Some new properties of Quasi-copulas

Roger B. Nelsen
Department of Mathematical Sciences, Lewis & Clark College, Portland, Oregon, USA.

José Juan Quesada Molina
Departamento de Matemática Aplicada, Universidad de Granada, Granada, Spain.

José Antonio Rodríguez Lallena
Manuel Úbeda Flores
Departamento de Estadística y Matemática Aplicada, Universidad de Almería, Almería, Spain.

ABSTRACT

Every multivariate distribution function with continuous marginals can be represented in terms of an unique n-copula, that is, in terms of a distribution function on $[0,1]^n$ with uniform marginals. The notion of quasi-copula was introduced in [1] by Alsina *et al.* (1993) and was used by these authors and others to characterize operations on distribution functions that can or cannot be derived from operations on random variables. In [2], Genest *et al.* (1999) characterize the concept of quasi-copula in simpler operational terms. We now present a new simple characterization and some nice properties of these functions, all of them concerning the measure of a quasi-copula. We show that the features of that measure can be quite different to the ones corresponding to measures associated to copulas.

AMS 1991 *Subject Classifications*: Primary 60E05; Secondary 62H05, 62E10. (?)

Keywords: Quasi-copulas, Lipschitz condition, Copulas.

1 Introduction.

The term "copula", coined in [6] by Sklar (1959), is now common in the statistical literature. The importance of copulas as a tool for statistical analysis and modelling stems largely from the observation that the joint distribution H of a set of $n \geq 2$ random variables X_i with marginals F_i can be expressed in the form

$$H(x_1, x_2, ..., x_n) = C\{F_1(x_1), F_2(x_2), ..., F_n(x_n)\}$$

in terms of a copula C that is uniquely determined on the set $\mathrm{Ran}F_1 \times ... \times \mathrm{Ran}F_n$. For a detailed study we refer to [3].

In [1], Alsina *et al.* (1993) introduced recently the notion of "quasi-copula" in order to show that a certain class of operations on univariate distribution functions is not derivable from corresponding operations on random variables defined on the same probability space. The same concept was also used in [5] by Nelsen *et al.* (1996) to characterize, in a given class of operations on distribution functions, those that do derive from corresponding operations on random variables.

In [2], Genest *et al.* (1999) characterized the concept of quasi-copula in two different ways. The first one states that a function $Q : \mathbf{I}^2 \longrightarrow \mathbf{I}$ ($\mathbf{I}$=[0,1]) is a quasi-copula if, and only if, it satisfies (i) $Q(0,x) = Q(x,0) = 0$ and $Q(x,1) = Q(1,x) = x$ for all $0 \leq x \leq 1$; (ii) $Q(x,y)$ is non-decreasing in each of its arguments; (iii) The Lipschitz condition, $|Q(x_1,y_1) - Q(x_2,y_2)| \leq |x_1 - x_2| + |y_1 - y_2|$ for all x_1, x_2, y_1 and y_2 in $\mathbf{I}$.

Nelsen *et al.* (2000) have developed in [4] a method to find best-possible bounds on bivariate distribution functions with fixed marginals, when additional information of a distribution-free nature is known, by using quasi-copulas.

In this work, we present some new properties of quasi-copulas. In Section 2 we provide a new simple characterization of quasi-copulas. In Section 3 we prove some properties of the Q-volumes and, therefore, the signed measure associated to quasi-copulas. In the last section, we provide a result about approximation of quasi-copulas (in particular, about approximation of copulas) by quasi-copulas of a special type.

2 A new characterization of quasi-copulas.

The following theorem provides a new simple characterization of quasi-copulas in terms of the absolute continuity of their vertical and horizontal sections.

THEOREM 2.1. *Let $Q : \mathbf{I}^2 \longrightarrow \mathbf{I}$ be a function satisfying the boundary conditions $Q(t,0) = Q(0,t) = 0$, $Q(t,1) = Q(1,t) = t$ for every $t \in \mathbf{I}$. Then, Q is a quasi-copula if, and only if, for every x and y in $\mathbf{I}$, the functions $Q_x, Q_y : \mathbf{I} \longrightarrow \mathbf{I}$ defined by $Q_x(y) = Q_y(x) = Q(x,y)$ are absolutely continuous and satisfy that*

$$0 \leq \frac{\partial Q}{\partial x}(x,y), \frac{\partial Q}{\partial y}(x,y) \leq 1$$

for almost every x and y in $\mathbf{I}$, respectively.

Proof: First, we suppose that Q is a quasi-copula, and let x and y in $\mathbf{I}$. A well-known result of Real Analysis states that the Lipschitz condition satisfied by Q_x and Q_y is equivalent to the following: Q_x and Q_y are absolutely continuous and $|Q_x'(y)| \leq 1$ and $|Q_y'(x)| \leq 1$ a.e. in $\mathbf{I}$. Since both Q_x and Q_y are non-decreasing, we have more precisely that $0 \leq \frac{\partial Q}{\partial x}(x,y) \leq 1$ for almost all x in $\mathbf{I}$ and $0 \leq \frac{\partial Q}{\partial y}(x,y) \leq 1$ for almost all y in $\mathbf{I}$.

In the opposite direction the only thing to be proved is the non-decreasingness. Let x, x' in $\mathbf{I}$ such that $x < x'$. Since Q_y is absolutely continuous we have that $Q_y(x') - Q_y(x) = \int_x^{x'} Q'_y(t)dt \geq 0$, i.e., $Q(x', y) \geq Q(x, y)$. With a similar reasoning it is proved the non-decreasingness in the second variable. □

Theorem 2.1 is interesting in order to check easily whether a function Q satisfying the boundary conditions is a quasi-copula. For instance, now it is easy to show that the function $Q : \mathbf{I}^2 \longrightarrow \mathbf{I}$ defined by

$$Q(x,y) = \begin{cases} xy & \text{if } 0 \leq y \leq 1/4 \\ xy + (1/24)(4y-1)\sin(2\pi x) & \text{if } 1/4 \leq y \leq 1/2 \\ xy + (1/12)(1-y)\sin(2\pi x) & \text{if } 1/2 \leq y \leq 1 \end{cases}$$

is a quasi-copula (but it is not a copula: see [2]).

3 The Q-volume of a quasi-copula.

Let C be a copula, and $R = [x_1, x_2] \times [y_1, y_2]$ be any 2-box in $\mathbf{I}^2$. If $V_C(R)$ stands for the $C-$volume of R, that is, $V_C(R) = C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1)$. We know that there exists an unique measure μ_C on the σ-algebra of all Lebesgue measurable subsets of the unit square such that $\mu_C(R) = V_C(R)$. Such a measure is double stochastic, i.e., $\mu_C(A \times \mathbf{I}) = \mu_C(\mathbf{I} \times A) = \lambda(A)$ for every Lebesgue-measurable set $A \subset \mathbf{I}$, where λ is the Lebesgue measure on $\mathbf{I}$.

Now, if Q is a quasi-copula, we can define similarly a signed measure on the σ-algebra A. This measure is uniquely determined by its definition on the rectangles $R = [x_1, x_2] \times [y_1, y_2]$:

$$\mu_Q(R) = V_Q(R) = Q(x_2, y_2) - Q(x_2, y_1) - Q(x_1, y_2) + Q(x_1, y_1).$$

And again, $\mu_Q(A \times \mathbf{I}) = \mu_Q(\mathbf{I} \times A) = \lambda(A)$ for every Lebesgue-measurable set $A \subset \mathbf{I}$. We know that $0 \leq V_C(R) \leq 1$ if C is a copula. Now, if Q is a quasi-copula, what can be said about the bounds for the $Q-$volume of R?. The following result provides the answer to this question.

THEOREM 3.1. *Let Q be a quasi-copula, and $R = [x_1, x_2] \times [y_1, y_2]$ any 2-box in $\mathbf{I}^2$. Then, $-1/3 \leq V_Q(R) \leq 1$.*

Proof: The $Q-$volume of R is given by $V_Q(R) = Q(x_2, y_2) - Q(x_2, y_1) - Q(x_1, y_2) + Q(x_1, y_1)$. Since $Q(x_2, y_2) - Q(x_2, y_1) \leq y_2 - y_1 \leq 1$, and $-Q(x_1, y_2) + Q(x_1, y_1) \leq 0$, we obtain that $V_Q(R) \leq 1$. If one of the equalities $x_1 = 0$, $x_2 = 1$, $y_1 = 0$, $y_2 = 1$ holds we know (see [2]) that $V_Q(R) \geq 0$. Thus, let us suppose that $0 = x_0 < x_1 < x_2 < x_3 = 1$

and $0 = y_0 < y_1 < y_2 < y_3 = 1$. Now, we divide $\mathbf{I}^2$ into 9 rectangles, namely $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, $i, j = 1, 2, 3$. So, $R = R_{22}$. Let $V_Q(R_{ij}) = v_{ij}$ $(i, j = 1, 2, 3)$. We know that

1) $v_{ij} \geq 0$ if $(i, j) \neq (2, 2)$,

2) $v_{12} + v_{22} \geq 0$, $v_{22} + v_{32} \geq 0$.

If $v_{22} < -1/3$, then $v_{12} \geq -v_{22} > 1/3$ and $v_{32} \geq -v_{22} > 1/3$. Hence $x_1 > 1/3$ and $1 - x_2 > 1/3$, which implies that $x_2 - x_1 < 1/3$. On the other hand $v_{22} = x_2 - x_1 - v_{23} - v_{21}$. Since $v_{23} \leq \min\{1 - y_2, x_2 - x_1\}$ and $v_{21} \leq \min\{y_1, x_2 - x_1\}$ we have that $v_{22} \geq x_2 - x_1 - \min\{1 - y_2, x_2 - x_1\} - \min\{y_1, x_2 - x_1\} \geq x_2 - x_1 - (x_2 - x_1) - (x_2 - x_1) = -(x_2 - x_1) > -1/3$. Thus, we get a contradiction. Whence $v_{22} \geq -1/3$ and the proof is complete. □

The following theorem completes the previous one:

THEOREM 3.2. *Let Q be a quasi-copula, and $R = [x_1, x_2] \times [y_1, y_2]$ any 2-box in $\mathbf{I}^2$. Then $V_Q(R) = 1$ if and only if $R = \mathbf{I}^2$, and $V_Q(R) = -1/3$ implies that $R = [1/3, 2/3]^2$.*

Proof: It is immediate that $V_Q(\mathbf{I}^2) = 1$. Moreover, it is clear that $V_Q(R) \leq \min\{x_2 - x_1, y_2 - y_1\}$; thus, $V_Q(R) < 1$ if $R \neq \mathbf{I}^2$. Now, suppose that $V_Q(R) = -1/3$. Then, as in Theorem 3.1, we can obtain that $x_1 \geq 1/3$ and $x_2 \leq 2/3$. Then, $x_2 - x_1 \leq 1/3$. If we use the notation of that theorem, we have that $v_{22} = -1/3 \geq x_2 - x_1 - \min\{1 - y_2, x_2 - x_1\} - \min\{y_1, x_2 - x_1\} \geq -(x_2 - x_1)$, which implies that $x_2 - x_1 \geq 1/3$. So, $x_2 - x_1 = 1/3$ and then $x_1 = 1/3$ and $x_2 = 2/3$. Similar reasonings yield that $y_1 = 1/3$, $y_2 = 2/3$, and the proof is complete. □

Of course, there exists quasi-copulas Q such that $V_Q([1/3, 2/3]^2) = -1/3$, as the following example shows.

EXAMPLE 3.1. Let s_1, s_2 and s_3 be three segments in $\mathbf{I}^2$, respectively defined by the following functions: $f_1(x) = x + 1/3$, $x \in [0, 2/3]$; $f_2(x) = x$, $x \in [1/3, 2/3]$; and $f_3(x) = x - 1/3$, $x \in [1/3, 1]$. Let us spread uniformly a mass of 2/3 on each of s_1 and s_3, and a mass of -1/3 on s_2. Let $(u, v) \in \mathbf{I}^2$. If we define $Q(u, v)$ as the mass spreaded on the rectangle $[0, u] \times [0, v]$, then it is easy to see that $Q : \mathbf{I}^2 \longrightarrow \mathbf{I}$ is a quasi-copula such that $V_Q([1/3, 2/3]^2) = -1/3$.

We have seen that a 2-box can take a negative $Q-$measure up to $-1/3$. However, if we choose a certain number of nonoverlapping 2-boxes, their $Q-$measure can be so negative as we wish, by choosing an appropiate quasi-copula. As a consequence, other 2-boxes will have a $Q-$measure so large as we wish. The quasi-copulas that we construct in the theorems of Section 4 show this fact.

Finally, we prove one more result which characterizes the rectangles with maximum area such that its Q-volume is zero for some quasi-copula Q. The result is exactly the same if we restrict ourselves to copulas.

THEOREM 3.3. *Let $R = [x_1, x_2] \times [y_1, y_2]$ be a 2-box in $\mathbf{I}^2$ such that $V_Q(R) = 0$ for some quasi-copula Q. Then, the maximum possible area for R is $1/4$; in this case, R must be a square.*

Proof: Let $x_0 = y_0 = 0$, $x_3 = y_3 = 1$, $a_{ij} = V_Q([x_{i-1}, x_i] \times [y_{j-1}, y_j])$, $i, j = 1, 2, 3$. Our hypothesis is that $a_{22} = 0$ and we are looking for the maximum possible value for $(x_2 - x_1)(y_2 - y_1)$. Moreover, we know that the remaining eight $a_{ij}'s$ are nonnegative (see [2]). Thus, we have that

$$a_{21} \leq y_1, \quad a_{23} \leq 1 - y_2, \quad a_{21} + a_{23} = x_2 - x_1,$$

$$a_{12} \leq x_1, \quad a_{32} \leq 1 - x_2, \quad a_{12} + a_{32} = y_2 - y_1.$$

These relations imply that that $x_2 - x_1 \leq 1 - y_2 + y_1$ and $y_2 - y_1 \leq 1 - x_2 + x_1$, i.e., $(x_2 - x_1) + (y_2 - y_1) \leq 1$.

So, it is clear that an upper bound for $(x_2 - x_1)(y_2 - y_1)$ is $(1/2)(1/2) = 1/4$. But we can reach this bound by taking

$$0 \leq x_1 < x_2 = x_1 + \frac{1}{2} \leq 1,$$

$$0 \leq y_1 < y_2 = y_1 + \frac{1}{2} \leq 1; \tag{1}$$

whence $a_{21} = y_1$, $a_{23} = 1 - y_2 = (1/2) - y_1$, $a_{12} = x_1$, $a_{32} = 1 - x_2 = (1/2) - x_1$ and $a_{11} = a_{13} = a_{31} = a_{33} = 0$. $\square$

We can construct not only proper quasi-copulas, but also copulas such that the measure associated to a square of side $1/2$ in $\mathbf{I}^2$ be equal to zero. If $R = [x_1, x_2] \times [y_1, y_2]$ satisfies conditions (1), we can consider, for instance, the copula whose mass is spreaded in the following manner: a mass of x_1, y_1, $(1/2) - x_1$, $(1/2) - y_1$ uniformly distributed on the respective segments which join the pair of points $\{(0, y_1), (x_1, y_1 + (1/2))\}$, $\{(x_1, 0), (x_1 + (1/2), y_1)\}$, $\{(x_1 + (1/2), y_1), (1, y_1 + (1/2))\}$ and $\{(x_1, y_1 + (1/2)), (x_1 + (1/2), 1)\}$.

4 Approximations of quasi-copulas.

We begin this Section proving that the copula Π can be approximated by a quasi-copula with so much negative mass as desired. Theorem 4.1 is a particular case of Theorem 4.2, but we include the first one to make easier the understanding of the second one.

THEOREM 4.1. *Let $\varepsilon, M > 0$. Then there exists a quasi-copula Q such that:*
(a) $\exists S \subset \mathbf{I}^2$ *satisfying $\mu_Q(S) < -M$ (μ_Q is the signed measure associated to Q).*
(b) $|Q(x, y) - \Pi(x, y)| < \varepsilon$ *for all x, y in $\mathbf{I}$.*

Proof: Let m be an odd number in $\mathbf{N}$ such that $m \geq 4/\varepsilon$ and $(m-1)^2/4m > M$. We divide $\mathbf{I}^2$ into m^2 squares, namely

$$R_{ij} = \left[\frac{i-1}{m}, \frac{i}{m}\right] \times \left[\frac{j-1}{m}, \frac{j}{m}\right],$$

for $i, j = 1, 2, ..., m$. Each R_{ij}, $i, j = 1, 2, ..., m$, is divided in the same manner into m^2 squares, namely

$$R_{ijkl} = \left[\frac{(i-1)m+k-1}{m^2}, \frac{(i-1)m+k}{m^2}\right] \times \left[\frac{(j-1)m+l-1}{m^2}, \frac{(j-1)m+l}{m^2}\right],$$

with $k, l = 1, 2, ..., m$.

We are going to define a signed measure on $\mathbf{I}^2$ in the following manner: Let $r = (m+1)/2$; for every (i,j) such that $i, j = 1, 2, ..., m$, we spread uniformly a mass of $1/m^3$ on the squares R_{ijkl}, with $1 \leq k \leq r$ and $l = r-k+1, r-k+3, ..., r+k-3, r+k-1$, and a mass of $-1/m^3$ on the squares R_{ijkl} with $2 \leq k \leq r$ and $l = r-k+2, r-k+4, ..., r+k-4, r+k-2$. The remaining squares R_{ijkl} with $1 \leq k \leq r$ have mass zero. We spread mass on the squares R_{ijkl} with $k > r$ symmetrically with respect $k = r$, i.e., the mass on R_{ijkl} $(k > r)$ coincides with the mass on the square $R_{ij(m+1-k)l}$.

Thus, the sum of the positive masses $1/m^3$ on each R_{ij} is

$$[2\{1+2+...+r-1\}+r]\left(\frac{1}{m^3}\right) = \frac{r^2}{m^3} = \frac{(m+1)^2}{4m^3},$$

and the sum of the negative masses $-1/m^3$ on each R_{ij} is $-(m-1)^2/4m^3$.

If $S = \bigcup\{R_{ijkl}|$ the mass spreaded on R_{ijkl} is $-1/m^3\}$, then we obtain that the mass spreaded on S is $-(m-1)^2/4m < -M$.

For every (x,y) in $\mathbf{I}^2$, let $Q(x,y)$ be the mass spreaded on $[0,x] \times [0,y]$. Then Proposition 3 in [2] implies that the function Q is a quasi-copula. Moreover, $V_Q(R_{ij}) = 1/m^2 = V_\Pi(R_{ij})$ for all (i,j). As a consequence, for every $i, j = 0, 1, 2, ..., m$, we have that $Q(i/m, j/m) = \Pi(i/m, j/m)$.

Now, let $(x,y) \in \mathbf{I}^2$. We have that $|x - i/m| < 1/m$ and $|y - j/m| < 1/m$ for some (i,j). Then, $|Q(x,y) - \Pi(x,y)| \leq |Q(x,y) - Q(i/m, j/m)| + |Q(i/m, j/m) - \Pi(i/m, j/m)| + |\Pi(i/m, j/m) - \Pi(x,y)| \leq 2|x - i/m| + 2|y - j/m| < 4/m \leq \varepsilon$. □

The following theorem generalizes the previous one to every quasi-copula.

THEOREM 4.2. *Let $\varepsilon, M > 0$, and Q a quasi-copula. Then, there exists a quasi-copula $\bar{Q}$ and a set $S \subset \mathbf{I}^2$ such that:*

(a) $\mu_{\bar{Q}}(S) < -M$.

(b) $|Q(x,y) - \bar{Q}(x,y)| < \varepsilon$ *for all* x, y *in* $\mathbf{I}$.

Proof: Let m, R_{ij} and R_{ijkl} be as in Theorem 4.1. Let $q_{ij} = V_Q(R_{ij})$ for all (i,j). We know (Proposition 3 in [2]) that $q_{ij} \geq 0$ whenever either i or j is equal either 1 or m. Observe that $\sum_{ij} q_{ij} = 1$.

Consider any square R_{ij}. We spread mass on the squares R_{ijkl} in a similar manner to the previous theorem, but taking q_{ij}/m and $-q_{ij}/m$ instead of $1/m^3$ and $-1/m^3$, respectively.

If $q_{ij} > 0$, the positive mass spreaded on R_{ij} is $r^2 q_{ij}/m = q_{ij}(m+1)^2/4m$, and the negative one is $-q_{ij}(m-1)^2/4m$.

If $q_{ij} = 0$ no mass is spreaded on R_{ij}.

If $q_{ij} < 0$, the positive and negative masses spreaded on R_{ij} are, respectively, $-q_{ij}(m-1)^2/4m$ and $q_{ij}(m+1)^2/4m$.

For every (x, y) in $\mathbf{I}^2$, let $\bar{Q}(x, y)$ be the mass spreaded on $[0, x] \times [0, y]$. Then Proposition 3 of [2] implies that the function $\bar{Q}$ is a quasi-copula (since Q is a 2-quasi-copula and again by using Proposition 3 in [2]).

The whole negative mass considered to define $\bar{Q}$ is

$$\sum_{q_{ij}<0} \frac{(m+1)^2}{4m} q_{ij} + \sum_{q_{ij}>0} -\frac{(m-1)^2}{4m} q_{ij} \leq -\frac{(m-1)^2}{4m} \sum q_{ij} = -\frac{(m-1)^2}{4m} < -M.$$

And similar reasonings to those showed in Theorem 4.1 yields that $|Q(x, y) - \bar{Q}(x, y)| < \varepsilon$ for all (x, y) in $\mathbf{I}^2$. □

References

[1] Alsina, C.; Nelsen, R. B.; and Schweizer, B., (1993), "On the characterization of a class of binary operations on distribution functions", *Statist. Probab. Lett.*, **17**, 85-89.

[2] Genest, C.; Quesada Molina, J. J.; Rodríguez Lallena, J. A.; and Sempi, C., (1999), "A Characterization of Quasi-copulas", *J. Multivariate Anal.* **69**, 193-205.

[3] Nelsen, R. B., (1999), *An Introduction to Copulas*, Springer-Verlag, New York.

[4] Nelsen, R. B.; Quesada Molina, J. J.; Rodríguez Lallena, J. A.; and Úbeda Flores, M., (2000), "Best-possible Bounds on Sets of Bivariate Distribution Functions", to appear.

[5] Nelsen, R. B.; Quesada Molina, J. J.; Schweizer, B.; and Sempi, C., (1996), "Derivability of some operations on distribution functions", in *Distributions with Fixed Marginals and Related Topics*, (L. Rüschendorf, B. Schweizer, M. D. Taylor, Eds.), IMS Lecture Notes-Monograph Series Number 28, pp. 233-243, Hayward, CA.

[6] Sklar, A. (1959). "Fonctions de répartition à n dimensions et leurs marges", *Publ. Inst. Statist. Univ. Paris* **8**, 229-231.