Nonlinear Dynamics (2006) 44: 351–357
DOI: 10.1007/s11071-006-2019-0

Entropy-Based Independence Test

ANDREIA DIONÍSIO[1,*], RUI MENEZES[2], and DIANA A. MENDES[2]
[1]*Centre of Business Studies (CEFAGE-UE) Management Department, University of Évora, Largo dos Colegiais 2, 7000 Évora, Portugal;* [2]*ISCTE, Av. Forcas Armadas, 1649-026 Lisbo, Portugal;* *Author for correspondence (e-mail: andreia@uevora.pt; fax: +00351-266-742-494)*

(Received: 20 September 2004; accepted: 9 June 2005)

Abstract. This paper presents a new test of independence (linear and non-linear) among distributions based on the entropy of Shannon. The main advantages of the presented approach are the fact that this measure does not need to assume any type of theoretical probability distribution and has the ability to capture the linear and non-linear dependencies, without requiring the specification of any kind of dependence model.

Key words: entropy, independence test, mutual information, non-linear serial dependence, stock index markets

1. Introduction

The notion of "independence", "distance" and "divergence" between distributions has been central in statistical inference and econometrics from the earliest stages. This is also evident in the work of Kullback and Leibler, H. Jeffreys, H. Akaike, E. Shannon, Hartley, A. Renyi, E. Maasoumi, C. Granger, H. White, A. Zellner and many others (see e.g. [1]). Some authors, such as Cover et al. [2] and Maasoumi [1], moved by the "elegance" and the potential power of information theory, brought a new way of interpretation and motivation for the research in statistical inference. In addition, the axiomatic systems in information theory suggest principles of decomposition that distinguish between different information functions and "entropies", and identify desirable measures, decision criteria and indices [1].

Several measures have been used as independence tests and/or dependence measures in this field. The most known measure of dependence between random variables is the Pearson correlation coefficient. However, this is nothing but a normalized covariance and only accounts for linear (or linearly transformed) relationships (see e.g. [3, 4]). In general, this statistic may not be helpful to capture serial dependence when there are non-linearities in the data. In this context, it seems that a measure of global dependence is required, that is, a measure that captures both linear and non-linear dependencies without requiring the specification of any kind of model of dependence. Urbach [5] defends a strong relationship between entropy, dependence and predictability. This relation has been studied by several authors, namely Granger and Lin [3], Maasoumi and Racine [4], Darbellay and Wuertz [6]. On the basis of the arguments mentioned earlier, we aim to evaluate in this paper the efficiency of a new entropy-based independence test without requiring the specification of mean-variance models and theoretical distribution probabilities. Thus, in the next section we discuss the subject of information and predictability in the context of entropy, and we then illustrate our test using evidence based on empirical financial data.