Statistics and Probability Letters 99 (2015) 36–43

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

A hybrid test for the isotonic change-point problem

Gang Shen[a,*], Karl D'Silva[b]

[a] Department of Statistics, North Dakota State University, United States
[b] Department of Mathematics and Statistics, St. Cloud State University, United States

ARTICLE INFO

Article history:
Received 11 October 2014
Received in revised form 30 December 2014
Accepted 31 December 2014
Available online 9 January 2015

Keywords:
Change-point
Isotonic
Minimax
Mixing
Nonparametric

ABSTRACT

This work studies minimaxity of the asymptotic risk of Brillinger's test and proposes a hybrid test for the isotonic change-point problem. The new test turns out to have asymptotic power 1 under some local alternatives and outperforms the other major tests in detecting some weak isotonic signals. An application to live data related with the Amazon flooding issue is illustrated.

1. Introduction

Many recent researches on the topic of areal and global natural catastrophic events like increasing rainfall and flooding height, or global warming are related with the isotonic change-point problem in statistics, which is aimed at answering the question whether the seemingly increasing intensity is caused by natural variability or an underlying isotonic trend induced by human activity. It can essentially be put into a statistical hypothesis test: Homogeneity of means versus monotonicity of means for the observed time series. To be specific, suppose the time series X_k observable at time $k = 1, \ldots, n$ has the form of $X_k \triangleq \mu_k + Z_k$, where Z_k is a zero mean strictly stationary noise with positive spectral density and μ_k is the deterministic signal satisfying $\mu_1 \le \mu_2 \le \cdots \le \mu_n$, we want to run a hypothesis test with

$$\begin{aligned} H_0: \ \mu_k &= \mu_{k+1}, \quad \text{for all } k = 1, \ldots, n-1, \\ H_a: \ \mu_k &< \mu_{k+1}, \quad \text{for some } k = 1, \ldots, n-1. \end{aligned}$$

The challenge of this problem is that the signal μ_k has an arbitrary unknown monotone structure and the change-points (the time when a jump in μ_k occurs) can be as many as the number of observations in the time series.

To consider temporal dependence among the noise Z_k, we assume the following mixing condition: Let $\mathcal{F}_a^b$ be the σ-algebra generated by $\{Z_k; a \le k \le b\}$ and $\beta_n \triangleq E \sup\{|P(A \mid \mathcal{F}_1^m) - P(A)|;\ A \in \mathcal{F}_{m+n}^\infty,\ m \ge 1\}$, there exists some $\delta > 0$ such that $n^{1+\delta}\beta_n \to 0$. The class of random processes satisfying this mixing condition includes m-dependent processes and Gaussian stationary processes with a spectral density bounded below such that $\operatorname{cov}(Z_0, Z_n) = O(n^{-3-\delta_1})$, $\delta_1 > \delta > 0$, vide Doukhan (1985, Section 2.1.1, p. 59). The assumption of a mixing condition is fairly common in literature for time series with short-range dependence but not necessarily a linear process, see Doukhan (1985), Rosenblatt (2000), or Brodsky and Darkhovsky (1993).

* Corresponding author.
E-mail address: gang.shen@ndsu.edu (G. Shen).

http://dx.doi.org/10.1016/j.spl.2014.12.025
0167-7152/

In this work, we additionally require that the marginal cdf of Z_1, $F_0(\cdot)$, belongs to a location family with positive pdf, like Shen and Xu (2013) (hereinafter abbreviated as [SX]). The mixing condition above then guarantees $\sigma^2 = \gamma_0 + 2\sum_{k=1}^{+\infty}\gamma_k < \infty$, where

$$\gamma_k \triangleq \text{cov}\,(F_0(Z_1), F_0(Z_{k+1})), \quad k = 0, 1, \ldots \tag{1}$$

because $|\gamma_k| \le 4\beta_k$ by Doukhan (1985, Section 1.2.2, Lemma 3).

With the assumption of the mixing condition on noise, Brillinger (1989) applies a contrast test to the isotonic change-point problem. Wu et al. (2001) consider a penalized likelihood ratio test based on the isotonic regression estimates of μ_k. Wu's test has a seemingly larger power than Brillinger's test, however it suffers inflated type I error, particularly when the time series X_k has autocorrelation. Wu's test relies on an undesirable tuning parameter and stabilizes slowly in distribution to an integral of quite a complicated stochastic process which does not have an explicit distribution. [SX] develop a U-test with a flavor of the classical Mann–Whitney test, which is asymptotically equivalent to a contrast of the cdf transformed data. Although all these tests have asymptotic power one under H_a with suitable assumptions on μ_k, the numerical study in [SX] shows that the U-test has a larger power than Wu's test and Brillinger's test when μ_k is of a linear or staircase pattern. In this work, we study the minimaxity of the asymptotic risk of Brillinger's test and propose a hybrid test which has the feature of both Brillinger's test and U-test. The new test turns out to have a larger asymptotic power than all these tests in detecting sort of weak isotonic signals.

For simplicity, we adopt Landau notation in this work for comparing the order of two positive sequences, viz., with $a_n/b_n \to c$, (1) $a_n = \theta(b_n)$ and $b_n = \theta(a_n)$, if $0 < c < \infty$; (2) $a_n = O(b_n)$ and $b_n = \Omega(a_n)$, if $0 \le c < \infty$; (3) $a_n = o(b_n)$ and $b_n = \omega(a_n)$, if $c = 0$.

The remaining part of this paper is organized as follows: A literature review is provided in Section 2; the new test statistic and the main results are introduced in Section 3; a numerical study of our test power compared with other tests is presented in Section 4; an application is illustrated in Section 5; discussion is contained in Section 6. In order not to interrupt introduction of the main result, proof of the claims is deferred to the Appendix.

2. Literature review

2.1. The U-test

[SX] propose the U-test with a flavor of the classical Mann–Whitney test for the isotonic change-point problem on correlated data with short-range dependence, which outperforms every well-known test in literature when the sequence of deterministic signals μ_k is linearly increasing or has a single upward jump. They consider the statistic

$$U_n \triangleq \frac{\sqrt{3}}{\binom{n}{2}} \sum_{i,j=1}^{n} \left(\frac{j-i}{n+1}\right)^{+} \left\{\mathbb{1}_{\{X_i < X_j\}} - \frac{1}{2}\right\}, \tag{2}$$

where $x^+ = \max\{x, 0\}$, and report that under H_0, $\sqrt{n}U_n \xrightarrow{d} N(0, \sigma^2)$, where σ^2 is as defined in Eq. (1). [SX] indicate that under H_0, U_n is asymptotically equivalent to

$$T_n = \frac{2\sqrt{3}}{n-1} \sum_{k=1}^{n} \left\{\frac{k}{n+1} - \frac{1}{2}\right\} F(X_k) \tag{3}$$

in the sense of $E[\sqrt{n}(U_n - T_n)]^2 \to 0$, where $F(\cdot)$ is the cdf of X_1. Clearly $\sqrt{n}T_n$ is a contrast of $F(X_k)$ with coefficients $u_k = \frac{2\sqrt{3n}}{n-1}\{\frac{k}{n+1} - \frac{1}{2}\}$, since $\sum_{k=1}^{n} u_k = 0$. Moreover, the coefficients series u_k has the following properties: (i) monotone increasing; (ii) $\sum_{k=1}^{n} u_k^2 \to 1$; (iii) $\sum_{k=1}^{n-s_1} u_k u_{k+s_1} \to 1$ and $\sum_{k=1}^{n-s_l}\prod_{i=0}^{l} u_{k+s_i} \to 0$, for any $l \ge 2$ and $0 = s_0 \le \cdots \le s_l < n$ (fixed). These observations motivate us to seek a new test of the form $\frac{\sum_{k=1}^{n}\lambda_k F(X_k)}{\{\sum_{k=1}^{n}\lambda_k^2\}^{1/2}}$ with coefficients series λ_k in the class $M \triangleq M_1 \cap M_2$, which may outperform the U-test in some scenarios under H_a, where

$$M_1 \triangleq \left\{(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n | \lambda_1 \le \cdots \le \lambda_n,\ \sum_{k=1}^{n}\lambda_k = 0\right\},$$

$$M_2 \triangleq \left\{(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n \left| \frac{\sum_{k=1}^{n-s_1}\lambda_k\lambda_{k+s_1}}{\sum_{k=1}^{n}\lambda_k^2} \to 1,\ \frac{\sum_{k=1}^{n-s_l}\prod_{i=0}^{l}\lambda_{k+s_i}}{\left\{\sum_{k=1}^{n}\lambda_k^2\right\}^{l/2}} \to 0,\ l \ge 2\right.\right\}.$$