1) - \frac{t_2 - t_1}{t_3 - t_1} \times \{B(t_3) - B(t_1)\}\right]$$

$$\Xi(B, t_1, t_2, t_3)$$

$$= \frac{1}{(t_3 - t_1)^2} \times \left(\int_{t_1}^{t_2} \left[B(s) - B(t_1) - \frac{s - t_1}{t_2 - t_1} \times \{B(t_2) - B(t_1)\}\right]^2 ds \right)$$

$$+ \int_{t_2}^{t_3} \left[B(t_3) - B(s) - \frac{t_3 - s}{t_3 - t_2} \times \{B(t_3) - B(t_2)\}\right]^2 ds$$

When there is no change point, the limiting distribution of the T_n statistic is fractional Brownian motion T(B). And there is no need to preset the number of change points, so this statistic is called unsupervised.

Theorem 2.2. The limiting distribution of H_A

When exists $m \ge 1$, $k_0 = 1 < k_1 < \cdots < k_m < n = k_{m+1}$, such that

$$E(X_\mathrm{i}) \ne E(X_{\mathrm{i}+1}), \mathrm{i} \in \{k_1, \ldots, k_m\}$$

while $E(X_\mathrm{i}) = E(X_{\mathrm{i}+1})$ in other position , then:

$$T \to \infty \quad (19)$$

When there is a change point, the limit distribution of statistics tends to ∞。

The test procedure above involves a pruning parameter ε, which controls the minimum length (proportionally) of the part amount. As Shao and Zhou (2013) [7] commented, fine-tuning parameters in autonormalization are different from smoothing parameters, such as truncated lags in long-term variance estimation (Newey and West, 1987 [8], Liu and Wu, 2010 [9], Politis, 2011[10]), the window size is dependent on the bootstrap (Shao, 2010, Zhou, 2013) or the width of the subsampling method (Hall and Jing, 1996; Politis et al., 1999; Zhang et al., 2013), such as considering the effect of trimming in the limiting distribution and its approximation. Zhou and Shao (2013), as well as Huang et al. (2015)[11], performed a series of numerical analyzes and concluded that the empirical choice of ε = 0.1 produce satisfactory self-regularizing methods. The provided test has the potential to generalize to other quantity-ratio means, such as quantiles, which we discuss in subsequent sections.

3 Numerical simulation

This article uses the R language arfima class library to generate long-memory time series data. Such libraries can set data change points number and bias. In this paper, number of generated sequence data points is 10000, and the deviation system d is divided into: -0.4, -0.3, -0.2, -0.1,0, 0.1, 0.2, 0.3, 0.4.

Table 1. Size and power of T_n for real time series

	No change point		With change point	
	Size		Power	
d	10%	5%	10%	5%
-0.4	0.8613	0.9210	0.9117	0.9618
-0.3	0.8987	0.9637	0.9031	0.9561
-0.2	0.9118	0.9355	0.9033	0.9481
0	0.8825	0.9374	0.8998	0.9564
0.1	0.9135	0.9557	0.8919	0.9404
0.2	0.8924	0.9375	0.9142	0.9552
0.3	0.9080	0.9479	0.9037	0.9582
0.4	0.8695	0.9524	0.8982	0.9489

It can be seen that our statistics can provide better Size and Power for different long-memory time series when d is different.

Table 2 Critical value of mean change point statistics under different dependence strengths

	SZ101		SZ102		SZ103		T	
ρ	90%	95 %	90%	95%	90%	95%	90%	95%
-0.4	0.876	0.933	0.872	0.972	0.900	0.928	0.861	0.921
-0.3	0.827	0.954	0.878	0.927	0.845	0.935	0.898	0.963
-0.2	0.856	0.934	0.785	0.884	0.722	0.827	0.911	0.935
0	0.825	0.967	0.934	0.942	0.947	0.938	0.882	0.937
0.1	0.925	0.928	0.937	0.946	0.936	0.946	0.913	0.955
0.2	0.925	0.926	0.937	0.949	0.948	0.931	0.892	0.937
0.3	0.934	0.912	0.933	0.945	0.946	0.932	0.908	0.947
0.4	0.923	0.965	0.956	0.934	0.934	0.945	0.869	0.952

Among them, SZ10 is the statistic of Shao and Zhang (2010), and T is the statistic of this paper.It can be seen from Table 2 that this statistic has a relatively stable performance.

4 Conclusion

In this paper, based on the self-regularization method, the multi-variation point test problem under unsupervised conditions is studied based on the time series of long-memory multi-variation points. From the simulation experiments, the proposed new method still has good properties in the case of no uniform assumed change point number.

References

[1] GOMBAY, E. and L. HORVATH (1990, 06). Asymptotic distributions of maximum likeli-hood tests for change in the mean. Biometrika 77(2), 411–414.
[2] Kiefer, N. M., T. J. Vogelsang, and H. Bunzel (2000). Simple robust testing of regression hypotheses. Econometrica 68(3), 695–714.
[3] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. Journal of the American Statistical Association 96(455), 1066–1076.
[4] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. The Annals of Probability 12 (1), 141–153.
[5] Shao, X. (2010). A self-normalized approach to confidence interval construction in time series. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72(3), 343–366.
[6] Zhou, Z. and Shao, X. (2013). Inference for linear models with dependent errors. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 75:323–343.
[7] Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3):703–708.
[8] Liu, W. and Wu, W. B. (2010). Asymptotics of spectral density estimates. Econometric Theory, 26:1218–1245.
[9] Politis, D. N. (2011). Higher-order accurate, positive semi-definite estimation of large-sample 24 covariance and spectral density matrices. Econometric Theory, 27:703–744.
[10] Huang, Y., Volgushev, S., and Shao, X. (2015). On self-normalization for censored dependent data. Journal of Time Series Analysis, 36:109–124.
[11] Huang, Y., Volgushev, S., and Shao, X. (2015). On self-normalization for censored dependent data. Journal of Time Series Analysis, 36:109–124.