Statistical Methodology 14 (2013) 39–61

Contents lists available at SciVerse ScienceDirect

Statistical Methodology

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

A semiparametric maximum likelihood ratio test for the change point in copula models

Salim Bouzebda [a,*], Amor Keziou [b]

[a] Laboratoire de Mathématiques Appliquées-L.M.A.C., Université de Technologie de Compiègne, B.P. 529, 60205 Compiègne cedex, France

[b] Laboratoire de Mathématiques de Reims (EA 4535), Université de Reims Champagne-Ardenne. UFR Sciences, Moulin de la Housse, B.P. 1039, 51687 Reims, France

ARTICLE INFO

Article history:
Received 23 July 2012
Received in revised form
15 December 2012
Accepted 27 February 2013

Keywords:
Dependence function
Multivariate rank statistics
Semiparametric inference
Copulas
Change point
Likelihood ratio processes
Brownian bridge
Weighted approximations
Extreme value asymptotics

ABSTRACT

In the present paper, a semiparametric maximum-likelihood-type test statistic is proposed and proved to have the same limit null distribution as the classical parametric likelihood one. Under some mild conditions, the limiting law of the proposed test statistic, suitably normalized and centralized, is shown to be double exponential, under the null hypothesis of no change in the parameter of copula models. We also discuss the Gaussian-type approximations for the semiparametric likelihood ratio. The asymptotic distribution of the proposed statistic under specified alternatives is shown to be normal, and an approximation to the power function is given. Simulation results are provided to illustrate the finite sample performance of the proposed statistical tests based on the double exponential and Gaussian-type approximations.

1. Introduction and motivations

Change point models are used in many fields, for instance, archeology, econometrics, epidemiology, medicine, reliability and financial applications. For a broader presentation of the field of change-point analysis along with statistical applications, we refer the reader to the monographs by Brodsky and Darkhovsky [10], Csörgő and Horváth [18], Basseville and Nikiforov [2] and Chen and Gupta [15]. As mentioned in [17], change point problems have originally arisen in the context of quality control; see [45,46]. Interesting applications of change point models in the analysis of DNA sequences may be

* Corresponding author.
E-mail addresses: bouzebda@gmail.com, salim.bouzebda@utc.fr (S. Bouzebda), amor.keziou@univ-reims.fr (A. Keziou).

1572-3127/$ – see front matter
http://dx.doi.org/10.1016/j.stamet.2013.02.003

found in [25,8] for the segmentation of such sequences, in this setting, in general there are multiple changes in the sequences. We may refer to [33,23,1,12,35] for more recent references.

We start by giving some notation and definitions that are needed for the forthcoming sections. Consider a random vector $\mathbf{X} := (X_1, \dots, X_d)^\top \in \mathbb{R}^d$, $d \geq 1$, with the joint cumulative distribution function [d.f.]

$$\mathbb{F}(\mathbf{x}) := \mathbb{F}(x_1, \dots, x_d) := \mathbb{P}(\mathbf{X} \leq \mathbf{x}) := \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$$

with continuous marginal d.f.'s $F_j(x_j) := \mathbb{P}(X_j \leq x_j)$, for $j = 1, \dots, d$. Throughout the paper, vectors are written with bold letters, e.g., $\mathbf{x} = (x_1, \dots, x_d)^\top$ is a d-dimensional vector and $\mathbf{x}^\top$ denotes the transpose of $\mathbf{x}$. Inequalities $\mathbf{x} \leq \mathbf{y}$ are understood componentwise, i.e., $x_j \leq y_j$ for all $j = 1, \dots, d$. The characterization theorem of Sklar [52] implies that there exists a unique d-variate function $\mathbb{C}(\cdot)$, such that,

$$\mathbb{F}(\mathbf{x}) = \mathbb{C}(F_1(x_1), \dots, F_d(x_d)), \quad \text{for all } \mathbf{x} \in \mathbb{R}^d. \tag{1.1}$$

The function $\mathbb{C}(\cdot)$ is called a copula and it is in itself a joint cumulative distribution function on $[0, 1]^d$ with uniform marginals. Conversely, for any marginal distributions $F_1(\cdot), \dots, F_d(\cdot)$, and any copula function $\mathbb{C}(\cdot)$, the function $\mathbb{C}(F_1(x_1), \dots, F_d(x_d))$ is a multivariate distribution function with given marginal distributions $F_1(\cdot), \dots, F_d(\cdot)$. In the monographs by Nelsen [43] and Joe [37] the reader may find detailed ingredients of the modeling theory as well as surveys of the commonly used copulas. For in depth and overview historical notes, we refer the reader to [49]. Schweizer and Sklar [50, Chapter 6], provide a mathematical account of developments on copulas over three decades. We can refer also to [53], where the author has sketched the proof of (1.1), developed some of its consequences, and surveyed some of the work on copulas. Copulas have proved to be a flexible and versatile tool in the analysis of dependency structures. More specifically, copula $\mathbb{C}(\cdot)$ "couples" the joint distribution function $\mathbb{F}(\cdot)$ to its univariate marginals, capturing as such the dependence structure between the components of $\mathbf{X} = (X_1, \dots, X_d)^\top$. Indeed, most conventional measures of dependence can be explicitly expressed in terms of the copula. This feature has led to successful applications in actuarial science and survival analysis (see, e.g., [24,19]). In the literature on risk management and, more generally, in mathematical economics and mathematical finance modeling, a number of illustrations are provided (we refer the reader to books of [16,41]), in particular, in the context of asset pricing and credit risk management. Many useful multivariate models for dependence between $X_1, \dots, X_d$ turn out to be generated by *parametric* families of copulas of the form $\left\{\mathbb{C}(\cdot, \boldsymbol{\theta}); \boldsymbol{\theta} := (\theta_1, \dots, \theta_p)^\top \in \boldsymbol{\Theta}\right\}$, typically indexed by a vector valued parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$, for $p \geq 1$ (see, e.g., [38,39,36] among others). In this framework, semiparametric inference methods, based on *pseudo-likelihood*, proposed by Oakes [44], have been investigated by a number of authors (see, e.g., [26,51,57,13,54,3,6] and the references therein); we may refer also to [4,5,7] for semiparametric estimation based on ϕ-divergences, including as a particular case the semiparametric likelihood estimators. Although the idea of our testing approach follows that in [29], we allow for infinite-dimensional nuisance parameters (marginal distributions) in our proposed test. Hence, our tests are semiparametric, and the limiting distribution is distribution free. Gombay and Horváth [29] have established asymptotic distribution of the likelihood ratio (LR) statistic for parametric models. Their results are not directly applicable here since the two-step estimation of the semiparametric copula model depends on the estimates of unknown marginal distributions. These results are not only useful in their own right but essential to the construction of our semiparametric likelihood ratio (SLR) tests. Conventionally, the copula parameter is assumed to remain constant over time. However, empirical evidence suggests that the dependence structure is likely to vary due to some financial adjustments and critical social events. To cope with the change point, in the copula setting, for known marginals, [21] suggested a change point test for the copula parameter based on maximum likelihood estimation. Guégan and Zhang [32] analyzed two types of copula parameter changes and proposed a test based on generalized likelihood ratio assuming that the copula family is known a priori. However, neither of these authors have examined in detail the theoretical properties of the test. On the other hand, in [21], the problem has been considered in a parametric framework since the marginals have been assumed uniform. In the present work, in the derivation of our asymptotic results, we state explicitly the regularity conditions and we give a rigorous proof of asymptotic properties of the test both under the null and