entropy

MDPI

Article

The Structure of the Class of Maximum Tsallis–Havrda–Chavát Entropy Copulas

Jesús E. García [1], Verónica A. González-López [1,*] and Roger B. Nelsen [2]

1 Department of Statistics, University of Campinas, Rua Sérgio Buarque de Holanda, 651, Campinas, São Paulo 13083-859, Brazil; jg@ime.unicamp.br
2 Department of Mathematical Sciences, Lewis & Clark College, 0615 SW Palatine Hill Rd., Portland, OR 97219, USA; nelsen@lclark.edu
* Correspondence: veronica@ime.unicamp.br; Tel.: +55-19-3521-6057

Academic Editors: Julio Stern and Adriano Polpo
Received: 21 May 2016; Accepted: 14 July 2016; Published: 19 July 2016

Abstract: A maximum entropy copula is the copula associated with the joint distribution, with prescribed marginal distributions on $[0,1]$, which maximizes the Tsallis–Havrda–Chavát entropy with $q = 2$. We find necessary and sufficient conditions for each maximum entropy copula to be a copula in the class introduced in Rodríguez-Lallena and Úbeda-Flores (2004), and we also show that each copula in that class is a maximum entropy copula.

Keywords: Tsallis–Havrda–Chavát entropy; Copula; Full Bayesian Significance Test

MSC: 62Hxx; 62B10; 62F15

1. Introduction

In thermodynamics, entropy is a measure of randomness or disorder, and in statistics, entropy is a measure of uncertainty in a probability distribution. During the past one and a half centuries, a number of such measures have been proposed. In terms of a continuous bivariate distribution with density $f(x,y)$, three well known entropies are the following:

$$\begin{aligned} B(f) &= -\int\int f(x,y)\ln(f(x,y))dxdy \text{ [Boltzmann–Gibbs–Shannon (BGS)];} \\ R_q(f) &= \frac{1}{1-q}\log_2\Big(\int\int f(x,y)^q dxdy\Big),\ q \geq 0,\ q \neq 1 \text{ [Rényi];} \\ T_q(f) &= \frac{1}{q-1}\Big(1-\int\int f(x,y)^q dxdy\Big),\ q \geq 0,\ q \neq 1 \text{ [Tsallis–Havrda–Chavát (THC)].} \end{aligned}$$

These measures are related to one another. For example, the limit as q approaches 1 of the THC entropy (see [1]) is the BGS entropy; and the THC and Rényi entropies for a common value of q are monotone functions of each other.

One objective in the study of entropy is to identify the joint distributions that maximize entropy subject to given margins, and to identify the copulas associated with those distributions. In [2] Pougaza and Mohammad-Djafari use Lagrange multipliers to show that when the support of $f(x,y)$ is $\mathbf{I}^2$, $\mathbf{I} = [0,1]$, the joint probability density function that maximizes BGS entropy is $f(x,y) = f_1(x)f_2(y)$, the product of the two marginal probability density functions. Thus, in the case of BGS entropy, every joint distribution maximizing BGS entropy has the same copula, namely the independence copula.

Pougaza and Mohammad-Djafari [2] also identified the joint distributions that maximize THC entropy in the case $q = 2$ (a case of interest since it coincides with Simpson's diversity index [3]).

In [4] the authors construct several families of copulas by considering the THC entropy with index $q = 2$ and particular expressions for the marginals. Our goal in this paper is to study the copulas associated with those joint distributions. We show that the family of maximum entropy copulas (in the THC $q = 2$ sense) coincides with the family of copulas studied previously by Rodríguez-Lallena and Úbeda-Flores [5]. In practical terms, this result guarantees that researchers using maximum entropy copulas (in the THC $q = 2$ sense) can find in [5] all the relevant properties of the copulas in [2].

We proceed as follows. In the next section we review the maximization of THC entropy in the case $q = 2$. Since not every solution to the Lagrange multiplier optimization in [2] yields a proper bivariate distribution, in Section 3 we establish necessary and sufficient conditions for the solutions to be proper bivariate distributions and identify the associated copulas. We illustrate the results with several examples.

2. Preliminaries

Consider the bivariate probability density function $f(x,y)$ with $(x,y) \in \mathbf{I}^2$, $\mathbf{I}$=[0,1], which maximizes the THC entropy $T_2(f)$, given by

$$T_2(f) = 1 - \int_0^1 \int_0^1 f^2(x,y)dxdy, \tag{1}$$

with the following constraints

$$\int_0^1 f(x,y)dy = f_1(x), \ \forall x \in \mathbf{I}, \tag{2}$$

$$\int_0^1 f(x,y)dx = f_2(y), \ \forall y \in \mathbf{I}, \tag{3}$$

$$\int_0^1 \int_0^1 f(x,y)dxdy = 1. \tag{4}$$

We use the notations $T_2(f)$ and $T_2(F)$ to refer to the THC ([1]) entropy of the density f and the entropy of the cumulative distribution F, respectively.

We note that distributions with the same copula but different marginals (hence different joint densities) may have different entropies (1). For example, consider the following cases, all with the product copula and different margins: $F_1(x,y) = xy,\ x,y \in \mathbf{I}, T_2(F_1) = 0$; $F_2(x,y) = x^2y^2,\ x,y \in \mathbf{I}, T_2(F_2) = -\frac{7}{9}$; and $F_3(x,y) = \sqrt{xy},\ x,y \in \mathbf{I}, T_2(F_3) = -\infty$.

According to [2] for absolutely continuous distribution functions $F_1(x)$ and $F_2(y)$, the joint density and distribution functions $f(x,y)$ and $F(x,y)$ on $\mathbf{I}^2$ with maximum entropy given by Equation (1), under the constraints (2)–(4) are $f(x,y) = f_1(x) + f_2(y) - 1$ and

$$F(x,y) = yF_1(x) + xF_2(y) - xy,\ \ x,y \in \mathbf{I} \tag{5}$$

for appropriate functions F_1 and F_2, and with a maximum entropy (ME) copula

$$C(u,v) = uF_2^{-1}(v) + vF_1^{-1}(u) - F_2^{-1}(v)F_1^{-1}(u),\ u,v \in \mathbf{I}. \tag{6}$$

Not every choice of cumulative distribution functions F_1 and F_2 in (5) yields a bivariate distribution function F. For example, let $F_1(x) = x^3$ and $F_2(y) = y^3$. Then $f(x,y) = 3x^2 + 3y^2 - 1$, which equals $-\frac{1}{3}$ when $x = y = \frac{1}{3}$.

Since the only maximum entropy copula for BGS entropy is the product copula, the entropy (1) identifies models of greater flexibility in terms of dependence structure than models derived from BGS entropy.

Given real functions $f, g : \mathbf{I} \to \mathbb{R}$, define

$$\alpha = \inf\{f'(u) : u \in A\}, \beta = \sup\{f'(u) : u \in A\}, \text{ where } A = \{u \in \mathbf{I} : f'(u) \text{ exists }\}, \tag{7}$$

$$\gamma = \inf\{g'(v) : v \in B\}, \delta = \sup\{g'(v) : v \in B\}, \text{ where } B = \{v \in \mathbf{I} : g'(v) \text{ exists }\}, \tag{8}$$

and consider the function C given by

$$C(u,v) = uv + f(u)g(v), \text{ for all } u,v \in \mathbf{I}. \tag{9}$$

According to Theorem 2.3 in [5], C in (9) is a copula, if and only if (i) $f(0) = f(1) = g(0) = g(1) = 0$; (ii) f and g are absolutely continuous and (iii) $\min\{\alpha\delta, \beta\gamma\} \geq -1$, where α, β, γ and δ are introduced in Equations (7) and (8). Furthermore, in such a case, C is absolutely continuous. We call a copula given by Equation (9) a RU (Rodríguez-Lallena and Úbeda-Flores) copula.

3. Results

We initiate this section with a result, showing that every ME copula (Equation (6)) has the form of a RU copula (Equation (9)). Since each inverse cumulative distribution function F_i^{-1} is absolutely continuous and nondecreasing on its domain $\mathbf{I}$, the inverse F_i of F_i^{-1} is an absolutely continuous distribution function on $\mathbf{I}$. We find necessary and sufficient conditions on the functions F_1 and F_2 (in terms of the derivatives of their inverses) for the ME copula in Equation (6) to be a RU copula. These conditions also insure that the functions F_1 and F_2 yield a proper joint distribution function in (5).

Theorem 1.

i *Every ME copula has the functional form of a RU copula.*

ii *Let $F_1^{-1}(u)$ and $F_2^{-1}(v)$ be inverse distribution functions on $\mathbf{I}$, and set $A = \{u \in \mathbf{I} : \frac{d}{du}F_1^{-1}(u) \text{ exists}\}$, $B = \{v \in \mathbf{I} : \frac{d}{dv}F_2^{-1}(v) \text{ exists}\}$, $m_1 = \inf\{\frac{d}{du}F_1^{-1}(u) : u \in A\}$, $M_1 = \sup\{\frac{d}{du}F_1^{-1}(u) : u \in A\}$, $m_2 = \inf\{\frac{d}{dv}F_2^{-1}(v) : v \in B\}$, $M_2 = \sup\{\frac{d}{dv}F_2^{-1}(v) : v \in B\}$. Then the function $C(u,v)$ in Equation (6) is a copula-in fact a RU copula-if and only if* $\max\{(M_1-1)(M_2-1), (m_1-1)(m_2-1)\} \leq 1$.

Proof.

i Equation (6) is equivalent to

$$C(u,v) = uv + [u - F_1^{-1}(u)][F_2^{-1}(v) - v], \tag{10}$$

which has the form of a RU copula with $f(u) = u - F_1^{-1}(u)$ and $g(v) = F_2^{-1}(v) - v$.

ii Since $f'(u) = 1 - \frac{d}{du}F_1^{-1}(u)$ and $g'(v) = \frac{d}{dv}F_2^{-1}(v) - 1$, $\alpha = 1 - M_1 = \inf\{f'(u) : u \in A\}$, $\beta = 1 - m_1 = \sup\{f'(u) : u \in A\}$, $\gamma = m_2 - 1 = \inf\{g'(v) : v \in B\}$, and $\delta = M_2 - 1 = \sup\{g'(v) : v \in B\}$. The condition $\min\{\alpha\delta, \beta\gamma\} \geq -1$ given in Theorem 2.3 [5] is equivalent to $\max\{(M_1-1)(M_2-1), (m_1-1)(m_2-1)\} \leq 1$, and the conclusion follows.

□

Example 1. Consider the function $C(u,v) = vu^{a+1} + uv^{b+1} - u^{a+1}v^{b+1}$ for $a, b \geq 0$, which has the form of Equation (6) with $F_1^{-1}(u) = u^{a+1}$ and $F_2^{-1}(v) = v^{b+1}$. It follows from differentiation that $m_1, m_2 = 0$ or 1, $M_1 = a+1$ and $M_2 = b+1$. Hence C is a copula if and only if $ab \leq 1$. C also has the form $uv[1-(1-u^a)(1-v^b)]$, a generalized Farlie–Gumbel–Morgenstern copula.

Example 2. Let $F_1^{-1}(u) = 3u^2 - 2u^3$ and $F_2^{-1}(v) = (1+\theta)v - \theta(3v^2 - 2v^3)$ on $\mathbf{I}$, with $\theta \in [-1, 2]$. Here $m_1 = 0$, $M_1 = \frac{3}{2}$. In addition $m_2 = 0$, $M_2 = 3$, since $g'(v) = \theta f'(v)$ (see Proof of Theorem 1) and for $-1 \leq \theta \leq 2$ the supremum and the infimum of $g'(v)$ will occur when $\theta = -1$ or $\theta = 2$. So $m_2 = 1 + \min\{-f'(0), -f'(1), 2f'(\frac{1}{2})\} = 1 + (-1) = 0$ and $M_2 = 1 + \max\{-f'(\frac{1}{2}), 2f'(0), 2f'(1)\} = 1 + 2 = 3$. Hence $\max\{(M_1-1)(M_2-1), (m_1-1)(m_2-1)\} = 1$. In this case the functions C in Equation (10) are $C(u,v) = uv + \theta uv(1-u)(1-v)(1-2u)(1-2v)$, copulas with cubic sections in Example 3.15 in [6].

We now state and prove a converse of Theorem 1 by showing that every RU copula is a ME copula.

Theorem 2. *Let $C(u,v) = uv + f(u)g(v)$ be a RU copula. Then there exists a constant $k > 0$ such that setting $F_1^{-1}(u) = u - kf(u)$ and $F_2^{-1} = \frac{1}{k}g(v) + v$ yields the ME copula $C(u,v) = uF_2^{-1}(v) + vF_1^{-1}(u) - F_2^{-1}(v)F_1^{-1}(u)$.*

Proof. Let $C(u,v) = uv + f(u)g(v)$ be a RU copula. We need to show that there exists a $k > 0$ such that $G_1(u) = u - kf(u)$ and $G_2(v) = \frac{1}{k}g(v) + v$ are inverse distribution functions on $\mathbf{I}$. Clearly, from Theorem 2.3 in [5], $G_1(0) = G_2(0) = 0$ and $G_1(1) = G_2(1) = 1$, so it suffices to show that $G_1'(u) = 1 - kf'(u) \geq 0$ and $G_2'(v) = \frac{1}{k}g'(v) + 1 \geq 0$ for some $k > 0$. Since $\min\{\alpha\delta, \beta\gamma\} \geq -1$ from Theorem 2.3 in [5], both $\alpha\delta \geq -1$ and $\beta\gamma \geq -1$. But $\beta > 0$ and $\gamma < 0$ imply that $0 < -\gamma \leq \frac{1}{\beta}$, hence there exists a $k > 0$ such that $-\gamma \leq k \leq \frac{1}{\beta}$. Thus $f'(u) \leq \beta$ implies that $kf'(u) \leq k\beta \leq 1$, and hence $G_1'(u) \geq 0$. Similarly $g'(v) \geq \gamma$ implies that $\frac{1}{k}g'(v) \geq \frac{1}{k}\gamma \geq -1$, and hence $G_2'(v) \geq 0$. Thus $G_1(u)$ and $G_2(v)$ are inverse distribution functions on $\mathbf{I}$. A similar result follows from $\alpha\delta \geq -1$ setting $G_1(u) = kf(u) + u$ and $G_2(v) = v - \frac{1}{k}g(v)$. □

Example 3. The function $C(u,v) = uv - \frac{uv(1-u)(1-v)}{(1+u)(1+v)}$ is a RU copula with $f(u) = -\frac{u(1-u)}{1+u}$ and $g(v) = \frac{v(1-v)}{1+v}$, with $\alpha = -1$, $\beta = \frac{1}{2}$, $\gamma = -\frac{1}{2}$, and $\delta = 1$. Setting $k = 1$ yields $F_1^{-1}(u) = u - f(u) = \frac{2u}{1+u}$ and $F_2^{-1}(v) = v + g(v) = \frac{2v}{1+v}$. Thus $F_1(x) = \frac{x}{2-x}$, $F_2(y) = \frac{y}{2-y}$, and the maximum entropy bivariate cumulative distribution function $F(x,y)$ with margins $F_1(x)$ and $F_2(y)$, is $F(x,y) = \frac{xy(x+y-xy)}{(2-x)(2-y)}$ (see Equation (5)) with $T_2(F) = -\frac{1}{3}$. As a point of comparison, in the independence case we have $T_2(F_1F_2) = -\frac{13}{36} < -\frac{1}{3}$. The functions $F_1(x)$ and $F_2(y)$ are not unique. If we set $F_1^{-1}(u) = u + f(u) = \frac{2u^2}{1+u}$ and $F_2^{-1}(v) = v - g(v) = \frac{2v^2}{1+v}$ we obtain $F_1(x) = \frac{1}{4}(x + \sqrt{x^2+8x})$, $F_2(y) = \frac{1}{4}(y + \sqrt{y^2+8y})$, and a different $F(x,y)$ with $T_2(F) = -\infty$.

Example 4. Let $F_1^{-1}(u) = u + a^2u(1-u)$ and $F_2^{-1}(v) = v - av(1-v)$ on $\mathbf{I}$ with $a \in [-1,1]$. Then the copula C in Equation (6) with maximum entropy is $C(u,v) = uv + a^3uv(1-u)(1-v)$, a Farlie–Gumbel–Morgenstern copula with parameter $\theta = a^3$. Note that $m_1 = 1 - a^2$, $M_1 = 1 + a^2$, $m_2 = \min\{1-a, 1+a\}$, and $M_2 = \max\{1-a, 1+a\}$. The joint distribution F given by Equation (5) has entropy $T_2(F) = 2 - \frac{1}{2a^2}\ln\left(\frac{1+a^2}{(1+a)^{1-a}(1-a)^{1+a}}\right)$ for $a \neq 0$, and $T_2(F) = 0$ for $a = 0$, where $F_1(x) = \frac{1+a^2-\sqrt{1+2a^2+a^4-4a^2x}}{2a^2}$, $f_1(x) = \frac{1}{\sqrt{1+2a^2+a^4-4a^2x}}$ and $F_2(y) = \frac{a-1+\sqrt{(a-1)^2+4ay}}{2a}$, $f_2(y) = \frac{1}{\sqrt{1-2a+a^2+4ay}}$.

The next example is an application to real data. We use a Farlie–Gumbel–Morgenstern copula to model the dependence between two indicators, under this model we test the independence assumption between such indicators, since it is the solution indicated by the BGS entropy. The probability density function in the Farlie–Gumbel–Morgenstern copula case is $c(u,v) = 1 + \theta(1-2u)(1-2v)$, $u, v \in \mathbf{I}$ with $\theta \in [-1,1]$. The Farlie–Gumbel–Morgenstern copulas are solutions that maximize THC entropy in certain cases (see Example 4). Because we use a data set with moderate sample size, to test the independence, we apply the Full Bayesian Significance Test (FBST) [7], a Bayesian approach especially designed to test precise hypothesis. The aim of this application is to use the family of ME (or RU) copulas derived from THC entropy in Example 4 to test the hypothesis of the independence assumption.

Example 5. The database is composed by paired observations $\{(x_i, y_i)\}_{i=1}^n$ coming from the bivariate vector (X, Y). The observations are taken from each one of the 50 American states and the District of Columbia. The observations of X are "Gonorrhea Rates per 100,000 Population" collected for each of the n=51 units by the Centers for Disease Control and Prevention in 2011. The data can be obtained from Table 14 in [8]. See also [9]. The observations of Y are records of poverty "400% FPL and over" records the upper section of the distribution of the total population by Federal Poverty Level (FPL) for the 51 units, in 2011. The data can be obtained from the Kaiser Family Foundation [10]. The hypothesis of independence in the case of Farlie–Gumbel–Morgenstern copula is translated to test $\theta = 0$, which leads naturally to apply the procedure FBST as follows: $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$, compute $Ev = 1 - \int_T g(\theta|D)d\theta$, where $T = \{\theta : -1 \leq \theta \leq 1 \text{ and } g(\theta|D) > g(0|D)\}$ and $D := \{(u_i, v_i)\}_{i=1}^n, u_i = \frac{rank(x_i)}{n}, v_i = \frac{rank(y_i)}{n}$, $i = 1, \ldots, n$, with posterior density on θ given by $g(\theta|D)$, $\theta \in [-1, 1]$. In FBST, H_0 is rejected when Ev is small, for details about this methodology see [11]. In order to explore some possibilities we compare the results using a prior density on θ proportional to $(\frac{1+\theta}{2})^{a-1}(\frac{1-\theta}{2})^{b-1}$ with a and b hyperparameters of the distribution. We fixed the hyperparameters to consider five settings (i) uniform distribution on $[-1, 1]$, (ii) prior density with a mode on $\theta = 0$, (iii) prior density with a mode on $\theta = -0.5$, (iv) prior density with a mode on $\theta = 0.5$, (v) prior density with a mode near to $\theta = -1$. This last setting is included in concordance with the Spearman's coefficient observed in the data, $\rho_s = -0.32113$, that indicates a θ value near to -0.96.

According to our results (Table 1), we have evidence to reject the independence assumption under the Farlie–Gumbel–Morgenstern copula model, except in a setting as (iv), in which the prior distribution on θ and the likelihood of the data expose some conflict of information, allowing us to disregard this case.

Table 1. Evidence against $H_0 : \theta = 0$, prior density on $\theta \propto (\frac{1+\theta}{2})^{a-1}(\frac{1-\theta}{2})^{b-1}$.

Setting	a	b	Ev
(i)	1	1	0.02686
(ii)	2	2	0.05580
(iii)	2	4	0.01702
(iv)	4	2	0.28670
(v)	2	10	0.00050

4. Conclusions

In [2], Pougaza and Mohammad-Djafari studied bivariate distributions with maximum entropy (in the Tsallis–Havrda–Chavát sense with $q = 2$) for the purpose of generating new families of copulas, which we call ME copulas. In this paper we show that every ME copula coincides with a member of a family studied by Rodríguez-Lallena and Úbeda-Flores in [5], copulas we call RU copulas. Furthermore, we show that every RU copula is a ME copula, and provide illustrative examples.

Acknowledgments: The authors gratefully acknowledge the financial support provided by FAEPEX-PRP of University of Campinas. The authors also wish to express their gratitude to three referees and an associate editor for their helpful comments on an earlier draft of this paper.

Author Contributions: The authors of this paper jointly conceived the idea for this paper, discussed the agenda for the research, performed the theoretical and numerical calculations, and prepared each draft of the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BGS: Boltzmann–Gibbs–Shannon
THC: Tsallis–Havrda–Chavát

ME: maximum entropy
RU: Rodríguez-Lallena and Úbeda-Flores
FBST: Full Bayesian Significance Test

References

1. Tsallis, C. Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
2. Pougaza, D.B.; Mohammad-Djafari, A. Maximum entropies copulas. In Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Chamoni, France, 4–9 July 2011; Volume 1305, pp. 329–336.
3. Ellerman, D. Counting distinctions: On the conceptual foundations of Shannon's information theory. *Synthese* **2009**, *168*, 119–149.
4. Pougaza, D.B.; Mohammad-Djafari, A. New copulas obtained by maximizing Tsallis or Rényi entropies. In Proceedings of AIP Conference, Garching, Germany, 15–20 July 2012; p. 238.
5. Rodríguez-Lallena, J.A.; Úbeda-Flores, M. A new class of bivariate copulas. *Stat. Probab. Lett.* **2004**, *3*, 315–325.
6. Nelsen, R.B. *An Introduction to Copulas*; Springer: Berlin/Heidelberg, Germany, 2006.
7. De Bragança Pereira, C.A.; Stern, J.M. Evidence and credibility: Full Bayesian significance test for precise hypotheses. *Entropy* **1999**, *1*, 99–110.
8. 2011 Sexually Transmitted Disease Surveillance. Available online: http://www.cdc.gov/std/stats11/ (accessed on 15 July 2016).
9. Table 14. Gonorrhea—Reported Cases and Rates by State/Area and Region in Alphabetical Order, United States and Outlying Areas, 2007–2011. Available online: http://www.cdc.gov/std/stats11/tables/14.htm (accessed on 15 July 2016).
10. Distribution of the Total Population by Federal Poverty Level (above and below 400% FPL). Available online: http://kff.org/other/state-indicator/population-up-to-400-fpl/ (accessed on 15 July 2016).
11. Madruga, M.R.; Esteves, L.G.; Wechsler, S. On the bayesianity of Pereira-Stern tests. *Test* **2001**, *10*, 291–299.