Revista Colombiana de Estadística
January 2020, Volume 43, Issue 1, pp. 3 to 20
DOI: http://dx.doi.org/10.15446/rce.v43n1.78054

Relationship Between Kendall's tau Correlation and Mutual Information

Relación entre la correlación tau de Kendall e información mutua

Mohammad Bolbolian Ghalibaf[a]

Department of Statistics, Faculty of Mathematics and Computer Science, Hakim Sabzevari Univercity, Sabzevar, Iran

Abstract

Mutual information (MI) can be viewed as a measure of multivariate association in a random vector. However, the estimation of MI is difficult since the estimation of the joint probability density function (PDF) of non-Gaussian distributed data is a hard problem. Copula function is an appropriate tool for estimating MI since the joint probability density function of random variables can be expressed as the product of the associated copula density function and marginal PDF's. With a little search, we find that the proposed copulas-based mutual information is much more accurate than conventional methods such as the joint histogram and Parzen window-based MI. In this paper, by using the copulas-based method, we compute MI for some family of bivariate distribution functions and study the relationship between Kendall's tau correlation and MI of bivariate distributions. Finally, using a real dataset, we illustrate the efficiency of this approach.

***Key words*:** Copula function; Kendall's tau correlation; Mutual information.

Resumen

La información mutua (MI) puede ser vista como una medida de asociación multivariante en un vector aleatorio. Sin embargo, la estimación de MI es difícil ya que la estimación de la función de densidad de probabilidad conjunta (PDF) de datos distribuidos no gaussianos es un problema difícil. La función copula es una herramienta apropiada para estimar el MI ya que la función de densidad de probabilidad de las variables aleatorias se puede expresar como el producto de la función de densidad de cópula asociada y de los PDF marginales. Con una pequeña búsqueda, encontramos que la información mutua propuesta basada en cópulas es mucho más precisa que

[a]PhD. E-mail: m.bolbolian@hsu.ac.ir, m.bolbolian@gmail.com

los métodos convencionales, como el histograma de la articulación y el MI basado en ventana de Parzen. En este artículo, al utilizar el método basado en cópulas, calculamos el MI para algunas familias de funciones de distribución bivariadas y estudiamos la relación entre la correlación tau de Kendall y el MI de las distribuciones bivariadas. Finalmente, usando un conjunto de datos real, ilustramos la eficiencia de este enfoque.

***Palabras clave*:** Función de cópula; Correlación tau de Kendall; Información mutua.

1. Introduction

One way of determining the measure of dependence between two random variables is using the information theory. Some measures such as entropy, mutual information, and quadratic mutual information play an important role in dependence measuring of bivariate distributions and some papers have written in this subject. The mutual information (also known as Kullback-Leibler divergence) is a general measure of the dependence between two random variables. It expresses the quantity of information one has obtained on X by observing Y. For two discrete variables, the MI between them is given by Shannon & Weaver (1949). The MI for random vector with underlying multivariate Gaussian distribution is given by Kullback (1952), Kullback (1959). Bell (1962) has used MI as a measure of dependence and Joe (1989) has presented the relative entropy measures of multivariate dependence. MI was calculated for the multivariate t distribution by Guerrero-Cusumano (1996*a*), Guerrero-Cusumano (1996*b*) and Mercier (2005) have arrived at the MI in Cuadras-Auge family of distributions. Also, Arellano-Valle, Contreras-Reyes & Genton (2013) represent MI for the full symmetric class of multivariate elliptical distributions and then extend it to the more flexible families of multivariate skew-elliptical distributions.

MI has been applied widely in signal processing such as image registration, and feature selection. However, an efficient method to estimate MI accurately is necessary. As an instance, in digital image processing, three approaches: specific multivariate distribution assumption such as multivariate Gaussian distribution, the joint histogram (Maes, Collignon, Vandermeulen, Marchal & Suetens 1997) and the Parzen window (Kwak & Choi 2002) are usually used to estimate MI. Multivariate distribution models such as the multivariate Gaussian are often employed. Note, however, that the distributions of the image pixel intensities in the real world may not obey the Gaussian or other certain probability distributions. Furthermore, known models of multivariate distributions require that the associated marginal distributions are consistent. However, the marginal distributions usually are arbitrary. The joint histogram method computes the normalized joint histograms of pixel intensities for the overlapping parts of two images, where the joint histogram counts the number of occurrences of pixel pairs. The number of bins is difficult to confirm, a smaller or greater bin number than the optimal number of bins is known to yield over-smoothed density and highly sparse density, respectively. As for the Parzen window method, the selections of kernel function

and parameter for kernel function are difficult to confirm. Furthermore, both the joint histogram and the Parzen window cannot estimate the continuous form of the joint probability density function. Zeng & Durrani (2011) introduce a novel method using copula density function to estimate MI with the continuous form. By using copulas-based mutual information Blumentritt & Schmid (2012) estimate MI in Frank copula and Clayton copula by Monte Carlo simulations and describe the estimators for MI. Kumar (2012) and Dobrowolski & Kumar (2014) compute MI in two-parameter Marshall-Olkin family of copulas.

In this paper, we estimate MI for some family of bivariate distribution functions by using copulas-based mutual information, but for calculating we use numerical integration. In Section 2, we represent copulas-based mutual information. MI for some family of bivariate distribution functions compute in Section 3, also we study the relationship between Kendall's tau correlation and MI of bivariate distributions. In Section 4, we compare the MI from various copulas mentioned in Section 3 numerically and graphically. We examine the efficiency of this approach in Section 5 with a real dataset. Finally, Section 6 concludes done. For these purposes, we use the R software (R Development Core Team 2012) and for numerical integration, we apply the R package "cubature".

2. Copulas-Based Mutual Information

MI can be viewed as a measure of multivariate association in a random vector. This measure is directly related to Shannon's entropy. Assume that X and Y are input and output respectively of a stochastic system, then Shannon's entropy $H[X]$ represents the uncertainty of input X before output Y is observed while conditional entropy $H[X|Y]$ is the uncertainty of input X after output Y has been realized. The quantity is called mutual information (distance from independence) between X and Y.

$$\begin{aligned} MI(X,Y) &= H[X] - H[X|Y] \\ &= H[X] + H[Y] - H[X,Y] \\ &= H[X,Y] - H[X|Y] - H[Y|X], \end{aligned}$$

Thus MI measures the decrease in uncertainty of X caused by the knowledge of Y which is the same as the decrease in uncertainty of Y caused by the knowledge of X. The measure $MI(X,Y)$ indicates the amount of information of X contained in Y or the amount of information of Y contained in X. Obviously $MI(X,X) = H[X]$, the amount of information contained in X about itself. MI is the recommended measure in Kinney & Atwal (2014). The conventional MI for continuous variables has been defined as:

$$MI(X,Y) = \int\int_{X,Y} f_{XY}(x,y)\, log \frac{f_{XY}(x,y)}{f_X(x) f_Y(y)} dxdy, \tag{1}$$

where $f_{XY}(x,y)$ is the joint PDF, and $f_X(x)$ and $f_Y(y)$ are the marginal PDF's of variables X and Y, respectively.

The units of information depend on the base of the logarithm. If base 2, e or 10 is used, information is measured in bits, nats or bans respectively. We use base e in this paper. For more details, see Table 1.

Table 1: The units of information.

Base	Units	Conversion
2	bits	$1bit = 1bit$
e	nats	$1bit = log_e 2 (\approx 0.693) nats$
10	bans	$1bit = log_{10} 2 (\approx 0.301) bans$

MI is always greater than or equal to zero, with equality iff X and Y are independent; the higher the MI obtained the stronger the dependency between X and Y. It is lower than the entropy of both variable, and equality only occurs iff one variable is a deterministic function of the other:

$$0 \leq MI(X,Y) \leq \min(H[X], H[Y]). \tag{2}$$

Copulas provide a useful way to model different types of dependence structures explicitly. Instead of having one correlation number that encapsulates everything known about the dependence between two variable, copulas capture information on the level of dependence as well as whether the two variables exhibit other types of dependence. In particular, copula density functions offer a natural way to estimate MI instead of the joint and marginal probability density functions, marginal and joint entropy.

Let (X, Y) be a random vector with density function $f_{XY}(x, y)$, distribution function $F_{XY}(x, y)$ and marginals $F_X(x)$ and $F_Y(y)$. The copula function $C(u, v)$ is a bivariate distribution function with uniform marginals on $[0, 1]$, such that

$$F_{XY}(x, y) = C_F(F_X(x), F_Y(y)).$$

By Sklar's Theorem (Sklar 1959), this copula exists and is unique if F_X and F_Y are continuous. Also, the copula C_F is given by

$$C(u, v) = F(F_X^{-1}(u), F_Y^{-1}(v)), \quad \forall u, v \in [0, 1],$$

where, F_X^{-1} and F_Y^{-1} are quasi-inverses of F_X and F_Y respectively (Nelsen 2006). The partial derivatives $\frac{\partial C(u,v)}{\partial u}$ and $\frac{\partial C(u,v)}{\partial v}$ exist and density function of $C(u, v)$ is defined as:

$$\begin{aligned} c(u, v) &= \frac{\partial^2 C(u, v)}{\partial u \partial v} = \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial F_X(x) \partial F_Y(y)} \\ &= \frac{\partial^2 F_{XY}(x, y)}{f_X(x) f_Y(y) \partial x \partial y} = \frac{f_{XY}(x, y)}{f_X(x) f_Y(y)}. \end{aligned}$$

Therefore by substituting in Equation (1), copulas-based mutual information can be defined as:

$$MI(X,Y) = \int \int_{[0,1]^2} c(u, v) \ln c(u, v) du dv = E_C[\ln c(U, V)],$$

where $c(u, v)$ is copula density function of the copula $C(u, v)$ and E_C denotes the expectation with respect to the Lebesgue-Stieltjes measure P^C induced by C, (Zeng & Durrani 2011). It should be noted this result has been obtained in Jenison & Reale (2004) earlier for calculating entropy. Note that the copulas-based mutual information only relies on the copula density function that is determined by the copula parameter, and therefore only the copula parameter is required for estimation of MI. According to (2), $MI(X, Y)$ is not necessarily limited to $[0, 1]$, Joe (1989) defined

$$\delta = \sqrt{1 - \exp(-2MI)},$$

which is normalizing this index. The measure of δ is confined to the interval $[0, 1]$. If X and Y are independent then $\delta = 0$ and when the dependence is maximal, δ achieves to one.

Kendall's tau is another measure of concordance between two variables; this measure has been introduced by Kendall (1938). Let (X_1, X_2) and (Y_1, Y_2) be independent and identically distributed random vectors with distribution function F. In the bivariate case, the population version of Kendall's tau is defined as the probability of concordance minus the probability of discordance:

$$\tau(X_1, X_2) = P\{(X_1 - Y_1)(X_2 - Y_2) > 0\} - P\{(X_1 - Y_1)(X_2 - Y_2) < 0\}.$$

***Note* 1.** Kendall's tau associated by copula function $C(u, v)$ as follows:

$$\begin{aligned}\tau &= 1 - 4\int_0^1 \int_0^1 \left[\frac{\partial C}{\partial u}.\frac{\partial C}{\partial v}\right] dudv \\ &= 4\int_0^1 \int_0^1 C(u, v)dC(u, v) - 1 = 4E_C[C(U, V)] - 1,\end{aligned}$$

where E_C denotes the expectation with respect to the copula function C (Nelsen 2006). For (bivariate) Archimedean copulas, Kendall's tau can directly be calculated from the generator $\phi_C(t)$ of the copula through

$$\tau = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)}dt,$$

for more details see Genest & MacKay (1986*a*) and Genest & MacKay (1986*b*).

3. Mutual Information in Some Family of Bivariate Distribution

In this section, we compute the MI for some family of bivariate distribution. We consider the copula functions: Cuadras-Auge, Clayton, Frank, Gumbel, Raftery, Gaussian, and T-copulas. We consider these copulas because these have positively ordered and perfect dependency bound since its corresponding Kendall's tau correlation locates between $[0, 1]$. Meanwhile, the bound of Kendall's tau correlation

for some copulas may be limited, such as the FGM copula (Nelsen 2006) is limited on $[-2/9, 2/9]$.

It should be noted that the explicit form of MI for Cuadras-Auge copula, Gaussian copula, and T-copulas already been achieved and we obtained the MI of them using those formulas. Also the estimating of MI for Frank copula and Clayton copula has done by Monte Carlo simulations. However, we estimate MI of Frank and Clayton copulas again by using numerical integration and we calculate MI of Gumbel and Raftery copulas by using numerical integration.

3.1. Cuadras-Auge Copula

Cuadras & Auge (1981) have defined the copula

$$C(u,v) = [\min(u,v)]^{\theta}(uv)^{1-\theta}, \quad \theta \in [0,1].$$

Cuadras-Auge (C-A) family of bivariate distributions is obtained by considering a weighted geometric mean of the independence distribution and the upper Frechet-Hoeffding bounds. C-A family is a subfamily of the Marshall-Olkin family of copulas. C-A copula has a singular part in the diagonal $u = v$. The copula density function of this family is given by

$$c(u,v) = (1-\theta)[\max\{u,v\}]^{-\theta} + \theta u^{1-\theta} I_{\{u=v\}},$$

where $I_{\{u=v\}} = 1$ if $u = v$ and 0 otherwise, is the indicator function.

***Note* 2.** By using Note 1 it is easy to show that Kendall's tau in C-A family is given by

$$\tau = \frac{\theta}{2-\theta}. \tag{3}$$

For the first time, Mercier (2005) obtains MI of C-A copula with respect to dependence parameter θ and Kumar (2012) offers a simpler formula for computing MI.

Proposition 1. *Let (X, Y) be a random vector with C-A copula, then MI is given by*

$$MI = -\frac{2(1-\theta)}{2-\theta}\left[\log(1-\theta) + \frac{\theta}{2-\theta}\right]. \tag{4}$$

***Proof*.** See page 21 from Mercier (2005) and Section 4 from Kumar (2012). □

***Note* 3.** By substituting Equation (3) in Equation (4), we conclude that

$$MI = -(1-\tau)\left[\tau + \log(\frac{1-\tau}{1+\tau})\right]. \tag{5}$$

We compute the MI of C-A family for various values of τ, i.e. θ was chosen according to the values of Kendall's τ given in the first row. Results display in Table 2. In Figure 1, by using Equation (5), we depict the behavior of the MI versus the Kendall's τ.

TABLE 2: The values of MI and normalized MI with respect to the amounts of Kendall's τ for C-A copula.

τ	0	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	1
θ	0	0.182	0.333	0.400	0.461	0.571	0.667	0.750	0.823	0.857	0.889	0.947	1
MI	0	0.091	0.164	0.196	0.223	0.268	0.299	0.314	0.310	0.299	0.279	0.204	0
δ	0	0.407	0.529	0.569	0.600	0.644	0.671	0.683	0.680	0.671	0.654	0.579	0

According to Table 2 and Figure 1, we conclude that MI isn't necessarily an increasing function of the τ and θ. Also the highest MI value of C-A family occurs at $\tau = 0.632006$, and it is $MI = 0.315552$.

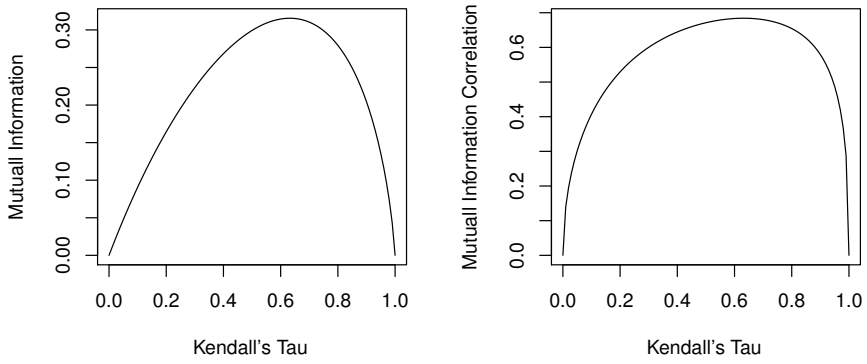

FIGURE 1: MI of C-A family for various values of τ.

3.2. Clayton Copula

The Clayton copula belongs to the class of Archimedean copulas. Its generator is $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$. This copula is referred to Clayton (1978), Cook & Johnson (1981) and Oakes (1982) and is defined as:

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}, \quad \theta \in (0, \infty).$$

When θ tends to infinity, the degree of association increases, i.e. the Clayton copula approaches the monotonic copula; for θ converging to zero, one obtains the independence copula. In particular, this family is positively ordered, and its members are absolutely continuous. The density of Clayton copula is

$$c(u, v) = (1 + \theta)(uv)^{-(1+\theta)}(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1+2\theta}{\theta}}.$$

***Note* 4.** By using Note 1 it is easy to show that Kendall's tau in Clayton family is given by

$$\tau = \frac{\theta}{\theta + 2}.$$

Blumentritt & Schmid (2012) estimate values of MI for Clayton copula by Monte Carlo simulations. By using numerical integration, we again compute the values of MI for this copula. Similar to Table 2, the quantity of θ was chosen according to the values of Kendall's τ. Results display in Table 3. Note that for $\tau = 0.9$, the R package "cubature" doesn't work, therefore, in this case, we compute MI by using Riemann integration with 10^8 cubes.

TABLE 3: The values of MI and normalized MI with respect to the amounts of Kendall's τ for Clayton copula.

τ	0	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
θ	0	0.222	0.500	0.667	0.857	1.333	2	3	4.667	6	8	18
MI	0	0.019	0.072	0.111	0.157	0.275	0.432	0.636	0.911	1.089	1.308	2.003
δ	0	0.192	0.366	0.446	0.520	0.651	0.760	0.848	0.915	0.942	0.963	0.991

3.3. Frank Copula

The Frank copula is introduced by Frank (1979), some of the statistical properties of this family were discussed in Nelsen (1986) and Genest (1987). This copula is an Archimedean copula and $\phi(t) = -\log((e^{-\theta t} - 1)/(e^{-\theta} - 1))$ is generator function. Thus, for $\theta > 0$, yielding

$$C(u, v) = -\frac{1}{\theta} log \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)}\right], \quad \theta \in (0, \infty),$$

for $\theta = 0$ it is defined as the independence copula. Frank copula is positively ordered and it is completely monotonic on $\theta > 0$, i.e. the degree of association increases as θ tends to infinity. The density of Frank copula is

$$c(u, v) = \frac{-\theta(e^{-\theta} - 1)e^{-\theta(u+v)}}{\left[(e^{-\theta u} - 1)(e^{-\theta v} - 1) + e^{-\theta} - 1\right]^2}.$$

***Note* 5.** By using Note 1 it is easy to show that Kendall's tau in Frank family is given by

$$\tau = 1 - \frac{4}{\theta}[1 - D_1(\theta)],$$

where D_1 is the Debye function of order 1, Debye function of order k is defined as:

$$D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt.$$

Blumentritt & Schmid (2012) estimate values of MI for Frank copula by Monte Carlo simulations. By using numerical integration, we again compute the values of MI for this copula according to the amounts of Kendall's τ. Results display in Table 4.

Analogously to the Clayton copula, for $\tau = 0.9$, the R package "cubature" doesn't work, therefore, in this case, we compute MI by using Riemann integration with 10^8 cubes.

TABLE 4: The values of MI and normalized MI with respect to the amounts of Kendall's τ for Frank copula.

τ	0	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
θ	0	0.907	1.861	2.372	2.917	4.161	5.736	7.930	11.411	14.138	18.191	38.281
MI	0	0.011	0.046	0.072	0.105	0.193	0.316	0.485	0.723	0.881	1.082	1.800
δ	0	0.149	0.296	0.367	0.435	0.566	0.684	0.788	0.874	0.910	0.941	0.986

3.4. Gumbel Copula

Gumbel copula was first discussed by Gumbel (1960), hence many authors refer to it as the Gumbel family. However, because Gumbel's name is attached to another Archimedean family and this family also appears in Hougaard (1986), Hutchinson & Lai (1990) refer to it as the Gumbel-Hougaard family. This copula is an Archimedean copula with generator $\phi(t) = (-\ln t)^\theta$. The Gumbel copula is defined as

$$C(u,v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}, \quad \theta \in [1,\infty),$$

$\theta = 0$ is implied the independence copula. This family is positively ordered, and its members are absolutely continuous. The density of Gumbel copula is

$$c(u,v) = \frac{C(u,v)}{uv} \frac{[(-\ln u)(-\ln v)]^{\theta-1}}{[(-\ln u)^\theta + (-\ln v)^\theta]^{2-\frac{1}{\theta}}} \left\{[(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}} + \theta - 1\right\}.$$

***Note* 6.** By using Note 1 it is easy to show that Kendall's tau in Gumbel family is given by

$$\tau = 1 - \frac{1}{\theta}.$$

By using numerical integration we compute values of MI for Gumbel copula according to the amounts of Kendall's τ. Results display in Table 5.

Table 5: The values of MI and normalized MI with respect to the amounts of Kendall's τ for Gumbel copula.

τ	0	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
θ	1	1.111	1.250	1.333	1.428	1.667	2	2.500	3.333	4	5	10
MI	0	0.018	0.064	0.096	0.136	0.237	0.375	0.562	0.821	0.991	1.205	1.884
δ	0	0.188	0.346	0.418	0.487	0.615	0.727	0.821	0.898	0.929	0.954	0.988

3.5. Raftery Copula

Raftery (1984) and Raftery (1985) introduced a one-parameter ($\theta \in [0,1]$) family of absolutely continuous (for $\theta \neq 1$) bivariate distributions with identically distributed exponential margins. The survival copulas for the Raftery family are given by

$$C(u,v) = \min(u,v) + \frac{1-\theta}{1+\theta}(uv)^{\frac{1}{1-\theta}} \left\{1 - [\max(u,v)]^{\frac{-(1+\theta)}{1-\theta}}\right\},$$

independence corresponds to for $\theta = 0$. This family is positively ordered and the density of Raftery copula is

$$c(u,v) = \frac{1}{1-\theta^2} \left\{(uv)^{\frac{\theta}{1-\theta}} + \theta[\min(u,v)]^{\frac{\theta}{1-\theta}}[\max(u,v)]^{\frac{-1}{1-\theta}}\right\}.$$

***Note* 7.** By using Note 1 it is easy to show that Kendall's tau in Raftery family is given by

$$\tau = \frac{2\theta}{3-\theta}.$$

By using numerical integration, we compute values of MI according to the amounts of Kendall's τ. Results display in Table 6.

Table 6: The values of MI and normalized MI with respect to the amounts of Kendall's τ for Raftery copula.

τ	0	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
θ	0	0.143	0.273	0.333	0.391	0.500	0.600	0.692	0.778	0.818	0.857	0.931
MI	0	0.032	0.098	0.142	0.192	0.315	0.473	0.676	0.947	1.123	1.340	2.023
δ	0	0.248	0.422	0.497	0.565	0.684	0.782	0.861	0.922	0.946	0.965	0.991

3.6. Gaussian Copula

The Gaussian copula is defined implicitly by

$$C(u,v) = \Phi_\theta(\Phi^{-1}(u), \Phi^{-1}(v)),$$

where Φ_θ is the distribution function of the bivariate normal distribution with zero means, variances one, and correlation parameter θ and Φ^{-1} denotes the inverse of

the univariate standard normal distribution function (For more details see Meyer 2013). Therefore the Gaussian copula expressed as

$$C(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{\exp\left\{-\frac{t^2+s^2-2\theta ts}{2(1-\theta^2)}\right\}}{2\pi\sqrt{1-\theta^2}} dtds, \quad \theta \in [0,1],$$

when $\theta = 0$, independence is implied. The Gaussian copula is positively ordered, and its members are absolutely continuous. The density of Gaussian copula is

$$c(u,v) = \sqrt{1-\theta^2}\ \exp\left\{-\frac{x^2+y^2-2\theta xy}{2(1-\theta^2)} + \frac{x^2+y^2}{2}\right\},$$

where $x = \Phi^{-1}(u)$ and $y = \Phi^{-1}(v)$.

***Note* 8.** By using Note 1, in the case of meta-elliptical distributions (Fang, Fang & Kotz 2002), that includes Gaussian and T-copulas, Kendall's tau is related to the correlation parameter θ as

$$\tau = \frac{2}{\pi}\sin^{-1}(\theta). \tag{6}$$

For the first time, Kullback (1952) obtains MI of the multivariate normal distribution with the title 'mean information' and Kullback (1959) studies it with more details.

Proposition 2. *Let (X,Y) be a random vector with Gaussian copula, then MI is given by*

$$MI = -\frac{1}{2}\log(1-\theta^2). \tag{7}$$

***Proof*.** See Equation (6.16), page 97 from Section 6 of Kullback (1952), Example 4.3 from page 8 and Equation (7.16), page 203 from Section 9.7 of Kullback (1959). □

By using Equations (6) and (7) we can compute MI of Gaussian family for various values of Kendall's τ, i.e. θ was chosen according to the values of τ given in the first row. Results display in Table 7.

Table 7: The values of MI and normalized MI with respect to the amounts of Kendall's τ for Gaussian copula.

τ	0	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
θ	0	0.156	0.309	0.383	0.454	0.588	0.707	0.809	0.891	0.924	0.951	0.988
MI	0	0.012	0.050	0.079	0.115	0.212	0.346	0.531	0.790	0.960	1.174	1.855
δ	0	0.156	0.309	0.383	0.454	0.588	0.707	0.809	0.891	0.924	0.951	0.988

3.7. T-Copulas

Let $t_{\theta,\nu}$ denote the standardized multivariate Student's t distribution function with correlation parameter θ and ν degrees of freedom and t_ν^{-1} the inverse of the univariate distribution function of the Student's t distribution with ν degrees of freedom. The T-copula, with ν degrees of freedom, is given by

$$C(u,v) = t_{\theta,\nu}(t^{-1}(u), t^{-1}(v)),$$

therefore the T-copula expressed as

$$C(u,v) = \int_{-\infty}^{t^{-1}(u)} \int_{-\infty}^{t^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \left\{1 + \frac{t^2+s^2-2\theta ts}{\nu(1-\theta^2)}\right\}^{-\frac{\nu+2}{2}} dtds, \quad \theta \in [0,1],$$

$\theta = 0$ is implied the independence copula. This family is positively ordered, and its members are absolutely continuous. The density of T-copula is

$$c(u,v) = \frac{\Gamma(\frac{\nu+2}{2})\Gamma(\frac{\nu}{2})}{\left[\Gamma(\frac{\nu+1}{2})\right]^2\sqrt{1-\theta^2}} \frac{\left[1+\frac{q_\theta(t_\nu^{-1}(u),t_\nu^{-1}(v))}{\nu}\right]^{-\frac{\nu+2}{2}}}{\left[1+\frac{(t_\nu^{-1}(u))^2}{\nu}\right]^{-\frac{\nu+1}{2}}\left[1+\frac{(t_\nu^{-1}(v))^2}{\nu}\right]^{-\frac{\nu+1}{2}}},$$

where $q_\theta(x,y) = \frac{x^2+y^2-2\theta xy}{1-\theta^2}$ and $\Gamma(.)$ is Gamma function.

For the first time, Guerrero-Cusumano (1996*a*), Guerrero-Cusumano (1996*b*) obtains MI of multivariate t distribution and Calsaverini & Vicente (2009) offer a simpler formula for MI in T-copulas.

Proposition 3. *Let (X,Y) be a random vector with T-copula, it can be shown that MI can be decomposed as*

$$MI = MI_{Gauss}(\theta) + MI_{Excess}(\nu), \tag{8}$$

where $MI_{Gauss}(\theta) = -\frac{1}{2}log(1-\theta^2)$ is MI for the Gaussian copula and

$$MI_{Excess}(\nu) = 2log\left[\sqrt{\frac{\nu}{2\pi}}\,\beta(\frac{\nu}{2},\frac{1}{2})\right] - \frac{2+\nu}{\nu} + (1+\nu)\left[\psi(\frac{\nu+1}{2}) - \psi(\frac{\nu}{2})\right], \tag{9}$$

is an excess of information in the dependence with respect to the Gaussian copula. Here, $\psi(.)$ and $\beta(.)$ denotes the Digamma and Beta functions, respectively.

Proof. See Lemma 1 from Guerrero-Cusumano (1996*a*), Lemma 1 from Guerrero-Cusumano (1996*b*) and page 4 from Calsaverini & Vicente (2009). □

***Note* 9.** Note that MI_{Excess} of T-copula is constant function from the dependence parameter θ and consequently it is constant with respect to the Kendall's τ.

By using Equation (9), we can compute MI_{Excess} of T-copula for various values of degrees of freedom ν. Results display in Table 8, therefore for computing MI in T-copula with degrees of freedom ν, it is sufficient that add the constant value of Table 8 with values of MI from Gaussian copula (Table 7). In Figure 2, we depict the behavior of MI_{Excess} versus the degrees of freedom ν.

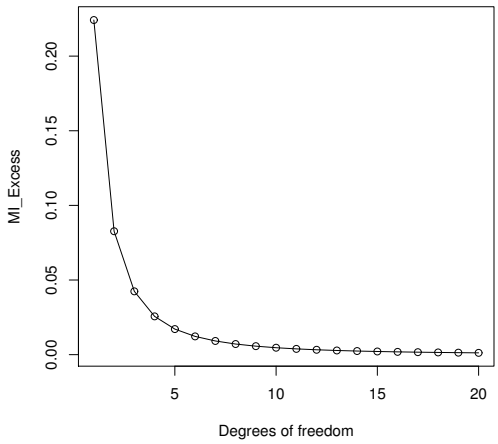

FIGURE 2: The MI_{Excess} of T-copula for various values of ν.

Table 8 and Figure 2 show that MI_{Excess} is a decreasing function of degrees of freedom ν, consequently when ν tends to infinity, the MI_{Excess} of T-copula tends to zero. Therefore, as we expect, when ν tends to infinity, the MI of T-copula tends to MI of Gaussian copula.

TABLE 8: The values of MI and normalized MI with respect to the amounts of Kendall's τ for T-copulas.

ν	1	2	3	4	5	6	7	8	9
MI_{Excess}	0.224	0.083	0.042	0.026	0.017	0.012	0.009	0.007	0.006
ν	10	12	15	20	30	50	100	200	500
MI_{Excess}	0.005	0.003	0.002	0.001	0.0005	0.0002	0.00005	0.00001	0.000002

4. Comparing mutual information

In this section, we compare MI correlation for some family of bivariate distribution mentioned in Section 3, visually. Note that, we don't have considered Cuadras-Auge copula, because MI correlation of this family is not increasing and it is limited on $[0, 0.6841022]$. Figure 3 shows the MI correlation from various copulas with respect to Kendall's τ correlation. For greater clarity, in the left graph, we draw the plots of Gaussian and T-copulas, and in the right graph, we consider

the plots of Gaussian, Clayton, Frank, Gumbel, and Raftery copulas. As be seen in figure, MI correlation of T-copulas, Clayton, Gumbel, and Raftery copulas is greater than Gaussian copula and MI correlation of Frank copula is lesser than Gaussian copula.

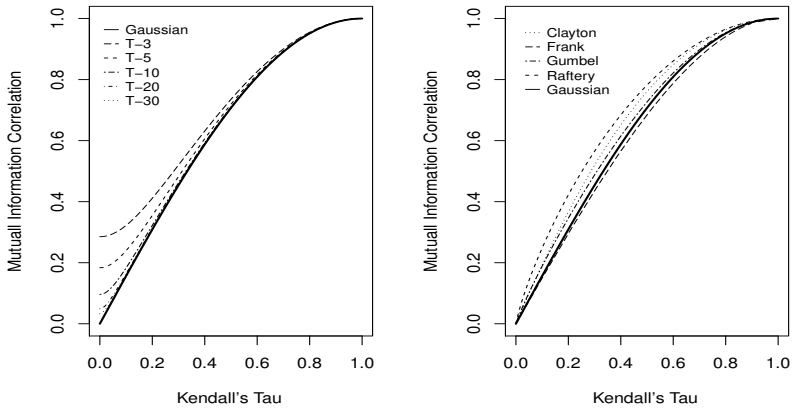

FIGURE 3: Comparing plots: the MI correlation from different copula with respect to Kendall's τ.

5. Data Analysis: Insurance Data

In this section, using a real dataset we illustrate the role of Kendall's τ in estimating MI. For this purpose, we apply Insurance dataset. Insurance data are given in data frame 'Insurance' in R software consist of the numbers of policyholders of an insurance company who were exposed to risk, and the numbers of car insurance claims made by those policyholders in the third quarter of 1973. In this dataset 'Holders' is numbers of policyholders and 'Claims' is numbers of claims. First, by the raw data, we compute sample Kendall's τ coefficient for two variables 'Holders' and 'Claims' in insurance data equal to 0.307. Then, by using the goodness of fit tests, we fit a copula function for two variables 'Holders' and 'Claims'. This work is done by 'copula' package in R software (For more details see Genest, Remillard & Beaudoin, 2009). We express the results in Table 9.

TABLE 9: Goodness of fit test for various copulas.

Copula Function	Statistic	$\hat{\theta}$	p-value
Frank	0.11971	2.9931	0.0004995
Gumbel	**0.0248**	**1.4254**	**0.1753**
Clayton	0.72023	0.49891	0.0004995
Normal	0.08931	0.45909	0.0004995
T-copula	0.11341	0.43387	0.0004995

According to p-values, we conclude that Gumbel copula is an appropriate copula for this dataset. Finally, using Gumbel copula and sample Kendall's τ coefficient, we estimate MI between two variables 'Holders' and 'Claims' equal to 0.142. We can control the results using 'mpmi' package in R software. By this package, we estimate MI equal to 0.157. Due to the small difference between the two estimated values, it can be concluded that we can employ copula function and Kendall's τ for estimating MI.

6. Conclusion

In this paper, by using the copula function, we have estimated the MI of some bivariate distributions. We have shown that, except Cuadras-Auge copula, the MI for mentioned copulas is increasing the function of Kendall's τ, i.e. Kendall's τ and MI have a direct relationship. But Cuadras-Auge copula is a counterexample, and it proves that this result is not true, generally. This example shows that the range of MI correlation is not necessarily $[0, 1]$, even if Kendall's τ is within $[0, 1]$.

Acknowledgement

The author would like to thank the reviewers for their valuable suggestions and comments.

[Received: February 2019 — Accepted: July 2019]

References

Arellano-Valle, R. B., Contreras-Reyes, J. E. & Genton, M. G. (2013), 'Shannon Entropy and Mutual Information for Multivariate Skew-Elliptical Distributions', *Scandinavian Journal of Statistics* **40**(1), 42–62.

Bell, C. B. (1962), 'Mutual information and maximal correlation as measures of dependence', *The Annals of Mathematical Statistics* **33**(2), 587–595.

Blumentritt, T. & Schmid, F. (2012), 'Mutual information as a measure of multivariate association: analytical properties and statistical estimation', *Journal of Statistical Computation and Simulation* **82**(9), 1257–1274.

Calsaverini, R. S. & Vicente, R. (2009), 'An information-theoretic approach to statistical dependence: Copula information', *EPL (Europhysics Letters)* **88**(6), 68003.

Clayton, D. G. (1978), 'A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence', *Biometrika* **65**(1), 141–151.

Cook, R. D. & Johnson, M. E. (1981), 'A family of distributions for modeling non-elliptically symmetric multivariate data', *Journal of the Royal Statistical Society, Series B* **43**(2), 210–218.

Cuadras, C. M. & Auge, J. (1981), 'A continuous general multivariate distribution and its properties', *Communications in Statistics-Theory and Methods* **10**(4), 339–353.

Dobrowolski, E. & Kumar, P. (2014), 'Some properties of the Marshall-Olkin and generalized Cuadras-Auge families of copulas', *Australian Journal of Mathematical Analysis and Applications* **11**(1), 1–13.

Fang, H. B., Fang, K. T. & Kotz, S. (2002), 'The meta-elliptical distributions with given marginals', *Journal of Multivariate Analysis* **82(1)**, 1–16.

Frank, M. J. (1979), 'On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$', *Aequationes Mathematicae* **99**(1), 194–226.

Genest, C. (1987), 'Frank's family of bivariate distributions', *Biometrika* **74**(3), 145–159.

Genest, C. & MacKay, R. J. (1986*a*), 'Copules archimediennes et families de lois bidimensionnelles dont les marges sont données', *Canadian Journal of Statistics* **14**(2), 145–159.

Genest, C. & MacKay, R. J. (1986*b*), 'The joy of copulas: bivariate distributions with uniform marginals', *The American Statistician* **40**(4), 280–283.

Genest, C., Remillard, B. & Beaudoin, D. (2009), 'Goodness-of-fit tests for copulas: A review and a power study', *Insurance: Mathematics and Economics* **44**(2), 199–213.

Guerrero-Cusumano, J. L. (1996*a*), 'A measure of total variability for the multivariate t distribution with applications to finance', *Information Sciences* **92**(1), 47–63.

Guerrero-Cusumano, J. L. (1996*b*), 'An asymptotic test of independence for multivariate t and Cauchy random variables with applications', *Information Sciences* **93**(1), 33–45.

Gumbel, E. J. (1960), 'Distributions des valeurs extrêmes en plusieurs dimensions', *Publications de l'Institut de statistique de l'Université de Paris* **9**, 171–173.

Hougaard, P. (1986), 'A class of multivanate failure time distributions', *Biometrika* **73**(3), 671–678.

Hutchinson, T. P. & Lai, C. D. (1990), *Continuous bivariate distributions emphasising applications*, Rumsby Scientific Publishing, Adelaide.

Jenison, R. L. & Reale, R. A. (2004), 'The shape of neural dependence', *Neural computation* **16**(4), 665–672.

Joe, H. (1989), 'Relative entropy measures of multivariate dependence', *Journal of the American Statistical Association* **84**(405), 157–164.

Kendall, M. G. (1938), 'A new measure of rank correlation', *Biometrika* **30**(1/2), 81–93.

Kinney, J. B. & Atwal, G. S. (2014), 'Equitability, mutual information, and the maximal information coefficient', *of the National Academy of Sciences* **111**(9), 3354–3359.

Kullback, S. (1952), 'An application of information theory to multivariate analysis', *The Annals of Mathematical Statistics* **23**(1), 88–102.

Kullback, S. (1959), *Information Theory and Statistics*, Wiley, New York.

Kumar, P. (2012), 'Statistical Dependence: Copula functions and mutual information based measures', *Journal of Statistics Applications and Probability: An International Journal* **1**(1), 1–14.

Kwak, N. & Choi, C. H. (2002), 'Input feature selection by mutual information based on Parzen window', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(12), 1667–1671.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. & Suetens, P. (1997), 'Multimodality image registration by maximization of mutual information', *IEEE transactions on Medical Imaging* **16**(2), 187–198.

Mercier, G. (2005), Mesures de dépendance entre images rso, Technical report, GET/ENST Bretagne, Tech. Rep. RR-2005003-ITI.

Meyer, C. (2013), 'The bivariate normal copula', *Communications in Statistics-Theory and Methods* **42**(13), 2402–2422.

Nelsen, R. B. (1986), 'Properties of a one-parameter family of bivariate distributions with specified marginals', *Communications in Statistics-Theory and Methods* **15**(11), 3277–3285.

Nelsen, R. B. (2006), *An Introduction to Copulas*, Springer, New York.

Oakes, D. (1982), 'A model for association in bivariate survival data', *Journal of the Royal Statistical Society, Series B* **44**(3), 414–422.

R Development Core Team (2012), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org

Raftery, A. E. (1984), 'A continuous multivariate exponential distribution', *Communications in Statistics-Theory and Methods* **13**(8), 947–965.

Raftery, A. E. (1985), 'Some properties of a new continuous bivariate exponential distribution', *Statistics and Decisions, Supplement Issue* **2**, 53–58.

Shannon, C. & Weaver, W. (1949), *The Mathematical Theory of Communication*, University of Illinois Press, Urbana.

Sklar, A. (1959), 'Fonctions de répartition à n dimensions et leurs marges', *Publications de l'Institut de statistique de l'Université de Paris* **8**, 229–231.

Zeng, X. & Durrani, T. S. (2011), 'Estimation of mutual information using copula density function', *Electronics Letters* **47**(8), 493–494.