International Journal of Approximate Reasoning 65 (2015) 11–23

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

A measure of mutual complete dependence in discrete variables through subcopula

Qingsong Shan [a,b], Tanes Wongyang [c], Tonghui Wang [a,b,*], Santi Tasena [c]

[a] College of Sciences, Northwest Agriculture and Forest University, China
[b] Department of Mathematical Sciences, New Mexico State University, USA
[c] Department of Mathematics, Faculty of Sciences, Chiang Mai University, Thailand

ARTICLE INFO

Article history:
Received 2 February 2015
Received in revised form 3 April 2015
Accepted 11 April 2015
Available online 22 April 2015

Keywords:
Mutual complete dependence
Subcopula
Functional relationship
Dependence measure
Bilinear extension

ABSTRACT

Siburg and Stoimenov [12] gave a measure of mutual complete dependence of continuous variables which is different from Spearman's ρ and Kendall's τ. In this paper, a similar measure of mutual complete dependence is applied to discrete variables. Also two measures for functional relationships, which are not bijection, are investigated. For illustration of our main results, several examples are given.

1. Introduction

It is well known that there are several types of associations between random variables, such as Pearson's correlation coefficient, Spearman's ρ and Kendall's τ. However, they are not suitable for nonlinear cases since Pearson's correlation coefficient only measure linear relationship of random variables and Spearman's ρ and Kendall's τ are measures of monotonic relationships. In order to investigate the nonlinear relationships between two random variables, their joint distributions are often used. Because joint distributions contain not only the information of association between random variables, but also their marginal distributions. Since we are only investigating the associations or functional relationships between variables, which are free of their marginals, the use of copulas or subcopulas is needed. The importance of our measures of the functional relationships is that it can be suitable for both linear and nonlinear relationships.

According to Sklar's Theorem [13], for any two random variables X and Y with joint and marginal distribution functions $H_{X,Y}$, F_X and G_Y respectively, there exists a unique function $C : Range(F_X) \times Range(G_Y) \to [0, 1]$ corresponding to $H_{X,Y}$, that is, $H_{X,Y}(x, y) = C(F_X(x), G_Y(y))$. This function is called a **subcopula**. Subcopulas corresponding to continuous random variables are known as **copulas**. Since copulas describe the relationships among random variables, several measures of dependence among random variables have been constructed through copulas. Definitions for multivariate measures of concordance are given in Dolati and Úbeda-Flores [3] and Taylor [15]. The popular dependence measures such as Spearman's ρ, Kendall's τ and Gini's γ are measures of concordance. For references, see Joe [5], Nelsen [7], Schmid and Schmidt [11], Taylor [15], and Wolff [16]. Also details of dependence measures are discussed in Lecture Notes of Nguyen [10].

Besides concordance, there is another relationship called mutual complete dependence, see Lancaster [6].

* Corresponding author at: Department of Mathematical Sciences, New Mexico State University, USA. Tel.: +1 (575)646 2507; fax: +1 (575)646 1064.
E-mail addresses: qingsong@nmsu.edu (Q. Shan), t.wongyang@gmail.com (T. Wongyang), twang@nmsu.edu (T. Wang), santi.tasena@cmu.ac.th (S. Tasena).

http://dx.doi.org/10.1016/j.ijar.2015.04.005

Definition 1.1. Let X and Y be two random variables defined on the same probability space. Then Y is **completely dependent** on X if Y is a measurable function of X. Precisely, Y is completely dependent on X if there is a function φ such that

$$P(Y = \varphi(X)) = 1.$$

The random variables X and Y are said to be **mutually completely dependent (MCD)** if Y is completely dependent on X and X is completely dependent on Y.

Note that Spearman's ρ and Kendall's τ are measures for concordance, while MCD is a measure of functional relationships between X and Y. For the case where X and Y are continuous, the measure of MCD through copula is given in Siburg and Stoimenov [12]. Tasena and Dhompongsa [14] extended this measure to higher dimensions. But this measure is not suitable for discrete variables due to the fact that derivative of subcopula does not make sense. Thus a new measure of MCD is needed for discrete variables.

In this paper, similar measures for MCD of discrete random variables through both subcopula and conditional distributions are defined and their properties are discussed. These measures are equal to 1 if and only if X and Y are MCD (i.e., Y is a bijective function of X). For non-bijective functional relations between X and Y, two related measures are obtained.

The paper is organized as follows. Basic properties of copulas related to MCD are discussed in Section 2. Measures for MCD and functional relationship of discrete random variables are given in Section 3. The MCD measure under E-process extension and the bilinear interpolation of a subcopula is discussed in Section 4. Relationship among different types of dependence measures are discussed in Section 5. For illustration of our results, several examples are given.

2. Preliminaries

First, we give the definitions of subcopula and copula given in Nelsen [8].

Definition 2.1. A two-dimensional **subcopula** (or **2-subcopula**) is a function C with the following properties:

(i) $D(C) = D_1 \times D_2$, where $D(C)$ is the domain of C, D_1 and D_2 are subsets of $I = [0, 1]$ containing 0 and 1;
(ii) For every u in D_1 and every v in D_2,

$$C(u, 0) = 0 = C(0, v) \quad \text{and} \quad C(u, 1) = u \quad C(1, v) = v;$$

(iii) For every u_1, u_2 in D_1, v_1, v_2 in D_2 such that $u_1 \le u_2$ and $v_1 \le v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ge 0.$$

Subcopulas for which $D_1 = D_2 = [0, 1]$ are called **copulas**.

The following result of Nelsen [8] will be used in the proofs of our main results.

Lemma 2.1. *Let $C \in \mathcal{C}$ be a subcopula, then*

$$\left|C(u_j, v_j) - C(u_i, v_i)\right| \le \left|u_j - u_i\right| + \left|v_j - v_i\right|$$

for any (u_i, v_i), (u_j, v_j) in $D(C)$.

From Lemma 2.1, it's easy to see

$$\frac{C(u_j, v_i) - C(u_i, v_i)}{u_j - u_i} \le 1$$

and

$$\frac{C(u_i, v_j) - C(u_i, v_i)}{v_j - v_i} \le 1,$$

where $i < j$.

Remark 2.1. The modified Sobolev norm of a copula C, introduced by Darsow and Olsen [1], is given by

$$\|C\| = \left(\int_{I^2} |\nabla C|^2 \, du \, dv \right)^{1/2},$$