Journal of Statistical Planning and Inference 142 (2012) 1506–1515

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

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

A test when the Fisher information may be infinite, exemplified by a test for marginal independence in extreme value distributions

Svetlana Litvinova, Mervyn J. Silvapulle *

Department of Econometrics and Business Statistics, Monash University, P.O. Box 197, Caulfield East, Melbourne 3145, Australia

ARTICLE INFO

Article history:
Received 13 May 2010
Received in revised form
22 May 2011
Accepted 15 December 2011
Available online 10 January 2012

Keywords:
Cramér–von Mises criterion
L^2 distance
Minimum distance
Nonstandard inference
Non-regular
Score test

ABSTRACT

For testing independence between the components of the multivariate logistic extreme value distribution, the Fisher Information is infinite and the score statistic converges only at the rate of $O(n^{-1/2}\log n)$. Consequently, the asymptotic optimality properties of the usual likelihood based methods do not necessarily carry over to this case. Motivated by this specific problem, this paper develops a general method based on the Cramér–von Mises discrepancy. This method does not require finite Fisher information. It is shown that this test statistic converges at the regular rate of $O(n^{-1/2})$. The test statistic has a closed form and it can be computed easily without using an iterative method or numerical integration. The asymptotic critical values can be obtained from a table for chi-square distribution. In a simulation study involving the test of marginal independence in bivariate extreme value distributions, the test proposed in this paper performed better than its main competing ones. While the method was motivated by, and developed for, a particular topic in inference for multivariate extreme value distributions, it is applicable more broadly than just for inference in multivariate extreme value distributions.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The standard likelihood based tests of a simple parametric hypothesis namely, the likelihood ratio, Wald and score tests, are asymptotically optimal when the Fisher information is finite and some other regularity conditions are satisfied. If the Fisher information turns out to be infinite, then there may well be other tests that are asymptotically more powerful than the foregoing likelihood based ones. and suitable tests are usually constructed on a case by case basis. In this paper, we consider the specific example of testing for independence of the components of multivariate extreme value distributions for which the Fisher information is infinite, and develop a test that is applicable more widely.

Probability models for multivariate extreme values play an important role in some studies to assess the risks based on multivariate observations. For example, one may be interested to study a probability model for the maximum movements of a bridge at several supporting points, or a model for the maxima of sea level and wind speed (Gupta and Manohar, 2005; Coles, 2001). For some such studies, the relevant probability models are based on multivariate extreme value [MEV] distributions. They arise as the limiting distribution of suitably standardized componentwise maxima of multivariate random variables. By restricting attention to the logistic parametric family of multivariate extreme value distributions,

* Corresponding author. Tel.: +61 3 9903 2559; fax: +61 3 9903 2007.
E-mail addresses: svetlana.litvinova@monash.edu (S. Litvinova), mervyn.silvapulle@monash.edu, Mervyn.Silvapulle@buseco.monash.edu.cn (M.J. Silvapulle).

0378-3758/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2011.12.031

Tawn (1988) pointed out that the likelihood based score statistic for testing independence between the components of a multivariate extreme value distribution suffers from some nonregular properties. In particular, the Fisher information is infinite and the score statistic has infinite variance. Consequently, the score statistic converges in distribution only at the rate of $O(n^{-1/2}\log n)$ instead of the regular parametric rate $O(n^{-1/2})$.

Recently, Ramos and Ledford (2005) (see also, Ledford and Tawn, 1996) proposed ways of modifying the score statistic studied by Tawn (1988), to attain the regular $O(n^{-1/2})$ rate of convergence. However, the more fundamental issue of infinite Fisher information for MEV distribution resulting in the inapplicability of standard likelihood based methods, without careful modifications on a case by case basis, highlights the need for an alternative method that does not require finite Fisher information and is applicable to the class of MEV distributions. The objective of this paper is to develop such a general method, and to apply it to improve upon the foregoing score and modified score statistics for testing independence in MEV distributions. A hypothesis testing problem involving infinite Fisher information was studied by Li et al. (2009). This was specifically for the case of mixtures of distributions and is not applicable to the type of problems studied in this paper.

For simplicity of exposition, in what follows, we shall consider only the bivariate case. Let $\boldsymbol{X}=(X_1,X_2)$ and let F denote its cumulative distribution function which we assume to be an extreme value distribution. Because we are interested in testing a hypothesis concerning the dependence function corresponding to F, we shall adopt the standard procedure of assuming, without loss of generality, that the marginal distribution of each component is unit Fréchet, given by $pr(X_j \le x_j)=\exp(-x_j^{-1})$, for $x_j>0$ and $j=1,2$ (see Tawn, 1988, p. 399). If X_1 and X_2 are independent then $F(x_1,x_2)=\exp\{-(x_1^{-1}+x_2^{-1})^{-1}\}$. Otherwise, $F(x_1,x_2)=\exp\{-V(x_1,x_2)\}$ for some function $V(x_1,x_2)$.

To simplify statistical modeling and inference on $F(x_1,x_2)$, different parametric families have been proposed for $V(x_1,x_2)$, for example, see Ledford and Tawn (1996). The logistic and mixed models are perhaps the most studied ones. In this paper also, we restrict to these families for exposition, but the results obtained in the following sections are applicable to much larger parametric families.

Consider the logistic model:

$$F(x_1,x_2;\theta)=\exp\{-(x_1^{-1/\theta}+x_2^{-1/\theta})^{\theta}\}, \quad 0<\theta\le 1,\ x_1>0,\ x_2>0. \tag{1}$$

Thus, marginal independence is equivalent to $\theta=1$. Let $\{(X_{i1},X_{i2}),i=1,\dots,n\}$ denote n independent and identically distributed random variables from the F in (1). Tawn (1988) showed that the score statistic for testing

$$H_0:\theta=1 \quad \text{against} \quad H_1:\theta<1 \tag{2}$$

has a term of the form $\sum X_{i1}X_{i2}/(X_{i1}+X_{i2})$, and the summand $X_{i1}X_{i2}/(X_{i1}+X_{i2})$ has infinite variance. Consequently, an unattractive feature of the score statistic is that it converges slower than the regular $n^{-1/2}$ rate.

To overcome the difficulty just mentioned, Ramos and Ledford (2005) (see also Ledford and Tawn, 1996) proposed to modify the score statistic by censoring the observations falling in the joint tail region $R_{11}\equiv\{(x_1,x_2): x_1\ge a_1, x_2\ge a_2\}$ for some suitably chosen threshold values a_1 and a_2. The resulting statistic converges at the regular $n^{-1/2}$ rate. In this sense, this is a regular version of the non-regular score test of Tawn (1988), and hence Ramos and Ledford (2005) refer to it as the *regularized score statistic*. It is natural to ask, whether or not it would be possible to improve upon the method of Ramos and Ledford (2005) by controlling the effect of observations in the extreme tail region more gradually than that by censoring? A contribution of this paper is to develop such a test.

An alternative approach to testing independence is to apply the Cramér–von Mises test (Genest et al., 2007; Genest and Rémillard, 2004; Genest et al., 2006; Deheuvels and Martynov, 1996). This is a nonparametric test, and hence has a different methodological basis. In a simulation study, the new test proposed in this paper performed better than the foregoing Cramér–von Mises test, and also better than the score test of Tawn (1988) and the regularized score test of Ramos and Ledford (2005).

The plan of the paper is the following. The testing method is developed in Section 2 and some simulation results are provided in Section 3 to compare several tests of independence. The proofs are relegated to Appendix.

2. The main results

Let $\boldsymbol{X}_i$ denote $(X_{i1},\dots,X_{id})^{\top}$ and $\{\boldsymbol{X}_i,\ i=1,\dots,n\}$ be n independent and identically distributed d-dimensional random vectors with a common distribution function $F_0(\boldsymbol{x})$. Let $F_n(\boldsymbol{x})=n^{-1}\sum I(X_{i1}\le x_1,\dots,X_{id}\le x_d)$, the empirical distribution function of $\{\boldsymbol{X}_i,\ i=1,\dots,n\}$. Let $F(\boldsymbol{x};\boldsymbol{\theta})$, $\boldsymbol{\theta}\in\Theta\subset\mathbb{R}^m$ be a family of cumulative distribution functions.

Suppose that $F_0(\cdot)$ belongs to the parametric family $F(\cdot;\boldsymbol{\theta})$. Let $\boldsymbol{\theta}_T$ in Θ denote the true population value of $\boldsymbol{\theta}$. Thus, $F_0(\cdot)=F(\cdot;\boldsymbol{\theta}_T)$. We wish to test

$$H_0:\boldsymbol{\theta}_T=\boldsymbol{\theta}_0 \quad \text{vs} \quad H_1:\boldsymbol{\theta}_T\ne\boldsymbol{\theta}_0, \tag{3}$$

where $\boldsymbol{\theta}_0$ is a given point in Θ. Suppose that the Fisher information of $F(\cdot;\boldsymbol{\theta})$ at $\boldsymbol{\theta}_0$ may be infinite, and that $\boldsymbol{\theta}_0$ may be a boundary point of Θ. Therefore, the usual likelihood ratio, score and Wald tests may not be asymptotically optimal, and as noted previously, these test statistics may converge slower than the regular $n^{-1/2}$ rate.

The starting point for the method developed in this paper is the well-known Cramér–von Mises criterion $\omega^2=\int\{F_n(x)-F_0(x)\}^2 w\{F_0(x)\}dF_0(x)$, which is a measure of discrepancy between F_n and F_0 (for example, see Durbin and Knott, 1972; Deheuvels and Martynov, 1996; Nikulin, 1999; Escanciano and Jacho-Chávez, 2010). This and similar

Download English Version:

https://daneshyari.com/en/article/1147951

Download Persian Version:

https://daneshyari.com/article/1147951

Daneshyari.com