ARTICLE IN PRESS

Journal of Econometrics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Econometrics

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

# Testing conditional independence via empirical likelihood☆

Liangjun Su[a,*], Halbert White[b]

[a] School of Economics, Singapore Management University, Singapore

[b] Department of Economics, UCSD, La Jolla, CA 92093-0508, United States

ARTICLE INFO

*Article history:*
Available online xxxx

*JEL classification:*
C12
C14
C22

*Keywords:*
Conditional independence
Empirical likelihood
Granger causality
Local smoothed bootstrap
Nonlinear dependence
Nonparametric regression
U-statistics

ABSTRACT

We construct two classes of smoothed empirical likelihood ratio tests for the conditional independence hypothesis by writing the null hypothesis as an infinite collection of conditional moment restrictions indexed by a nuisance parameter. One class is based on the CDF; another is based on smoother functions. We show that the test statistics are asymptotically normal under the null hypothesis and a sequence of Pitman local alternatives. We also show that the tests possess an asymptotic optimality property in terms of average power. Simulations suggest that the tests are well behaved in finite samples. Applications to some economic and financial time series indicate that our tests reveal some interesting nonlinear causal relations which the traditional linear Granger causality test fails to detect.



## 1. Introduction

Recently there has been a growing interest in testing the conditional independence (CI) of two random vectors $Y$ and $Z$ given a third random vector $X : Y \perp Z|X$. Linton and Gozalo (1997) propose two nonparametric tests of CI for independent and identically distributed (IID) variables based on generalized empirical distribution functions. Fernandes and Flores (1999) employ a generalized entropy measure to test CI but rely heavily on the choice of suitable weighting functions to avoid distributional degeneracy. Delgado and González-Manteiga (2001) propose an omnibus test of CI using the weighted difference of the estimated conditional distributions under the null and the alternative. Su and White (2007, 2008) consider testing CI based on conditional-characteristic-function-based moment conditions and the Hellinger distance between two conditional densities, respectively. Amaro de Matos and Fernandes (2007) and Chen and Hong (2012) propose nonparametric tests for the Markov property (a special case of CI) based on the comparison of densities and generalized cross spectrums, respectively. Song (2009) studies an asymptotically pivotal test of CI via the probability integral transform. Huang (2010) proposes a test of CI based on the estimation of the maximal nonlinear conditional correlation. Bergsma (2011) proposes a test for CI by means of the partial copula. Corradi et al. (2012) consider testing CI between two volatility processes. Bouezmarni et al. (2012) and Bouezmarni and Taamouti (2012) propose tests for CI by comparing Bernstein copulas using the Hellinger distance and conditional distributions using the $L_2$-distance, respectively. Su and Spindler (2013) consider testing for asymmetric information (a special case of conditional dependence) by comparing conditional distributions with both continuous and discrete variables. Huang et al. (2013) develop a flexible test for CI based on the generically comprehensively revealing functions of Stinchcombe and White (1998).[1]

In this paper, we propose two new classes of tests for CI based on empirical likelihood (EL). The motivation is as follows. First, the equality of two conditional distributions can be expressed in terms of an infinite sequence of conditional moment restrictions. Second, there are many powerful tests available in the literature

☆ The authors sincerely thank the guest editors and anonymous referees for their many insightful comments and suggestions that lead to a substantial improvement of the presentation. They also thank Qihui Chen for his outstanding research assistance. Su gratefully acknowledges the Singapore Ministry of Education for Academic Research Fund under grant number MOE2012-T2-2-021.

* Correspondence to: School of Economics, SMU, 90 Stamford Road, Singapore 178903, Singapore. Tel.: +65 6828 0386.

*E-mail address:* ljsu@smu.edu.sg (L. Su).

[1] Most of the aforementioned papers came out after the first version of this paper in 2003. Note that for categorical data, the literature traces back to Rosenbaum (1984) and Yao and Tritchler (1993).

http://dx.doi.org/10.1016/j.jeconom.2014.04.006
0304-4076/© 2014 Elsevier B.V. All rights reserved.

to test for conditional moment restrictions, including EL-based tests. Third, EL has been shown to share some key properties with parametric likelihood such as Wilks' theorem and Bartlett correctability. Owen (1988, 1990, 1991) studies inference based on the nonparametric likelihood ratio, which is particularly useful in testing moment restrictions. Kitamura (2001) investigates the asymptotic efficiency of moment restriction tests for a finite number of unconditional moments in terms of large deviations and demonstrates the optimality of EL for testing such unconditional moment restrictions. Tripathi and Kitamura (2003, **TK** hereafter) and Kitamura et al. (2004) extend the EL paradigm to test for a finite number of conditional moment restrictions and show that their test possesses an optimality property in large samples and behaves well in small samples. As yet, it remains unknown whether one can extend EL methods to test for an infinite collection of conditional moment restrictions, and, if so, whether the test continues to possess some optimality property and behaves reasonably well in finite samples. These issues are the focus of this paper.

The contributions of this paper lie primarily in four directions. First, we show that a smoothed empirical likelihood ratio (SELR) can be used to test hypotheses that can be expressed in terms of an infinite collection of conditional moment restrictions, indexed by a nuisance parameter, $\tau$, say. Corresponding to each $\tau$, one can construct a SELR. Then one obtains a test statistic by integrating $\tau$ out. After being appropriately centered and rescaled, the resulting test statistic is shown to be asymptotically distributed as $N(0, 1)$ under the null. Second, we study the asymptotic distribution of the test statistic under a sequence of local alternatives and show that our test is asymptotically optimal in that it attains the maximum average local power with respect to a certain space of functions for the local alternatives. Third, unlike most work in the EL literature, including that of **TK**, our tests allow for data dependence and thus are applicable to time series data.[2] Fourth, our paper offers a convenient approach to testing distributional hypotheses via an infinite collection of conditional moment restrictions. It further extends the applicability of the EL method. A variety of interesting and important hypotheses other than CI in economics and finance, including conditional goodness-of-fit, conditional homogeneity, conditional quantile restrictions, and conditional symmetry, can also be studied using our approach.

It is well known that distributional Granger non-causality (Granger, 1980) is a particular case of CI. Our tests can be directly applied to test for Granger non-causality with no need to specify a particular linear or nonlinear model. Using the same techniques as in Su and White (2008), it is also easy to show that our tests can be applied to the situation where not all variables of interest are continuously valued and some have to be estimated from the data. In particular, our tests apply to situations where limited dependent variables or discrete conditioning variables are involved, and to parametrically or nonparametrically generated regressors/residuals. For brevity, however, we only focus on the case where all random vectors are observed and continuously valued.

The remainder of this paper is organized as follows. In Section 2, we treat a simple version of our tests based on CDF's in order to lay out the basic framework for our SELR tests for CI. In Section 3, we study the asymptotic distributions of the test statistics under both the null hypothesis and a sequence of local alternatives, and show the asymptotic optimality of our tests in terms of average local power. We discuss a version of our SELR tests based on smoother moment conditions that has better finite sample power properties in Section 4. We examine the finite sample performance of our smoother SELR test via Monte Carlo simulations in Section 5, and we apply it to some macroeconomic and financial time series data in Section 6. Final remarks are contained in Section 7. All technical details are relegated to the Appendix.

## 2. Test statistic based on the CDFs

In this paper, we are interested in testing whether $Y$ and $Z$ are independent conditional on $X$, where $X$, $Y$ and $Z$ are vectors of dimension $d_1$, $d_2$ and $d_3$, respectively. The data consist of $n$ identically distributed but weakly dependent observations $\{X_t, Y_t, Z_t\}_{t=1}^n$. For notational simplicity, we assume that $d_2 = 1$ throughout the paper.

### 2.1. Hypotheses

Let $f(x, y, z)$ and $F(x, y, z)$ denote the joint probability density function (PDF) and cumulative distribution function (CDF) of $(X_t, Y_t, Z_t)$, respectively. Below we make reference to several marginal densities of $f$ which we denote simply using the list of their arguments — for example $f(x, y) = \int f(x, y, z)dz$, $f(x, z) = \int f(x, y, z)dy$, and $f(x) = \int\int f(x, y, z)dydz$. This notation is compact, and, we hope, sufficiently unambiguous. Let $f(\cdot|\cdot)$ denote the conditional density of one random vector given another. We assume that $f(y|x, z)$ is smooth in $(x, z)$. Let $1(\cdot)$ be the usual indicator function, $F(\tau|x, z) \equiv E[1(Y_t \le \tau)|X_t = x, Z_t = z]$ and $F(\tau|x) \equiv E[1(Y_t \le \tau)|X_t = x]$. The null of interest is that conditional on $X$, the random vectors $Y$ and $Z$ are independent, i.e.,

$$\mathbb{H}_0: \ \Pr[F(\tau|X_t, Z_t) = F(\tau|X_t)] = 1 \quad \text{for all } \tau \in \mathbb{R}. \tag{2.1}$$

The alternative hypothesis is that for $\tau$ with a nontrivial volume of the support of $Y_t$,

$$\mathbb{H}_1: \ \Pr[F(\tau|X_t, Z_t) = F(\tau|X_t)] < 1. \tag{2.2}$$

In Section 4, we consider another approach based on a related condition involving the characteristic function. We treat $\mathbb{H}_0$ first because of its intuitive appeal.

### 2.2. Test statistics

Noting that $\mathbb{H}_0$ specifies an infinite collection of conditional moment restrictions that are indexed by $\tau$: $E[\varepsilon_t(\tau)|X_t, Z_t] = 0$ a.s. for all $\tau \in \mathbb{R}$ where $\varepsilon_t(\tau) = 1(Y_t \le \tau) - F(\tau|X_t)$, we can test $\mathbb{H}_0$ by testing a single conditional moment restriction for given $\tau$,

$$\mathbb{H}_0(\tau): \ \Pr[F(\tau|X_t, Z_t) = F(\tau|X_t)] = 1, \tag{2.3}$$

based on the EL principle, and obtain the final test statistic by integrating out $\tau$. Due to the use of integration, it is computationally expensive to calculate the statistic. Therefore we also consider a weaker version of the CI hypothesis:

$$\mathbb{H}_0': \ \Pr[F(Y_t|X_t, Z_t) = F(Y_t|X_t)] = 1, \tag{2.4}$$

which is implied by (2.1). Su and Spindler (2013) and Bouezmarni and Taamouti (2012) independently propose a $L_2$-distance-based test for (2.4) by comparing the weighted difference between the nonparametric kernel estimates of $F(Y_t|X_t, Z_t)$ and $F(Y_t|X_t)$.

To proceed, we first consider a SELR test statistic for $\mathbb{H}_0(\tau)$. Let $p_{ts} \equiv p_{(Y_s;X_t,Z_t)}$ denote the probability mass placed at $(Y_s; X_t, Z_t)$ by a discrete distribution with support $\{Y_s\}_{s=1}^n \times \{(X_t, Z_t)\}_{t=1}^n$. Let $\hat{\varepsilon}_s(\tau) \equiv 1(Y_s \le \tau) - \hat{F}_{h_2}(\tau|X_s)$, where

$$\hat{F}_{h_2}(\tau|x) \equiv n^{-1}\sum_{t=1}^{n} L_{h_2}(x - X_t)1(Y_t \le \tau)/\hat{f}_{h_2}(x), \tag{2.5}$$

[2] Chen et al. (2003) consider an EL goodness-of-fit test for time series. They integrate out the conditioning variable and employ only finite-dimensional parameter estimates in the constraints. This is quite different from our approach.