UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

POSGRADO EN CIENCIAS MATEMATICAS

FACULTAD DE CIENCIAS

DIAGONAL PROPERTIES OF THE EMPIRICAL COPULA AND APPLICATIONS. CONSTRUCTION OF FAMILIES OF COPULAS WITH GIVEN RESTRICTIONS.

T E S I S

QUE PARA OBTENER EL GRADO ACADEMICO DE

DOCTOR EN CIENCIAS

PRESENTA

ARTURO ERDELY RUIZ

DIRECTOR DE TESIS:

DR. JOSE MARIA GONZALEZ - BARRIOS MURGUIA

MEXICO, D. F. ENERO DE 2007

Agradezco profundamente a los miembros del jurado su tiempo y paciencia en la revisión de la presente tesis doctoral:

Dra. M.A. Begoña Fernández Fernández

Dr. José María González-Barrios Murguía (Director de Tesis)

Dr. Roger B. Nelsen

Dr. José Juan Quesada Molina

Dra. Beatriz Vaz de Melo Mendes

Contents

Introducción en español

El capítulo 1 comienza con tres secciones que resumen algunos conceptos y resultados importantes en relación a cópulas, haciendo énfasis en cópulas diagonales y arquimedianas. Luego se incluye una sección para introducir la *Ecuación Funcional de Schröder* y la relación que guarda con cópulas arquimedianas de acuerdo a un resultado de Frank (1996). El capítulo termina discutiendo algunas ideas y conceptos sobre concordancia, dependencia y cópulas.

El capítulo 2, inspirado en el hecho de que bajo la *condición de Frank* toda la información sobre una cópula arquimediana está contenida en su sección diagonal, se exploran algunas propiedades de la sección diagonal de la cópula empírica: cotas, incrementos, identificación y conteo de todas las distintas trayectorias que puede seguir la diagonal empírica, y finalmente, la distribución exacta de la diagonal empírica bajo la hipótesis de independencia, para los casos bidimensional y tridimensional. Este último resultado abrió la puerta hacia una propuesta de prueba no paramétrica de independencia, bajo el supuesto de que la cópula subyacente pertenece a la familia arquimediana, ya que la cópula que representa la independencia es del tipo arquimediano y satisface la condición de Frank, y por lo tanto es la única cópula arquimediana que tiene sección diagonal $\delta(u) = u^2$.

Después de los resultados del capítulo 2, una pregunta natural es si fuera de la familia arquimediana existen cópulas absolutamente continuas con la misma diagonal que la cópula que representa la independencia, pero diferentes de ella fuera de la diagonal. La respuesta es en el sentido positivo, y en el capítulo 3 se construye una amplia familia de cópulas absolutamente continuas con sección diagonal dada, que pueden diferir the otra cópula absolutamente continua casi en todas partes respecto a la medida de Lebesgue. Es importante estar consciente de esto, en caso de que se desee utilizar una prueba no paramétrica de independencia basada en la sección diagonal, pero fuera de la familia arquimediana.

En el capítulo 4 se da solución a un problema propuesto por Alsina, Frank y Schweizer (2003), mismo que reaparece en Alsina, Frank y Schweizer (2006):

> Es posible diseñar una prueba de independencia estadística basada en los supuestos de que la cópula subyacente es arquimediana y que su sección diagonal es $\delta(u) = u^2$?

Se propone una prueba no paramétrica de independencia por medio de un estadístico bivariado basado en la diagonal empírica y en su distribución exacta obtenida en el capítulo 2, lo cual permite obtener la distribución exacta de cualquier estadístico de prueba basado en la diagonal empírica. Se llevó a cabo un estudio de simulación para comparar la potencia de la prueba propuesta versus algunas bien conocidas puebas no paramétricas de la literatura estadística: Spearman, Blum-Kiefer-Rosenblatt, Kallenberg-Ledwina, bajo tres clases de cópulas, dos del tipo arquimediano y una tercera no arquimediana. Incluso se hace una comparación de la prueba propuesta versus la prueba localmente más potente basada en rangos.

Finalmente, en las conclusiones se hacen algunas observaciones y se discuten algunos problemas abiertos.

Introduction

Chapter 1 begins with three sections which summarize some important concepts and results regarding copulas, making emphasis on diagonal and Archimedean copulas. Then we include one section to introduce *Schröder's Functional Equation* and its relationship with Archimedean copulas via a result by Frank (1996). We end this chapter discussing some ideas and concepts around concordance, dependence, and copulas.

In Chapter 2, inspired in the fact that under *Frank's condition* all the information about an Archimedean copula is contained in its diagonal section, we explore some properties of the diagonal section of the empirical copula: bounds, one-step increments, labeling and counting all the different paths an empirical diagonal may follow, and finally, the exact distribution of the empirical diagonal under the hypothesis of independence, for the two-dimensional and three-dimensional cases. This last result opened the door toward a proposal of a nonparametric test for independence, under the assumption that the underlying copula belongs to the Archimedean family, since the copula that represents independence is of the Archimedean type and satisfies Frank's condition, and so it is the unique Archimedean copula with diagonal section $\delta(u) = u^2$.

After the results of Chapter 2, a natural question was if outside the Archimedean family there exist absolutely continuous copulas with the same diagonal section as the independence copula, but different from it outside the diagonal? The answer is in the positive sense, and in Chapter 3 we build a broad family of absolutely continuous copulas with a fixed diagonal, which can differ from another absolutely continuous copula almost everywhere with respect to Lebesgue measure. It is important to be aware of this, in case a nonparametric test for independence based on the diagonal section is used outside the Archimedean family.

In Chapter 4 we solve an open problem proposed by Alsina, Frank and Schweizer (2003), which again appeared in Alsina, Frank and Schweizer (2006):

> Can one design a test of statistical independence based on the assumptions that the copula in question is Archimedean and that its diagonal section is $\delta(u) = u^2$?

We propose a nonparametric test for independence via a bivariate statistic based on the empirical diagonal and its exact distribution obtained in Chapter 2, which allows to obtain the exact distribution of any test statistic based on the empirical diagonal. A simulation study is performed to compare the power of the proposed test against some well-known nonparametric tests in the statistical literature: Spearman, Blum-Kiefer-Rosenblatt, Kallenberg-Ledwina, for three classes of copulas, two of the Archimedean type and a third one non Archimedean. We also made a comparison of the proposed test's power against the locally most powerful rank test.

Finally, in the conclusions chapter we make some remarks and discuss some open problems.

Chapter 1

Preliminaries

1.1 Copulas: basic facts.

Let f and g be univariate probability density functions, that is $f, g \geq 0$ and $\int_{\mathbb{R}} f(x)\,dx = 1$, $\int_{\mathbb{R}} g(y)\,dy = 1$, and let F and G be their corresponding probability distribution functions, respectively, that is $F(x) = \int_{-\infty}^{x} f(t)\,dt$ and $G(y) = \int_{-\infty}^{y} g(t)\,dt$. If we define the parametric family of bivariate functions $h_\theta : \mathbb{R}^2 \to \mathbb{R}$ such that

$$h_\theta(x,y) := f(x)g(y)\big[\,1 + \theta(1-2F(x))(1-2G(y))\,\big]\,, \tag{1.1}$$

it is straightforward to verify that $\{h_\theta : -1 \leq \theta \leq 1\}$ is a family of bivariate probability density functions with marginal densities always equal to f and g, that is $h_\theta \geq 0$ and

$$\iint_{\mathbb{R}^2} h_\theta(x,y)\,dxdy = 1\;, \qquad \int_{\mathbb{R}} h_\theta(x,y)\,dy = f(x)\;, \qquad \int_{\mathbb{R}} h_\theta(x,y)\,dx = g(y)\,.$$

This is a typical example to show that knowledge of the marginal distributions of a random vector is not enough to assess the joint density function, even though marginal distributions are obtainable from the joint density. (1.1) is known as the *Farlie-Gumbel-Morgenstern family* of distributions and was discussed by Morgenstern (1956), Gumbel (1958, 1960), and Farlie (1960); however, according to Nelsen (2006a) it seems that the earliest publication of (1.1) is due to Eyraud (1938). According to Kotz and Seeger (1991):

> In the last decade [1980s] it has become increasingly important to consider dependence as more than an antithesis to independence, the latter being the basic concept of mathematical probability theory. As a result, several methods have been developed to impose dependence among random variables with given marginal distributions. The majority of bivariate methods are based on a well-known result due

to Hoeffding (1940) and Fréchet (1951) which says that given any two random variables, X and Y, with respective c.d.f.s. [cumulative distribution functions] $F_1(x)$ and $F_2(y)$, the class $\Psi(F_1, F_2) = \{H(x,y)\,|\,H$ is a bivariate c.d.f. with marginals $F_1(x)$ and $F_2(y)\}$ contains an upper bound, H^*, and a lower bound H_*. These are bounds with respect to the partial ordering $\prec$, [a transitive and antisymmetric relation] denoting stochastic dominance, that is if $H, H' \in \Psi(F_1, F_2)$ then $H \prec H'$ iff [if and only if] $H(x,y) \leq H'(x,y)\ \forall(x,y)$. Moreover, the so-called Fréchet bounds have general expressions in terms of F_1 and F_2, namely

$$H_*(x,y) = \max\{F_1(x) + F_2(y) - 1, 0\}\,, \tag{1.2}$$

$$H^*(x,y) = \min\{F_1(x), F_2(y)\}\,. \tag{1.3}$$

[...] Several parametrized subsets of $\Psi(F_1, F_2)$ which are linearly ordered with respect to $\prec$ [that is for all $H, H' \in \Psi(F_1, F_2)$ either $H \prec H'$, $H' \prec H$, or $H = H'$] have appeared in the literature [...] there are the Farlie-Gumbel-Morgenstern, among others [...] were constructed according to the viewpoint that the way to impose dependence is to increase (or decrease) everywhere the independent c.d.f. $F_1(x)F_2(y)$ without altering the marginals, thus creating a new c.d.f. closer in value to H^* (or H_*) [...]

Fréchet (1951, 1957) and Féron (1956) made important contributions to the question of determining the relationship between a multidimensional probability distribution function and its lower dimensional margins. An effective answer to this question emerged as a result of the collaboration between Abe Sklar and Berthold Schweizer. According to Schweizer (1991) their collaboration began in the context of probabilistic metric spaces in 1957. By 1958 they had made some progress and submitted a note describing their results to M. Fréchet, and an exchange of letters began with him. In one of them, Fréchet raised this question about determining the relationship between a multidimensional probability distribution function and its lower dimensional margins.

Abe Sklar answered this question for one-dimensional margins. For example, let (X,Y) be a random vector with joint distribution function $H(x,y)$, then the marginal distribution functions of X and Y are $F(x) := H(x,\infty)$ and $G(y) := H(\infty,y)$, respectively. Sklar (1959) proved that there exists a function C, which he called *copula*, which links the joint distribution function to its marginals:

$$H(x,y) = C\big(F(x), G(y)\big)\,. \tag{1.4}$$

Before giving a formal definition of a copula, we may get some motivation using well-known properties for bivariate probability distribution functions:

$$H(\infty,\infty)=1\,,\quad H(x,\infty)=F(x)\,,\quad H(\infty,y)=G(y)\,,\quad H(x,-\infty)=0=H(-\infty,y)\,,$$

and for every real numbers x_1,x_2,y_1,y_2 where $x_1\leq x_2$ and $y_1\leq y_2$ we have

$$H(x_2,y_2)-H(x_2,y_1)-H(x_1,y_2)+H(x_1,y_1)=\mathbb{P}\big\{(X,Y)\in[\,x_1,x_2\,]\times[\,y_1,y_2\,]\big\}\geq 0\,.$$

Using the above properties in combination with (1.4) we observe that

$$C(F(x),1)=F(x)\,,\quad C(1,G(y))=G(y)\,,\quad C(F(x),0)=0=C(0,G(y))\,,$$

$$C(F(x_2),G(y_2))-C(F(x_2),G(y_1))-C(F(x_1),G(y_2))+C(F(x_1),G(y_1))\geq 0\,,$$

for every real numbers x_1,x_2,y_1,y_2 where $x_1\leq x_2$ and $y_1\leq y_2$. If we define $u:=F(x)$ and $v:=G(y)$, we arrive to the following definition, which may be found in Schweizer and Sklar (2005) or Nelsen (2006a):

1.1. Definition. A *bivariate copula* is a function $C:[\,0,1\,]^2\to[\,0,1\,]$ with the following properties:

1. For every $u,\ v$ in $[\,0,1\,]$
$$C(u,0)=0=C(0,v) \tag{1.5}$$
and
$$C(u,1)=u\,,\;C(1,v)=v\,; \tag{1.6}$$

2. For every u_1,u_2,v_1,v_2 in $[\,0,1\,]$ such that $u_1\leq u_2$ and $v_1\leq v_2$
$$C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1)\geq 0\,. \tag{1.7}$$

Even though we gave a probabilistic motivation for the above definition of copula, it is important to notice that, strictly speaking, Definition 1.1 does not involve probabilistic concepts at all, it is just about a real valued function with domain the unit square which satisfies boundary conditions (1.5) and (1.6), and condition (1.7) which is called *2-increasing*. The same comment applies to the formal version of *Sklar's Theorem*, as we will see later.

It follows that C is nondecreasing in each variable (let $v_1=0$ or $u_1=0$ in 1.7) and uniformly continuous (since Definition 1.1 implies that C satisfies the Lipschitz condition $|C(u_2,v_2)-C(u_1,v_1)|\leq|u_2-u_1|+|v_2-v_1|$), for details of these and the following properties and definitions in this section see Nelsen (1995, 2006a) or Schweizer and Sklar (2005). An immediate consequence of these two properties is that the horizontal, vertical and diagonal sections of a copula C are all nondecreasing and uniformly continuous on $[0,1]$ where the *horizontal section, vertical section* and *diagonal section* are functions from $[0,1]$ to $[0,1]$ given by $t \mapsto C(t,a)$, $t \mapsto C(a,t)$, and $\delta_C(t) = C(t,t)$, respectively, with a fixed in $[0,1]$.

1.2. Theorem. *Let C be a copula. Then for every (u,v) in $[0,1]^2$*

$$\max(u+v-1,0) \leq C(u,v) \leq \min(u,v)\,. \tag{1.8}$$

It is straightforward to verify that the bounds in (1.8) are themselves copulas and are commonly denoted by $M(u,v) := \min(u,v)$ and $W(u,v) := \max(u+v-1,0)$. Thus for every copula C and every (u,v) in $[0,1]^2$ we have

$$W(u,v) \leq C(u,v) \leq M(u,v)\,. \tag{1.9}$$

The above inequality is the copula version of the *Fréchet-Hoeffding bounds* inequalities (1.2) and (1.3), which we shall encounter later in terms of distribution functions, after stating formally Sklar's theorem. We refer to M as the *Fréchet-Hoeffding upper bound* and W as the *Fréchet-Hoeffding lower bound.* A third important copula that we will frequently deal with is the *product copula* $\Pi(u,v) := uv$.

1.3. Definition. A *distribution function* is a function F with domain the extended real line (that is in $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$) such that

1. F is nondecreasing,
2. $F(-\infty) = 0$ and $F(\infty) = 1$.

1.4. Definition. A *bivariate joint distribution function* is a function H with domain the extended real plane (that is in $\overline{\mathbb{R}}^2 := \overline{\mathbb{R}} \times \overline{\mathbb{R}}$) such that

1. H is 2-increasing, that is for all x_1, x_2, y_1, y_2 in $\overline{\mathbb{R}}$ with $x_1 \leq x_2$ and $y_1 \leq y_2$

$$H(x_2,y_2) - H(x_2,y_1) - H(x_1,y_2) + H(x_1,y_1) \geq 0\,,$$

2. $H(x,-\infty) = 0 = H(-\infty,y)$ and $H(\infty,\infty) = 1$.

Here we state an important quotation from Nelsen (2006a):

> Note that there is nothing "probabilistic" in these definitions of distribution functions. Random variables are not mentioned, nor is left-continuity or right-continuity. All the [probability] distribution functions of one or two random variables usually encountered in statistics satisfy either the first or the second of the above definitions. Hence any results we derive for such distribution functions will hold when we discuss random variables, regardless of any additional restrictions that may be imposed.

1.5. Definition. The *margins* of a bivariate joint distribution function H are the functions F and G given by $F(x) := H(x, \infty)$ and $G(y) := H(\infty, y)$.

It is an immediate consequence of the above definitions that the margins of H are themselves distribution functions.

1.6. Theorem. ***Sklar (1959)*** *Let H be a bivariate joint distribution function with margins F and G. Then there exists a copula C such that for all x, y in $\overline{\mathbb{R}}$*

$$H(x, y) = C\big(F(x), G(y)\big). \tag{1.10}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $Ran\, F \times Ran\, G$. Conversely, if C is a copula and F and G are distribution functions, then the function H defined by (1.10) is a bivariate joint distribution function with margins F and G.

Notation: $Ran\, F := \{z \,:\, z = F(x) \text{ for some } x\}$. For details of the proof see either Sklar (1959, 1996a), Schweizer and Sklar (2005) or Nelsen (2006a). For a different proof, see Carley and Taylor (2002).

1.7. Definition. Let F be a distribution function. Then a *quasi-inverse* of F is any function $F^{(-1)}$ with domain $[\,0, 1\,]$ such that

1. If t is in $Ran\, F$, then $F^{(-1)}(t)$ is any number x in $\overline{\mathbb{R}}$ such that $F(x) = t$, that is for all t in $Ran\, F$
$$F\big(F^{(-1)}(t)\big) = t\,;$$
2. If t is not in $Ran\, F$, then
$$F^{(-1)}(t) = \inf\{x \,|\, F(x) \geq t\} = \sup\{x \,|\, F(x) \leq t\}\,.$$

If F is strictly increasing, then it has but a single quasi-inverse. Moreover, $F^{(-1)} = F^{-1}$ where F^{-1} is the usual inverse.

1.8. Corollary. *Let H be a bivariate joint distribution with continuous margins F and G, and let C be the unique copula such that (1.10) holds. Then for any (u,v) in $[0,1]^2$*

$$C(u,v) \;=\; H\big(F^{(-1)}(u), G^{(-1)}(v)\big)\,. \tag{1.11}$$

For the particular case of random variables, if X and Y are continuous random variables with marginal probability distribution functions F and G, respectively, and with joint probability distribution function H, Sklar's theorem implies that there exists a unique copula, which we may denote C_{XY}, such that $H(x,y) = C_{XY}\big(F(x), G(y)\big)$. Moreover, if C is any copula and F_1 and G_1 are (marginal) probability distribution functions, then $C\big(F_1(x), G_1(y)\big)$ is indeed a joint probability distribution function. It is important to observe what was mentioned by Mikusiński *et al* (1991):

> The copula C contains valuable information about the type of dependence that exists between random variables having C as their copula [...] One might think of a copula as a canonical representative of all distributions H that correspond to random variables X and Y which have a specific sort of relationship to each other.

Given a specific joint probability distribution function H of two continuous random variables with (marginal) probability distributions F and G, we may "extract" the associated copula C_{XY} using Corollary 1.8, and then build a new joint probability distribution H_1 with the same copula but different (marginal) probability distribution functions F_1 and G_1, that is $H_1(x,y) = C_{XY}\big(F_1(x), G_1(y)\big)$. So, for example, the exercise of building a bivariate distribution with standard normal margins that is not the standard bivariate normal becomes trivial. Also, we may extract the associated copula from the bivariate normal distribution (called *Gaussian copula*) and use it to build a new bivariate distribution with non-normal margins.

Here we have to be careful, for example we have a warning remark from Marshall (1996):

> The marginals F and G can be inserted into any copula, so they carry no direct information about the coupling; at the same time, any pair of marginals can be inserted into C, so C carries no direct information about the marginals. This being the case, it may seem reasonable to expect that the connections between the marginals of H are determined by C alone, and any question about these connections can be answered with knowledge of C alone.
>
> Of course, things are not that simple. Some problems stem from the fact that copulas are not unique when at least one marginal is discontinuous. In fact, the

> marginals can sometimes play a significant role as the copula in determining the way in which they are coupled in H; in the extreme case, degenerate marginals by themselves determine the joint distribution and any copula can be used. But interaction between the copula and the marginals is often critical; copulas can look quite different in different parts of their domain, and the relevant part is determined by the range of the marginals.
> [...] Much of the literature regarding copulas has been based upon the assumption that the marginals of H are continuous because this is a necessary and sufficient condition for the copula of H to be unique.

But even in the context of continuous random variables, it is tempting to consider Sklar's theorem as an infinite source for building multivariate distributions of all kinds by just choosing continuous margins and plugging them into any desired copula. Somehow, Marshall and Olkin (1967) make emphasis on the importance of doing some additional work when building a multivariate distribution:

> [...] The family of solutions
>
> $$H(x,y) = F(x)G(y)\{1-\alpha[1-F(x)][1-G(y)]\}\,, \quad |\alpha| \leq 1\,, \tag{1.12}$$
>
> due to Morgenstern (1956) has been studied by Gumbel (1960) when F and G are exponential. Gumbel also studied the bivariate distribution
>
> $$H(x,y) = 1 - e^{x} - e^{y} + e^{-x-y-\delta xy}\,, \quad 0 \leq \delta \leq 1\,, \tag{1.13}$$
>
> which has exponential marginals. However, we know of no model or other basis for determining how these distributions might arise in practice.

By the way, for continuous margins F and G and applying Sklar's theorem to (1.12) we immediately identify the underlying (parametric) family of copulas

$$C_\theta(u,v) = uv\big[\,1-\theta(1-u)(1-v)\,\big]\,, \quad |\theta| \leq 1\,, \tag{1.14}$$

known as the *Farlie-Gumbel-Morgenstern family* of copulas.

It would be nice to have a huge catalog of copulas which specifies the sort of dependence or probabilistic interpretation each copula captures. An immediate consequence of Sklar's theorem for a random vector (X,Y) of continuous random variables is that the product copula $\Pi(u,v) = uv$ is the copula of (X,Y) if and only if X and Y are independent. Fréchet

(1951) proved that $M(u,v) = \min(u,v)$ is the copula for (X,Y) if and only if X and Y are almost surely increasing functions of each other, and $W(u,v) = \max(u+v-1,0)$ is the copula for (X,Y) if and only if X and Y are almost surely decreasing functions of each other. Random variables with copula M are often called *comonotonic,* and those with copula W are often called *countermonotonic.* Mikusiński *et al* (1991) give probabilistic interpretations of other types of copulas, such as *Shuffles of Min, Hairpins,* and convex sums of copulas.

If U and V are continuous uniform random variables on $(0,1)$ then by Sklar's theorem we have that their joint probability distribution function H restricted to $[\,0,1\,]^2$ equals the associated copula C_{UV}. A copula is itself a bivariate distribution with uniform margins on $[\,0,1\,]$. So, as stated by Nelsen (2006a):

> [...] each copula C induces a probability measure on $[\,0,1\,]^2$ [...] Hence, at a intuitive level, the C-measure of a subset of $[\,0,1\,]^2$ is the probability that two continuous uniform $(0,1)$ random variables U and V with joint distribution C assume values in that subset.

1.9. Definition. For any copula C let $C(u,v) = A_C(u,v) + S_C(u,v)$, where

$$A_C(u,v) := \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s,t)\, dt ds\,, \tag{1.15}$$

$$S_C(u,v) := C(u,v) - A_C(u,v)\,. \tag{1.16}$$

If $C \equiv A_C$ on $[\,0,1\,]^2$ — that is, if considered as a joint distribution function, C has a joint density (usually referred as the *copula density*) given by $\partial^2 C(u,v)/\partial u \partial v$ — then C is *absolutely continuous,* whereas if $C \equiv S_C$ on $[\,0,1\,]^2$ — that is, if $\partial^2 C(u,v)/\partial u \partial v = 0$ almost everywhere in $[\,0,1\,]^2$ — then C is *singular.* Otherwise, C has an *absolutely continuous component* A_C and a *singular component* S_C (in this case neither A_C or S_C is a copula because neither has uniform $(0,1)$ margins). The C-measure of the absolutely continuous component is $A_C(1,1)$, and the C-measure of the singular component is $S_C(1,1)$. The *support of a copula* is the complement of the union of all open subsets of $[\,0,1\,]^2$ with C-measure equal to zero.

Most of the previous definitions and results are extended to the multivariate case, for details see Schweizer and Sklar (2005) and Nelsen (2006a), so we will just point out those which do not, as well as some issues that do not arise in the bivariate case.

The extensions of the bivariate copulas M, Π, and W to n dimensions are given by:

$$\begin{aligned} M^{(n)}(u_1,\ldots,u_n) &:= \min(u_1,\ldots,u_n)\,; \\ \Pi^{(n)}(u_1,\ldots,u_n) &:= u_1 u_2 \cdots u_n\,; \\ W^{(n)}(u_1,\ldots,u_n) &:= \max(u_1+u_2+\cdots+u_n-n+1,0)\,. \end{aligned} \tag{1.17}$$

$M^{(n)}$ and $\Pi^{(n)}$ are multivariate copulas (or n-copulas) for all $n \geq 2$, but $W^{(n)}$ fails to be an n-copula for any $n > 2$. However, we still have the n-dimensional version of the Fréchet-Hoeffding bounds (1.9): If C is any n-copula, then for every $\mathbf{u} = (u_1,\ldots,u_n)$ in $[0,1]^n$

$$W^{(n)}(\mathbf{u}) \leq C(\mathbf{u}) \leq M^{(n)}(\mathbf{u})\,. \tag{1.18}$$

Although the Fréchet-Hoeffding lower bound $W^{(n)}$ is never a copula for $n > 2$, the above inequality cannot be improved, see Nelsen (2006a):

1.10. Theorem. *For any $n \geq 3$ and any fixed $\mathbf{u} = (u_1,\ldots,u_n)$ in $[0,1]^n$, there exists an n-copula $C_{\mathbf{u}}$, which depends on $\mathbf{u}$, such that $C_{\mathbf{u}}(\mathbf{u}) = W^{(n)}(\mathbf{u})$.*

For $n = 2$ we may just talk about univariate margins, which by definition (1.1) are the identity function, that is $C(u,1) = u$ and $C(1,v) = v$. For $n \geq 3$ and n-copulas, the concept of m-margins is introduced, for $2 \leq m \leq n$, defining an m-variate function by setting $n-m$ of the arguments of C equal to 1. An n-copula thus has $\binom{n}{m}$ m-margins. It is straightforward to show that each m-margin of C is an m-copula. In the other direction, however, $\binom{n}{m}$ given m-copulas are not necessarily the m-margins of an n-copula; if they are, then the m-copulas are said to be *compatible.* Whether certain given m-copulas may or may not be m-margins of higher dimension copulas has become known as *the compatibility problem.* Sklar (1996a) gives some examples of *incompatibility.* There is work done on necessary and sufficient conditions for compatibility in 3-copulas: Dall'Aglio (1959, 1960, 1972), Quesada-Molina and Rodríguez-Lallena (1994), Joe (1997). For some cases in higher dimensions see Joe (1996, 1997) .

1.2 Diagonal copulas.

We begin this section with some definitions and lemmas from Schweizer and Sklar (2005):

1.11. Definition. A *binary operation* on a nonempty set S is a function $T : S \times S \to S$. A *binary system* is a pair (S, T).

1.12. Definition. Let (S,T) be a binary system. For any a in S, the *vertical section* of T at a is the function $\eta_a : S \to S$ defined by

$$\eta_a(x) := T(a, x)\,; \tag{1.19}$$

and the *horizontal section* of T at a is the function $h_a : S \to S$ defined by

$$h_a(x) := T(x, a)\,. \tag{1.20}$$

The *diagonal section* of the binary operation T is the function $\delta_T : S \to S$ defined by

$$\delta_T(x) := T(x, x)\,. \tag{1.21}$$

Bivariate copulas are a particular type of binary operations on $[\,0,1\,]$. If C is a bivariate copula then from the previous section we know that for the binary system $([\,0,1\,], C)$ and for any a in $[\,0,1\,]$ the sections η_a, h_a, and δ_C are all nondecreasing and uniformly continuous functions on $[\,0,1\,]$.

1.13. Definition. Let T be a binary operation on S. An element a of S is a *left null element* of T if $T(a,x) = \eta_a(x) = a$ for all x in S; a *right null element* of T if $T(x,a) = h_a(x) = a$ for all x in S; and a *null element* of T if it is both a left and right null element of T. Correspondingly, an element a of S is a *left identity* of T if $T(a,x) = \eta_a(x) = x$ for all x in S; a *right identity* of T if $T(x,a) = h_a(x) = x$ for all x in S; and an *identity* of T if it is both a left and a right identity of T. An element a of S is *idempotent* under T if $T(a,a) = \delta_T(a) = a$, that is if a is a fixed point of δ_T.

Thus (left or right) null elements and (left or right) identities are idempotent elements.

1.14. Lemma. *If a is a left null element and b a right null element of T, then $a = b$. If a is a left identity and b a right identity of T, then $a = b$.*

The above lemma implies that a binary operation can have at most one null element and one identity. Idempotent elements that are neither null elements nor identities can be of course much more numerous. In the case of copulas, 0 is the (unique) null element and 1 the (unique) identity.

1.15. Definition. The *dual of a bivariate copula* C is the function $\tilde{C} : [\,0,1\,]^2 \to [\,0,1\,]$ defined by

$$\tilde{C}(u,v) := u + v - C(u,v)\,. \tag{1.22}$$

1.16. Lemma. *The dual of a bivariate copula is a binary operation on* $[\,0,1\,]$, *with identity* 0 *and null element* 1, *which is continuous and nondecreasing in each variable, but not 2-increasing (see (1.7)). If* $\tilde{C}_1$ *and* $\tilde{C}_2$ *are the dual of copulas* C_1 *and* C_2*, respectively, then* $\tilde{C}_1 \le \tilde{C}_2$ *if and only if* $C_2 \le C_1$*. Hence every dual of a copula satisfies*

$$\tilde{M}(u,v) \le \tilde{C}(u,v) \le \tilde{W}(u,v)\,. \tag{1.23}$$

The above lemma implies that the dual of a bivariate copula is not a copula. As pointed out by Nelsen (2006a) if C is the copula of a pair of continuous random variables X and Y, with marginal probability distributions F and G, respectively, and joint probability distribution H, the dual of C expresses a probability of an event involving X and Y :

$$\mathbb{P}[X \le x \text{ or } Y \le y] = F(x) + G(y) - H(x,y) = \tilde{C}\big(F(x), G(y)\big)\,. \tag{1.24}$$

The random variables $\max(X,Y)$ and $\min(X,Y)$ are the order statistics for X and Y, and it is immediate to verify that their distribution functions are given by, respectively,

$$\mathbb{P}[\max(X,Y) \le t] = C\big(F(t), G(t)\big) \quad \text{and} \quad \mathbb{P}[\min(X,Y) \le t] = \tilde{C}\big(F(t), G(t)\big)\,, \tag{1.25}$$

and in the particular case when $F = G$ we get

$$\mathbb{P}[\max(X,Y) \le t] = \delta_C\big(F(t)\big)\,, \tag{1.26}$$

$$\mathbb{P}[\min(X,Y) \le t] = 2F(t) - \delta_C\big(F(t)\big)\,. \tag{1.27}$$

If we denote the diagonal section of the dual of a copula C by

$$\tilde{\delta}_C(t) := \tilde{C}(t,t) = 2t - \delta_C(t)\,, \tag{1.28}$$

then (1.27) may be rewritten as

$$\mathbb{P}[\min(X,Y) \le t] = \tilde{\delta}_C\big(F(t)\big)\,. \tag{1.29}$$

In a few words, $\delta_C\big(F(t)\big)$ and $\tilde{\delta}_C\big(F(t)\big)$ are the distribution functions of the order statistics $\max(X,Y)$ and $\min(X,Y)$, respectively, when X and Y are continuous random variables with a common marginal distribution F and copula C.

It immediately follows from the basic definitions and results given in the previous section that if δ is the diagonal section of a copula then

$$\delta(0) = 0 \quad \text{and} \quad \delta(1) = 1 \qquad \text{(that is 0 and 1 are fixed points of } \delta) ; \tag{1.30}$$

$$0 \le \delta(t_2) - \delta(t_1) \le 2(t_2 - t_1), \quad \text{for all } t_1, t_2 \text{ in } [0,1] \text{ with } t_1 \le t_2 ; \tag{1.31}$$

$$\max(2t-1, 0) \le \delta(t) \le t, \qquad \text{for all } t \text{ in } [0,1]. \tag{1.32}$$

As proposed by Fredricks and Nelsen (1997b), any function from $[0,1]$ to $[0,1]$ satisfying the above three properties will be called simply a *diagonal,* while the function $\delta_C(t) = C(t,t)$ will be referred to as the *diagonal section of* C.

1.17. Lemma. ***(Fredricks and Nelsen, 1997b)*** *Let δ be any diagonal and $\tilde{\delta}$ its dual. Then*

i) $\tilde{\delta}$ is a nondecreasing, absolutely continuous function mapping $[0,1]$ onto $[0,1]$, and

$$t \le \tilde{\delta}(t) \le \min(2t, 1), \quad \textit{for all } t \textit{ in } [0,1] ;$$

ii) if δ is differentiable at t_0 in the interior of $[0,1]$, then

$$0 \le \delta'(t_0) \le 2 \qquad \textit{and} \qquad 0 \le \tilde{\delta}'(t_0) \le 2 .$$

If δ is any diagonal, does there exist a copula C whose diagonal section is δ? This question has already been answered affirmatively by Fredricks and Nelsen (1997a, 1997b, 2002). They made the following remark: For any copula C and any (u,v) in $[0,1]^2$, the 2-increasing property (1.7) implies that $C(v,v) - C(u,v) - C(v,u) + C(u,u) \ge 0$; so that if C is *symmetric* (that is if $C(u,v) = C(v,u)$ for all (u,v) in $[0,1]^2$), then $2C(u,v) \le \delta_C(u) + \delta_C(v)$, or $C(u,v) \le (1/2)[\delta_C(u) + \delta_C(v)]$. Thus, if C is any symmetric copula, then $C(u,v) \le \min(u, v, (1/2)[\delta_C(u) + \delta_C(v)])$. The following theorem proved that this last upper bound (with δ_C replaced by any diagonal δ) is itself, like the Fréchet-Hoeffding upper bound, a copula:

1.18. Theorem. ***(Fredricks and Nelsen, 1997b)*** *Let δ be any diagonal, and set*

$$K_\delta(u,v) := \min\left(u\,,\; v\,,\; \frac{\delta(u) + \delta(v)}{2} \right). \tag{1.33}$$

Then K_δ is a copula whose diagonal section is δ (that is $\delta_K = \delta$). Moreover, if C is any symmetric copula with diagonal section δ, then $C \le K_\delta$ on $[0,1]^2$.

Copulas of the form (1.33) were called *diagonal copulas* by Fredricks and Nelsen (1997a, 1997b). Although copula K_δ is completely determined by its diagonal δ, it is important to mention than not every copula characterized by its diagonal is necessarily of the form (1.33), as proved, for example, by Fredricks and Nelsen (2002):

1.19. Proposition. ***(Fredricks and Nelsen, 2002)*** *Let δ be any diagonal and define B_δ on $[\,0,1\,]^2$ by*

$$B_\delta(u,v) := \begin{cases} u - \inf\limits_{u\,\leq\, t\,\leq v}[\,t - \delta(t)\,]\,, & u \leq v\,, \\ v - \inf\limits_{v\,\leq\, t\,\leq u}[\,t - \delta(t)\,]\,, & v \leq u\,. \end{cases}$$

Then B_δ is a symmetric copula with diagonal section δ. Moreover, if C is any copula with diagonal section δ, then $B_\delta \leq C$ on $[\,0,1\,]^2$.

In all cases, the results obtained by Fredricks and Nelsen (1997a, 1997b, 2002) lead to symmetric singular copulas. **If δ is any diagonal, does there exist an absolutely continuous copula C whose diagonal section is δ?** We get a partial answer in the case of *Archimedean copulas* (see next section). Archimedean copulas are always symmetric, but it is also possible to build absolutely continuous copulas, not necessarily symmetric, with a given diagonal (see chapter 3), as proved by Erdely and González-Barrios (2006a), Nelsen, Quesada-Molina *et al* (2006), and Nelsen (2006b).

1.3 Archimedean copulas.

We will review the construction of this particular type of copulas through some results from the point of view of *functional equations,* particularly in terms of certain solution of the *associativity equation.*

According to Aczél (1966) or Castillo and Ruiz (1993), a functional equation may be considered as an equation which involves independent variables, known functions, unknown functions, and constants. The fact that it is excluded the possibility of infinitely many variables or functions as well as the possibility of known and unknown operators and functionals excludes the consideration of differential and integral equations, and other equations which involve infinitesimal operators. The main interest in this field of study is the substitution of known or unknown functions by other known or unknown functions.

Among many kinds of problems regarding functional equations, there is the problem of representing multivariate functions by superpositions of functions of a smaller number of variables. Questions like, for example, given a bivariate function T, when is it possible to find univariate functions f and g such that

$$T(x,y) = g\big(f(x) + f(y)\big)? \tag{1.34}$$

By assuming certain properties on T it had been possible to have different *representation theorems* of the form (1.34). We are specifically interested in a theorem by Ling (1965), but before we proceed, we need to review some basic concepts involved with such result.

Let S be a nonempty set. $T : S \times S \to S$ is an *associative function* if

$$T\big(T(x,y),z\big) = T\big(x,T(y,z)\big)\,, \qquad \text{for all } x,y,z \text{ in } S\,. \tag{1.35}$$

According to Definition 1.11 we may consider T as a binary operation on S. When T is associative, the binary system (S,T) is what is known as a *semigroup.* Some authors prefer to use symbols such as $*$ to denote binary operators like T, so that when referring to a semigroup $(S,*)$ it is meant that

$$(x * y) * z = x * (y * z)\,, \qquad \text{for all } x,y,z \text{ in } S\,. \tag{1.36}$$

The T*-powers* (or simply the *powers under* $*$) of $x \in S$ are the elements of S given recursively by

$$x^1 := x \qquad \text{and} \qquad x^{n+1} := x^n * x \equiv T(x^n, x) \tag{1.37}$$

for all positive integers n. Since both addition and multiplication are commutative operations on the integers, by induction we have

$$x^{m+n} = x^m * x^n = x^n * x^m \equiv T(x^m, x^n) = T(x^n, x^m) \tag{1.38}$$

and

$$x^{mn} = (x^m)^n = (x^n)^m \tag{1.39}$$

for all positive integers m,n and x in S. If T is associative and has an identity element b then (1.37) can be extended to the nonnegative integers by defining $x^0 := b$ for all x except left and right null elements of T. So for the diagonal section we have $\delta_T(x) = x^2$ and simple induction yields

$$\delta_T^n(x) = x^{2^n} \tag{1.40}$$

for all integers $n \geq 0$ and all x in S.

If $f : [a,b] \to [0,\infty]$ is a continuous and strictly decreasing function, the *pseudo-inverse* of f is defined by

$$g(x) := \begin{cases} b\,, & \text{if } x \text{ is in } [0, f(b)]\,, \\ f^{-1}(x)\,, & \text{if } x \text{ is in } [f(b), f(a)]\,, \\ a\,, & \text{if } x \text{ is in } [f(a), \infty]\,. \end{cases} \tag{1.41}$$

It is immediate to verify that the pseudo-inverse g is continuous and nonincreasing.

1.20. Theorem. ***(Ling, 1965)*** *Let S be a closed interval $[a, b]$ and $T : S \times S \to S$ be an associative function satisfying the following conditions:*

(1) T is continuous,

(2) T is nondecreasing in each variable,

(3) The endpoint b is a left identity, that is $T(b, x) = x$ for all x in S,

(4) For all x in the interior of S, the diagonal section $\delta_T(x) = T(x, x) < x$.

Then there exists a continuous and strictly decreasing function $f : S \to [0, \infty]$ such that T is representable in the form

$$T(x, y) = g(f(x) + f(y)), \tag{1.42}$$

where g is the pseudo-inverse of f.

The function f in (1.42) is called *additive generator* of T. Another important remark is that if T satisfies the hypotheses of Ling's theorem, then by (1.42) we have that T is *symmetric*:

$$T(x, y) = T(y, x), \qquad \text{for all } x, y \text{ in } S. \tag{1.43}$$

Equivalently, $(S, *)$ is said to be a *commutative semigroup*, that is

$$x * y = y * x, \qquad \text{for all } x, y \text{ in } S. \tag{1.44}$$

1.21. Definition. Let T be an associative binary operation on the interval $[a, b]$ satisfying

1. T is nondecreasing in each variable, that is

$$T(x_1, y_1) \le T(x_2, y_2) \tag{1.45}$$

   for all x_1, x_2, y_1, y_2 in $[a, b]$ such that $x_1 \le x_2$ and $y_1 \le y_2$.

2. The endpoint b is an identity of T, that is

$$T(x, b) = T(b, x) = x \qquad \text{for all } x \text{ in } [a, b]. \tag{1.46}$$

Then T is said to be *Archimedean* if for any x, y in the interior of $[a, b]$ there is a positive integer m such that $x^m < y$. The binary system $([a, b], T)$ is then called an *Archimedean semigroup*.

A typical example of an Archimedean semigroup would be $([0,1], \cdot)$ with $\cdot$ the usual multiplication operator on the real line. The following lemma is from Schweizer and Sklar (2005):

1.22. Lemma. *Let $([a,b], T)$ be a semigroup, with T a continuous function. Then T is Archimedean if and only if $\delta_T(x) < x$ for all x in the interior of $[a,b]$.*

An immediate consequence of this lemma is that if $([a,b], T)$ is an Archimedean semigroup with T a continuous function then T admits Ling's representation (1.42). Moreover, the converse of Ling's theorem is also true:

1.23. Lemma. ***(Ling, 1965)*** *Let $[a,b]$ be a closed interval, $f : [a,b] \to [0,\infty]$ be a continuous and strictly decreasing function, and $g : [0,\infty] \to [a,b]$ the pseudo-inverse of f. Then the bivariate function T defined by*

$$T(x,y) := g\big(f(x) + f(y)\big)$$

is continuous and $([a,b], T)$ is a (commutative) Archimedean semigroup.

It is important to mention that the representation (1.42) is still possible under weaker assumptions than those of Ling's theorem, mainly without asking T to be continuous but with other conditions instead, see Krause (1981) or Schweizer and Sklar (2005). The following two results were taken from Schweizer and Sklar (2005):

1.24. Corollary. *If $([a,b], T)$ is an Archimedean semigroup with T a continuous function, then there is a c in $[a,b]$ such that the diagonal section $\delta_T(x) = a$ for x in $[a,c]$, while δ_T is strictly increasing on $[c,b]$.*

1.25. Corollary. *If $([a,b], T)$ is an Archimedean semigroup with T a continuous function, and T is strictly increasing in each variable on $]a,b]^2$, then $f(a) = \infty$, whence $g = f^{-1}$ and*

$$T(x,y) = f^{-1}\big(f(x) + f(y)\big) \tag{1.47}$$

for all x, y in $[a,b]$.

Within the context of Archimedean semigroups and the results reviewed so far in this section, we may now define a particular type of copulas:

1.26. Corollary. *Let C be a bivariate copula such that C is associative and $\delta_C(t) < t$ for all t in the open interval $]\,0,1\,[$. Then C admits Ling's representation (1.42), that is, there exists a continuous and strictly decreasing function $\varphi : [\,0,1\,] \to [\,0,\infty\,]$ such that C is representable in the form*

$$C(u,v) = \varphi^{[-1]}\big(\varphi(u) + \varphi(v)\big)\,, \tag{1.48}$$

where $\varphi^{[-1]}$ is the pseudo-inverse of φ.

The above corollary is an immediate consequence of Definition 1.1, the continuity of any copula, Lemma 1.22 and Ling's theorem. We should also notice that $([\,0,1\,], C)$ is therefore an Archimedean semigroup, C is symmetric (commutative), and as a consequence of Lemma 1.24 there is a t_0 in $[\,0,1\,]$ such that $\delta_C(t) = 0$ for t in $[\,0,t_0\,]$ and δ_C is strictly increasing on $[\,t_0,1\,]$.

1.27. Definition. An associative copula C such that $\delta_C(t) < t$ for all t in the open interval $]\,0,1\,[$ is called *Archimedean copula.*

In the particular case of an Archimedean copula, the definition of pseudo-inverse of its (additive) generator φ as in (1.41) becomes

$$\varphi^{[-1]}(t) := \begin{cases} \varphi^{-1}(t)\,, & 0 \leq t \leq \varphi(0)\,, \\ 0\,, & \varphi(0) \leq t \leq \infty\,. \end{cases} \tag{1.49}$$

We also note that in case $\varphi(0) = \infty$, then $\varphi^{[-1]} = \varphi^{-1}$, the usual inverse.

Even though we may use Lemma 1.23 and a continuous and strictly decreasing function $\varphi : [\,0,1\,] \to [\,0,\infty\,]$ to build a function $C : [\,0,1\,]^2 \to [\,0,1\,]$ such that $([\,0,1\,], C)$ is a commutative Archimedean semigroup, this is not sufficient to guarantee that C is indeed a copula, further properties are required on the generator φ, as proved in Alsina, Frank and Schweizer (2006) or Nelsen (2006a):

1.28. Theorem. *Let $\varphi : [\,0,1\,] \to [\,0,\infty\,]$ be a convex, continuous, strictly decreasing function such that $\varphi(1) = 0$. Then the function $C : [\,0,1\,]^2 \to [\,0,1\,]$ given by*

$$C(u,v) := \varphi^{[-1]}\big(\varphi(u) + \varphi(v)\big) \tag{1.50}$$

is an Archimedean copula.

If a function φ satisfies the hypothesis of the above theorem, it is called *(additive) generator of the Archimedean copula.* In case $\varphi(0) = \infty$ (and so $\varphi^{[-1]} = \varphi^{-1}$, the usual inverse), it is said that φ is a *strict generator*, and in this case we have $C(u,v) = \varphi^{-1}\big(\varphi(u) + \varphi(v)\big)$. In case $\varphi(0) < \infty$ it is called *non-strict generator.* It is also straight forward to verify that if φ is a generator of a copula C, then for any constant $k > 0$ we have that $k\varphi$ is also a generator of C.

For example, for any constant $k > 0$, if $\varphi_1(t) := -k \log t$ and $\varphi_2(t) := k(1-t)$ for t in $[\,0,1\,]$, it is straightforward to verify that φ_1 and φ_2 satisfy the hypothesis of Theorem 1.28 and so they generate Archimedean copulas. In fact, φ_1 is a strict generator and generates the product copula $\Pi(u,v) = uv$, while φ_2 is a non-strict generator and generates the Fréchet-Hoeffding lower bound copula $W(u,v) = \max(u+v-1,0)$.

In the case of the Fréchet-Hoeffding upper bound copula $M(u,v) = \min(u,v)$, it cannot be Archimedean since $\delta_M(t) = t$, although it is an associative copula. For methods to build associative non-Archimedean copulas see Mikusiński and Taylor (1999) or Schweizer and Sklar (2005).

Theorem 1.28 is also a powerful tool in building families of parametric Archimedean copulas by using parametrized generators. For example,

$$\varphi_\theta(t) := \frac{1}{\theta}\left(t^{-\theta} - 1\right), \qquad \theta \in [-1, \infty\,) \setminus \{0\}, \tag{1.51}$$

generates a parametric family of Archimedean copulas, discussed by Clayton (1978):

$$C_\theta(u,v) = \Big[\ \max\left(u^{-\theta} + v^{-\theta} - 1\,,\,0\right)\Big]^{-1/\theta}. \tag{1.52}$$

A list of different Archimedean families may be found in Nelsen (2006a), Alsina, Frank and Schweizer (2006) or De Matteis (2001).

The *level sets* of a copula C are given by $\{(u,v) \in [\,0,1\,]^2 \,:\, C(u,v) = t\}$. In the particular case of Archimedean copulas and for $t > 0$, the level sets are in fact *level curves* since $\varphi^{[-1]}\big(\varphi(u) + \varphi(v)\big) = t$ implies

$$v \,=\, \varphi^{[-1]}\big(\varphi(t) - \varphi(u)\big) = \varphi^{-1}\big(\varphi(t) - \varphi(u)\big), \tag{1.53}$$

where the replacement of $\varphi^{[-1]}$ by φ^{-1} is justified by the fact that $\varphi(t) - \varphi(u)$ lies in the interval $[\,0, \varphi(0)\,[$. For $t = 0$, $\{(u,v) \in [\,0,1\,]^2 \,:\, C(u,v) = 0\}$ is called the *zero set* of C. For many Archimedean copulas, their zero set is simply the union of the two line segments

$\{0\} \times [\,0,1\,]$ and $[\,0,1\,] \times \{0\}$ (for example, $\Pi(u,v) = uv$); for others, their zero set has positive area (for example, $W(u,v) = \max(u+v-1,0)$) and for such zero set the boundary curve $\varphi(u) + \varphi(v) = \varphi(0)$ is called the *zero curve* of C. The following theorem is proved in Nelsen (2006a):

1.29. Theorem. *The level curves of a bivariate Archimedean copula are convex.*

One important remark about the diagonal section of Archimedean copulas: most of the well-known families of Archimedean copulas have convex diagonals, but this is not true in general, as an example we have copula 4.2.18 in Nelsen's catalog (2006a) with, for example, parameter $\theta = 2$. Another example, provided by Mesiar (2006), using the non-strict generator

$$\varphi(u) := \begin{cases} 1 - \frac{3}{2}\,u\,, & 0 \leq u \leq \frac{1}{2}\,, \\ \frac{1-u}{2}\,, & \frac{1}{2} \leq u \leq 1\,, \end{cases}$$

which yields the non-convex diagonal section

$$\delta(u) := \begin{cases} 0\,, & 0 \leq u \leq \frac{1}{3}\,, \\ \frac{2}{3}\,(3u-1)\,, & \frac{1}{3} \leq u \leq \frac{1}{2}\,, \\ \frac{2}{3}\,u\,, & \frac{1}{2} \leq u \leq \frac{3}{4}\,, \\ 2u - 1\,, & \frac{3}{4} \leq u \leq 1\,. \end{cases}$$

It is important to mention that, recently, Durante, Quesada-Molina, and Sempi (2006) have introduced and studied a class of bivariate copulas depending on two univariate functions, which generalizes the Archimedean family.

For the multivariate case, we are led naturally to the problem of analyzing whether the function $C : [\,0,1\,]^n \to [\,0,1\,]$ given by

$$C(u_1,\ldots,u_n) = \varphi^{[-1]}\big(\varphi(u_1) + \varphi(u_2) + \cdots + \varphi(u_n)\big)\,, \quad n \geq 3\,, \tag{1.54}$$

is a copula, and in such case we would call it *multivariate Archimedean copula.* For this purpose we need the following definition from Widder (1946):

1.30. Definition. A function g is *completely monotonic* on an interval J if it is continuous and it has derivatives of all orders that alternate in sign, that is if g satisfies

$$(-1)^k \frac{d^k}{d\,t^k}\, g(t) \geq 0 \tag{1.55}$$

for all t in the interior of J and $k = 0, 1, 2, \ldots$

By a result from Widder (1946), if $g(t)$ is completely monotonic on $[\,0,\infty\,[$ and $g(c)=0$ for some finite $c>0$, then g must be identically zero on $[\,0,\infty\,[$. As a consequence, if the pseudo-inverse $\varphi^{[-1]}$ of an Archimedean generator φ is completely monotonic, it has to be positive on $[\,0,\infty\,[$ and so φ is a strict generator, that is $\varphi^{[-1]}=\varphi^{-1}$, the usual inverse.

Kimberling (1974) obtained necessary and sufficient conditions for a strict generator φ to generate multivariate Archimedean copulas. See also Nelsen (2006a) or Alsina, Frank and Schweizer (2006):

1.31. Theorem. *Let φ be a strict Archimedean generator. The function C given by (1.54) is a multivariate copula for all $n\geq 2$ if and only if φ^{-1} is completely monotonic on $[\,0,\infty\,[$.*

An immediate example is the case of the (strict) generator of the bivariate product copula given by $\varphi_1(t)=-\log t$, so $\varphi_1^{-1}(t)=e^{-t}$; it is clear that φ^{-1} is completely monotonic and, as expected, it generates the multivariate product copula $\Pi(u_1,\ldots,u_n)=u_1u_2\cdots u_n$.

Another example is the case of the Clayton family of copulas. In the bivariate case its generator is given by (1.51) and it is strict only if $\theta>0$, in which case φ_θ^{-1} is completely monotonic, generating the multivariate Clayton family of copulas given by

$$C_\theta(u_1,\ldots,u_n)=\Big(u_1^{-\theta}+\cdots+u_n^{-\theta}-n+1\Big)^{-1/\theta}. \tag{1.56}$$

1.4 Schröder's functional equation.

If C is an Archimedean copula with (additive) generator φ, according to (1.50) the diagonal section of C, $\delta(u)=C(u,u)$, may be directly expressed in terms of such generator:

$$\delta(u)=\varphi^{[-1]}\big(2\varphi(u)\big)\,,\quad u\in[\,0,1\,]. \tag{1.57}$$

The details of the proof of Ling's representation theorem (Theorem 1.20) show the important role of assumption (4), that is $\delta(u)<u$ for u in the interior of the closed interval under consideration. So one may ask, as observed by Darsow and Frank (1983), how much information about an Archimedean C is contained in its diagonal section. In other words, given δ, what can be said about φ?

First let's assume that φ is a strict generator. Then by (1.49) we have that $\varphi^{[-1]}=\varphi^{-1}$, the usual inverse, and (1.57) may be written as

$$\varphi\big(\delta(u)\big)=2\,\delta(u)\,,\qquad u\in[\,0,1\,]. \tag{1.58}$$

The above equation is a particular case of *Schröder's functional equation*

$$\varphi \circ \delta = \lambda \delta, \qquad \lambda \neq 0, 1, \text{ a constant,} \tag{1.59}$$

which has been studied in one form or another, according to Kuczma (1968), since the late nineteenth century:

> [(1.59)] appeared for the first time about 1870 [Schröder (1871)] in connection with the problem of continuous iteration (cf. Chapter IX). A fundamental theorem about the existence and uniqueness of analytic solutions of [(1.59)] was then proved by G. Koenigs [(1884)], and therefore equation [(1.59)] is often called also the *Koenigs equation* or the *Schröder-Koenigs equation* [...] The authors have mainly paid attention to analytic solutions on the complex plane. The Schröder equation for functions of a real variable has been dealt in the [twentieth] century.

In the abstract of Sungur and Yang (1996) they made the following statement:

> It is shown that for the Archimedean class, the diagonal copula uniquely determines the corresponding copula.

They defined in their article as "diagonal copula" what has been defined in the present work in (1.21) as *diagonal section.* The above statement by Sungur and Yang (1996) is wrong. Frank (1996) announced:

> We explore the following question [...] When is an associative copula C uniquely determined by its diagonal $\delta(x) = C(x,x)$? [...] a sufficient, but not necessary, uniqueness condition for Archimedean C is given by [the left derivative] $\delta\,'(1) = 2$. This is an almost immediate consequence of standard results on convex solutions of Schröder's equation (with a new, direct proof) via the representation of these copulas. We present some related conditions and illustrate non-uniqueness by constructing families of copulas having identical diagonals.

The above quotation from Frank (1996) is a report of meeting in the Thirty-third International Symposium on Functional Equations, May 21 - May 27, 1995, held in Caldes de Malavella, Catalonia, Spain. According to Schweizer (2006) the relevant paper has not been published. Such result is now included in the recently published book by Alsina, Frank and Schweizer (2006), pp. 151-155, including an example with the construction of different Archimedean copulas that have the same diagonal, which contradicts the statement by Sungur and Yang (1996). The problem of finding conditions under which an Archimedean copula is or is not uniquely determined by its diagonal section has been also studied in the context of *triangular norms*, see for example Klement, Mesiar and Pap (2000) and Klement and Mesiar (2005).

From a remark from Itô (1996) we have that if φ_1 is a solution of $\varphi\big(\delta(u)\big) = \lambda\,\delta(u)$ and if g_1 is a solution of $g\big(\delta(u)\big) = g(u)$ then it is straightforward to verify that the product $\varphi_1 g_1$ is also a solution of (1.59). A trivial solution for $g\big(\delta(u)\big) = g(u)$ is $g_1 = k$, being k any constant, and in such case we have that if φ_1 is a solution of (1.59) then $k\,\varphi_1$ is also a solution. The following result is a particular case of Theorem 6.6 in Kuczma (1968) (or Theorem 2.3.12 in Kuczma *et al* (1990)):

1.32. Theorem. *Let the function $\gamma : [\,0,1\,] \to [\,0,1\,]$ be such that $0 < \gamma(u) < u$ for all $u \in \,]\,0,1\,[\,$, and $\gamma\,'(0) = \frac{1}{2}$. If $s(u)$ is a solution of the functional equation*

$$s\big(\gamma(u)\big) \;=\; \frac{1}{2}\,s(u) \tag{1.60}$$

such that the function $s(u)/u$ is monotonic in $]\,0,1\,[$, then

$$s(u) \;=\; k \lim_{n\to\infty} 2^{\,n}\gamma^{\,n}(u)\,, \tag{1.61}$$

where $\gamma^{\,n}$ is the n-th iteration of γ, that is the composition of γ with itself n times, and k any constant.

The following result is part of what was announced by Frank (1996), and now included in Alsina, Frank and Schweizer (2006) only defining the functions needed for the proof. For completeness we present the proof of the following:

1.33. Theorem. ***(Frank, 1996)*** *If C is an Archimedean copula whose diagonal δ satisfies $\delta\,'(1-) = 2$ then it is uniquely determined by its diagonal.*

Proof: We first consider the case of a strict generator φ, then from (1.41), (1.49), and (1.57), we have that δ is continuous and strictly increasing in $[\,0,1\,]$, and so it has an inverse $\delta^{\,-1}$. By defining a function γ on $[\,0,1\,]$ via $\gamma(u) := 1 - \delta^{\,-1}(1-u)$ for u in $[\,0,1\,]$ we have that γ is continuous and strictly increasing, $\gamma(0) = 0$, and $0 < \gamma(u) < u$ for all $u \in]0,1[$. Now define $s(u) := \varphi(1-u)$ and substituting the definitions of γ and s in (1.58) we get that this functional equation is equivalent to (1.60) and so we may apply Theorem 1.32 by requiring $s(u)/u$ to be monotonic and $\lim_{u\to 0}[\,\gamma(u)/u\,] = \dfrac{1}{2}$. This last condition is fulfilled if γ is right-differentiable in zero and $\gamma\,'(0+) = \frac{1}{2}$ which is equivalent to $\delta\,'(1-) = 2$.

Since we are dealing with an Archimedean copula the generator φ is convex and so it is $s(u) = \varphi(1-u)$, then by Proposition 6.3.2 in Dudley (2002) we have that $s(u)/u$ is monotonic. Applying Theorem 1.32 we obtain the following formula for φ in terms of diagonal δ :

$$\varphi(u) = k \lim_{n\to\infty} 2^n \big[1 - \delta^{-n}(u)\big]\,, \tag{1.62}$$

where δ^{-n} is the composition of δ^{-1} with itself n times, and k is any positive constant, since we require that $\varphi \geq 0$. So far we have considered the case where φ is a strict generator. In the case of a non-strict generator, that is $\varphi(0) < \infty$, if we define $\alpha := \varphi^{-1}(\varphi(0)/2)$, then from (1.49) and (1.57) we get

$$\delta(u) = \begin{cases} \varphi^{-1}(2\varphi(u))\,, & \text{if} \quad \alpha \leq u \leq 1\,, \\ 0\,, & \text{if} \quad 0 \leq u \leq \alpha\,, \end{cases} \tag{1.63}$$

that is, $\delta = 0$ in $[\,0, \alpha\,]$ and δ is strictly increasing in $[\,\alpha, 1\,]$. We notice that $\alpha \leq \frac{1}{2}$ by (1.32). Then

$$\varphi(\delta(u)) = \begin{cases} 2\varphi(u)\,, & \text{if} \quad \alpha \leq u \leq 1\,, \\ \varphi(0)\,, & \text{if} \quad 0 \leq u \leq \alpha\,. \end{cases} \tag{1.64}$$

We have that for $0 \leq u \leq \alpha$, $\varphi(\delta(u)) = \varphi(0)$ since $\delta(u) = 0$ in such interval, so it just remains to solve $\varphi(\delta(u)) = 2\varphi(u)$ for $\alpha \leq u \leq 1$, and we proceed as in the case of a strict generator, by taking $w = (u-\alpha)/(1-\alpha)$ which takes values in $[0,1]$, and solving $\varphi(\delta(w)) = 2\varphi(w)$. □

From now on we will refer to the condition $\delta\,'(1-) = 2$ as *Frank's condition.* An important example of an Archimedean copula that satisfies Frank's condition is the case of the product copula $\Pi(u,v) = uv$, which characterizes a couple of independent continuous random variables, via Sklar's Theorem, and so it is uniquely determined by its diagonal section $\delta_\Pi(u) = u^2$. **This fact is crucial for the work done in this thesis.** Frank's condition is satisfied by 13 out of 22 copulas in the catalog of Archimedean copulas provided by Nelsen (2006a). From Alsina, Frank and Schweizer (2006) we have the following remark:

> [Frank's condition] is not nearly as stringent as it might seem at first sight. Thus, since $\delta_W{}'(1-) = \delta_\Pi(1-) = 2$, [where W is Frechet-Hoeffding's lower bound copula (1.9)] we have $\delta_C{}'(1-) = 2$ whenever $W \leq C \leq \Pi$.

As a consequence of the above remark, for all the parametric families of Archimedean copulas $\{C_\theta : \theta \in \Theta\}$ for which there exists a subset $\Theta_0 \subseteq \Theta$ such that $C_\theta \leq \Pi$ for all θ in Θ_0, the subfamily $\{C_\theta : \theta \in \Theta_0\}$ satisfies Frank's condition. This will be the case of *negative quadrant dependence* (see Definition 1.35 and (1.80) in section 1.5).

In Alsina *et al* (2006), Section 3.8, a counterexample is given, in order to show that if φ is a generator for an Archimedean copula C such that $\varphi'(1-) = 0$, or equivalently $\delta_C{}'(1-) < 2$, where δ_C is the diagonal of the copula C, then the diagonal does not characterize uniquely the generator φ. Alsina *et al* (2006) provide a parametric family of generators $\{\varphi_\beta : 0 < \beta \leq 1/(1+8\pi)\}$ such that their diagonal section $\delta_\beta = \delta_C$ but $C_{\beta_1} \neq C_{\beta_2}$ for $\beta_1 \neq \beta_2$. The details will be done to show that the given upper bound for β is not sharp, by providing a sharp one.

Let $0 < \beta \leq 1$ and define for $0 < x < 1$

$$\varphi_\beta(x) = \ln(x)^2 + 2^n \beta \sin\left(\frac{\ln(x)^2}{2^n}\right) \quad \text{if} \quad 2^{n+1}\pi \leq \ln(x)^2 < 2^{n+2}\pi, \tag{1.65}$$

for $n \in \mathbb{Z}$. First we observe that (1.65) is equivalent to

$$\varphi_\beta(x) = \ln(x)^2 + 2^n \beta \sin\left(\frac{\ln(x)^2}{2^n}\right) \quad \text{if} \quad \exp(-\sqrt{2^{n+2}\pi}) < x \leq \exp(-\sqrt{2^{n+1}\pi}), \tag{1.66}$$

for $n \in \mathbb{Z}$. Since

$$\lim_{n\to\infty} \exp(-\sqrt{2^{n+2}\pi}) = 0 \quad \text{and} \quad \lim_{n\to-\infty} \exp(-\sqrt{2^{n+2}\pi}) = 1.$$

Then $\varphi_\beta(x)$ is defined and it is continuous on $(0,1)$, to see that this holds, observe that

$$2^{n+1}\pi \leq \ln(x)^2 < 2^{n+1}\pi \quad \text{if and only if} \quad 2\pi \leq \frac{\ln(x)^2}{2^n} < 4\pi. \tag{1.67}$$

Therefore, $\sin\left(\frac{\ln(x)^2}{2^n}\right)$ has a whole period for each $n \in \mathbb{Z}$, in fact, for each $n \in \mathbb{Z}$

$$\begin{aligned} \varphi_\beta(\exp(-\sqrt{2^{n+1}\pi})-) &= 2^{n+1}\pi + 2^n \beta \sin(2\pi) \\ &= 2^{n+1}\pi \\ &= 2^{(n-1)+2}\pi + 2^{n-1}\beta \sin(4\pi) \\ &= \varphi_\beta(\exp(-\sqrt{2^{(n-1)+2}\pi})+). \end{aligned}$$

Now we observe that φ_β is differentiable and its derivative is given by

$$\varphi_\beta'(x) = \frac{2\ln(x)}{x}\left(1 + \beta \cos\left(\frac{\ln(x)^2}{2^n}\right)\right) \quad \text{if} \quad 2^{n+1}\pi \leq \ln(x)^2 < 2^{n+2}\pi, \tag{1.68}$$

for $0 < x < 1$, equivalently,

$$\varphi_\beta'(x) = \frac{2\ln(x)}{x}\left(1 + \beta \cos\left(\frac{\ln(x)^2}{2^n}\right)\right) \quad \text{if} \quad 2\pi \leq \frac{\ln(x)^2}{2^n} < 4\pi, \tag{1.69}$$

for $0 < x < 1$. We also have continuity of $\varphi'_\beta(x)$, since for each $n \in \mathbb{Z}$,

$$\begin{aligned}
\varphi'_\beta(\exp(-\sqrt{2^{n+1}\pi})-) &= \frac{-2\sqrt{2^{n+1}\pi}}{\exp(-\sqrt{2^{n+1}\pi})}(1+\beta\cos(2\pi)) \\
&= \frac{-2\sqrt{2^{n+1}\pi}}{\exp(-\sqrt{2^{n+1}\pi})}(1+\beta) \\
&= \frac{-2\sqrt{2^{(n-1)+2}\pi}}{\exp(-\sqrt{2^{(n-1)+2}\pi})}(1+\beta\cos(4\pi)) \\
&= \varphi'_\beta(\exp(-\sqrt{2^{n+1}\pi})+).
\end{aligned}$$

Now, we know that $\varphi_\beta(x)$ is convex if and only if $\varphi'_\beta(x)$ is increasing, see for example Pečarič *et al* (1992), Theorem 1.4. We want to find the values of β, if any, such that $\varphi_\beta(x)$ is convex on $(0,1)$.
Let $u^2 = \frac{\ln(x)^2}{2^n}$, for $0 < x < 1$, then $u = \frac{\ln(x)}{\sqrt{2^n}}$, hence $x = \exp(\sqrt{2^n}u)$. Also, for $n \in \mathbb{Z}$,

$$2^{n+1}\pi \le \ln(x)^2 < 2^{n+2}\pi \text{ if and only if } 2\pi \le u^2 < 4\pi \text{ if and only if } -\sqrt{4\pi} < u \le -\sqrt{2\pi}. \tag{1.70}$$

Therefore, (1.68) is equivalent to

$$\varphi'_\beta(u;n) = \frac{2\sqrt{2^n}u}{\exp(\sqrt{2^n}u)}\big(1+\beta\cos(u^2)\big) \text{ for } -\sqrt{4\pi} < u \le -\sqrt{2\pi} \text{ and every } n \in \mathbb{Z}. \tag{1.71}$$

Hence, to show that (1.68) is increasing for some $0 < \beta < 1$ is equivalent to show that (1.71) is increasing for some $0 < \beta < 1$. First we observe that for large positive values of n, $\varphi'_\beta(u;n)$ behaves like $\frac{2\sqrt{2^n}u}{\exp(\sqrt{2^n}u)}$ which decreases to $-\infty$ as $n \to \infty$, for every $-\sqrt{4\pi} < u \le -\sqrt{2\pi}$. On the other hand as n becomes smaller, or more negative, the denominator $\exp(\sqrt{2^n}u)$ in equation (1.71) approaches 1 and $\sqrt{2^n}u$ approaches zero. It can be shown that $\varphi'_\beta(u);n$ is increasing for "large" values of n, but for small values of n and some $\beta's$ it is not increasing anymore. In Table 1 we present the values of β, the first value of $n \in \mathbb{Z}$ for which $\varphi'_\beta(u;n)$, given in equation (1.71), is not increasing, and the corresponding interval in which φ_β ceases to be convex. Observe that from Table 1.1, the maximum value of n, for which φ_β is not convex, decreases with β and the corresponding interval in which φ_β ceases to be convex moves to the right with endpoints approaching 1.
Now we obtain the second derivative $\varphi''_\beta(u;n)$. We know that if $\varphi''_\beta(u;n) \ge 0$ for each $n \in \mathbb{Z}$, then $\varphi_\beta(u)$ is a convex function.

$$\begin{aligned}
\varphi''_\beta(u;n) &= \frac{2\sqrt{2^n}-2^{n+1}u}{\exp(\sqrt{2^n}u)}(1+\beta\cos(u^2)) - \frac{4\sqrt{2^n}\beta\sin(u^2)}{\exp(\sqrt{2^n}u)} \qquad (1.72)\\
&= \frac{2\sqrt{2^n}}{\exp(\sqrt{2^n}u)}\big(1+\beta\big(\cos(u^2)-2u^2\sin(u^2)\big)\big) - \frac{2^{n+1}u}{\exp(\sqrt{2^n}u)}(1+\beta\cos(u^2))
\end{aligned}$$

for $-\sqrt{4\pi} < u \leq -\sqrt{2\pi}$. We observe that the last expression on the right hand of (1.72) approaches 0 very quickly when $n \to -\infty$. We also observe that $\exp(\sqrt{2^n}u)$ approaches 1 uniformly for every $-\sqrt{4\pi} < u \leq -\sqrt{2\pi}$ as $n \to -\infty$. Therefore, $\varphi''_\beta(u; n)$ behaves like

$$2\sqrt{2^n}\left(1 + \beta\left(\cos(u^2) - 2u^2\sin(u^2)\right)\right) \quad \text{for} \quad 2\pi \leq u^2 < 4\pi, \tag{1.73}$$

as n tends to $-\infty$. Hence, letting $v = u^2$, the sign of $\varphi''_\beta(u : n)$ is given by the sign of

$$g(v) = K(1 + \beta(\cos(v) - 2v\sin(v))) \quad \text{for} \quad 2\pi \leq v < 4\pi, \tag{1.74}$$

where $K = 2\sqrt{2^n}$. Therefore, we only need to find the maximal value of β for which $g(v) \geq 0$ for any $2\pi \leq v < 4\pi$. This is equivalent to find the minimum of $k(v) = \cos(v) - 2v\sin(v)$ for $2\pi \leq v < 4\pi$. By analyzing the functions k and (1.74), we observe that if $\beta \leq 0.062548$ then $g(v) \geq 0$. Hence, the function $\varphi_\beta(x)$ is convex for $\beta \leq 0.062548$, which agrees with the results in Table 1.1.

In Alsina *et al* (2006), page 155, it is mentioned that $\varphi_\beta(x)$ is convex if $0 < \beta \leq \frac{1}{1+8\pi} = 0.038266$, which of course is true. However, this bound for β is not sharp as shown above, since convexity of φ_β also holds for $\beta \leq 0.062548$, in fact this a sharp bound for convexity of φ_β.

1.5 Dependence and copulas.

As a brief introduction to this section we quote part of the preface from Drouet and Kotz (2001):

> The concept of dependence permeates our Earth and its inhabitants in a most profound manner. Examples of interdependent meteorological phenomena in nature, interdependence in medical, social, and political aspects of our existence, not to mention economic structures, are too numerous to be cited individually. Moreover, the dependence is obviously not deterministic but of a stochastic nature.
> It is therefore somewhat surprising that the concepts and measures of dependence did not receive sufficient attention in the statistical literature, at least until as late as 1966 when the pioneering paper by E.L. Lehmann [see Lehmann (1966)] has appeared. The concept of correlation (and its modifications) introduced by F. Galton in 1885 dominated statistics during some 70 years of the 20-th century, practically serving as the only measure of dependence, often resulting in somewhat misleading conclusions. The last 20-th century have witnessed a rapid resurgence in investigations of dependence properties from statistical and probabilistic points

Table 1.1: Values of n and intervals for which $\varphi_\beta(x)$ ceases to be convex

β	maximal value of n for non convexity	corresponding interval
0.9	7	$[3.8 \times 10^{-18}, 4.8 \times 10^{-13})$
0.8	5	$[1.9 \times 10^{-9}, 6.9 \times 10^{-7})$
0.7	4	$[6.9 \times 10^{-7}, 4.4 \times 10^{-5})$
0.6	4	$[6.9 \times 10^{-7}, 4.4 \times 10^{-5})$
0.5	3	$[4.4 \times 10^{-5}, 0.00083)$
0.4	1	$[0.0066, 0.0288)$
0.3	1	$[0.0066, 0.0288)$
0.2	-1	$[0.08, 0.17)$
0.1	-5	$[0.53, 0.64)$
0.09	-6	$[0.64, 0.73)$
0.08	-7	$[0.73, 0.80)$
0.07	-10	$[0.89, 0.92)$
0.069	-10	$[0.89, 0.92)$
0.068	-11	$[0.92, 0.94)$
0.067	-11	$[0.92, 0.94)$
0.066	-12	$[0.94, 0.96)$
0.065	-13	$[0.96, 0.97)$
0.064	-14	$[0.97, 0.98)$
0.063	-18	$[0.993, 0.995)$
0.0626	-24	$[0.9991, 0.9994)$
0.06255	-35	$[0.99998, 0.99999)$
0.062549	-39	$[0.999995, 0.999996)$
0.062548	$-\infty$	does not exist

of view but the first –and to the best of our knowledge– the only text (of some 400 pages) devoted to dependence concepts (by Harry Joe) appeared as late as 1997 [see Joe (1997)]. Moreover, it seems to us that no Department of Statistics (or/and Mathematics) in the U.S.A. and Europe offer courses dealing specifically with dependence concepts and measures.

We recall that the *correlation* between two random variables X and Y is defined by

$$r(X, Y) := \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}},$$

provided the existence of second moments. Before we continue a critique against the use of correlation as a dependence measure, we review the following property for copulas, see Nelsen (2006a):

1.34. Theorem. *Let X and Y be continuous random variables with copula C_{XY}. If α and β are strictly increasing functions on Ran X and Ran Y, respectively, then $C_{\alpha(X)\beta(Y)} = C_{XY}$. Thus C_{XY} is invariant under strictly increasing transformations of X and Y.*

This result is also valid for the general multivariate case, see Schweizer and Sklar (2005). The above theorem in connection with the concept of correlation deserved the following comment by Embrechts *et al* (2003a):

> Copulas provide a natural way to study and measure dependence between random variables. As a direct consequence of [Theorem 1.34], copula properties are invariant under strictly increasing transformations of the underlying random variables. Linear correlation (or Pearson's correlation) is most frequently used in practice as a measure of dependence. However, since linear correlation is not a copula-based measure of dependence, it can often be quite misleading and should not be taken as the canonical dependence measure [...] The popularity of linear correlation stems from the ease with which it can be calculated and it is a natural scalar measure of dependence in elliptical distributions (with well-known members such as the multivariate normal and the multivariate t-distribution). However most random variables are not jointly elliptically distributed, and using linear correlation as a measure of dependence in such situations might prove very misleading. Even for elliptically jointly distributed random variables there are situations where using linear correlation [...] does not make sense. We might choose to model some scenario using heavy-tailed distributions such as t_2-distributions. In such cases the linear correlation coefficient is not even defined because of infinite second moments.

Moreover, we find in Embrechts *et al* (1999) a list of pitfalls within the use of linear correlation:

> Correlation is a minefield for the unwary. One does not have to search far in the literature of financial risk management to find misunderstanding and confusion. This is worrying since correlation is a central technical idea in finance [...] :

1. Correlation is simply a scalar measure of dependency; it cannot tell us everything we would like to know about the dependence structure of risks.
2. Possible values of correlation depend on the marginal distribution of the risks. All values between -1 and 1 are not necessarily attainable.
3. Perfectly positively dependent risks do not necessarily have a correlation of 1; perfectly negatively dependent risks do not necessarily have a correlation of -1.
4. A correlation of zero does not indicate independence of risks.
5. Correlation is not invariant under transformations of the risks. For example, $\log X$ and $\log Y$ generally do not have the same correlation as X and Y.
6. Correlation is only defined when the variances of the risks are finite. It is not an appropriate dependence measure for very heavy-tailed risks where variances appear infinite.

After these "complaints" against the use (and abuse) of correlation and its limitations in modeling under more general dependence relations, we proceed to show how copulas are a useful tool for this purpose.

In his pioneering work, Lehmann (1966) gives three different definitions of what he called *positive dependence.* We will just state one of them, and then we will discuss it under the framework of copulas. For random variables X and Y, the following definition compares the probability of any quadrant $X \leq x$, $Y \leq y$ under the joint distribution H of (X, Y) with the corresponding probability in the case of independence:

1.35. Definition. ***Quadrant dependence.*** We say that the pair (X, Y) or its [joint] distribution H is *positively quadrant dependent* if

$$\mathbb{P}(X \leq x, Y \leq y) \geq \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y)\,, \qquad \text{for all } x, y. \tag{1.75}$$

The dependence is strict if inequality holds for at least some pair (x, y). The family of all [joint] distributions H satisfying (1.75) will be denoted by $\mathfrak{F}_1$. Similarly, (X, Y) or H is *negatively quadrant dependent* if (1.75) holds with the inequality sign reversed, and the totality of negatively quadrant dependent distributions will be denoted by $\mathfrak{G}_1$. To simplify the notation we shall write $(X, Y) \in \mathfrak{F}$ to mean that the distribution of (X, Y) belongs to $\mathfrak{F}$.

1.36. Lemma. *(Lehmann, 1966):*

(i) $(X, X) \in \mathfrak{F}_1$ *for all* X

(ii) $(X, Y) \in \mathfrak{F}_1 \Leftrightarrow (X, -Y) \in \mathfrak{G}_1$

(iii) $(X, Y) \in \mathfrak{F}_1$ *implies* $\big(r(X), s(Y)\big) \in \mathfrak{F}_1$ *for all non-decreasing functions* r *and* s*. The concept of positive quadrant dependence is thus invariant under non-decreasing transformations (and similarly under non-increasing transformations) of both variables.*

(iv) The set of inequalities (1.75) is equivalent to that obtained by replacing one or both of the inequalities $X \leq x$ *or* $Y \leq y$ *by the corresponding* $X < x$ *or* $Y < y$*.*

(v) The set of inequalities (1.75) is equivalent to each of the following, where again the equality signs inside the probabilities are optional:

$$\mathbb{P}(X \leq x, Y \geq y) \;\leq\; \mathbb{P}(X \leq x)\mathbb{P}(Y \geq y) \tag{1.76}$$

$$\mathbb{P}(X \geq x, Y \leq y) \;\leq\; \mathbb{P}(X \geq x)\mathbb{P}(Y \leq y) \tag{1.77}$$

$$\mathbb{P}(X \geq x, Y \geq y) \;\geq\; \mathbb{P}(X \geq x)\mathbb{P}(Y \geq y). \tag{1.78}$$

Intuitively, X and Y are positively quadrant dependent if the probability that they are simultaneously small (or simultaneously large) is at least as large as it would be if they were independent. Examples of this kind of dependence are the case when $Y = s(X)$ for any random variable X and any non-decreasing function s, or $Y = X + V$ for any independent random variables X and V.

If X and Y have joint distribution function H, with continuous margins F and G, respectively, and copula C, then (1.75) is equivalent to

$$H(x, y) \;\geq\; F(x)G(y)\,, \qquad \text{for all } (x, y) \text{ in } \mathbb{R}^2\,, \tag{1.79}$$

and by virtue of Sklar's theorem:

$$C(u, v) \;\geq\; uv = \Pi(u, v)\,, \qquad \text{for all } (u, v) \text{ in } [\,0, 1\,]^2\,. \tag{1.80}$$

So in order to investigate quadrant dependence between two continuous random variables it suffices to analyze the underlying copula and compare it to the product copula.

For example, if the underlying copula is a member of the *Farlie-Gumbel-Morgenstern family* introduced in (1.14), then it is straightforward to verify that there is positive quadrant dependence whenever $\theta \geq 0$, and negative quadrant dependence whenever $\theta \leq 0$. We notice that for this family of copulas we have that $C_0 = \Pi$. Amblard and Girard (2002) proposed a semiparametric family of copulas, which includes as a particular case the Farlie-Gumbel-Morgenstern family, and it is given by:

$$C_\theta(u,v) = uv + \theta\psi(u)\psi(v)\,, \quad \theta \in [\,-1,1\,]\,, \tag{1.81}$$

where ψ must satisfy $\psi(0) = \psi(1) = 0$ and the Lipschitz condition

$$|\psi(u) - \psi(v)| \,\leq\, |u-v|\,, \qquad \text{for all } (u,v) \text{ in } [\,0,1\,]^2\,. \tag{1.82}$$

With $\psi(x) = x(1-x)$ we have the particular case of the Farlie-Gumbel-Morgenstern family. Since the *Amblard-Girard family* of copulas is characterized by ψ it is not a surprise that its authors proved that, for example, positive quadrant dependence may be determined by analyzing certain property of ψ : Let $\theta > 0$, X and Y continuous random variables with underlying Amblard-Girard copula, then X and Y are positively quadrant dependent if and only if either for all $u \in [\,0,1\,]$ $\psi(u) \geq 0$, or for all $u \in [\,0,1\,]$ $\psi(u) \leq 0$. We should notice that for this broad family of copulas, ψ plays a role similar to the generator of an Archimedean copula. It is important to mention that there exists a more general class of copulas, which may or may not be symmetric, that contains the Amblard-Girard family as a particular case, as proved by Rodríguez-Lallena and Úbeda-Flores (2004), by determining the cases in which the function C given by

$$C(u,v) = uv + f(u)g(v) \tag{1.83}$$

is a copula, for f and g real functions defined on $[\,0,1\,]$ and for all u,v in $[\,0,1\,]$.

In the case of Archimedean copulas, we note that if $\varphi(0) < \infty$ (that is, if φ is a non-strict generator) then the *support* of the Archimedean copula C it generates (recall Definition 1.9) is the set $\{(u,v) \in [\,0,1\,]^2 \,:\, \varphi(u) + \varphi(v) \leq \varphi(0)\}$; so it follows that copulas with such generator cannot have the positive quadrant dependence since for some $u,v > 0$ we have $C(u,v) = 0$. If X and Y are continuous random variables with an underlying Archimedean copula with strict generator φ, then they are positive quadrant dependent if the following functional inequality holds:

$$\varphi(u) + \varphi(v) \leq \varphi(uv)\,, \qquad \text{for all } (u,v) \text{ in } \,]\,0,1\,]^2. \tag{1.84}$$

Lehmann (1966) also proved that, provided that the covariance of X and Y exists, then $Cov(X,Y) \geq 0$ characterizes positive quadrant dependence, but we may need to talk about this type of dependence in presence of random variables whose moments do not exist, and this is a pitfall in relying on covariance, as quoted from Embrechts *et al* (1999) at the beginning of this section.

In addition to Lehmann (1966) we may find more on definitions of different types of dependence in Harris (1970), Esary and Proschan (1972), Shaked (1977), Block and Ting (1981), and Block *et al* (1982). The work of Kimeldorf and Sampson (1987, 1989) introduced a unified framework for studying and relating various concepts of positive dependence, beginning with a formal definition of what it should be understood by positive dependence. For an extensive list of types of dependence see Joe (1997), and for some equivalences of different types of dependence in terms of copulas see Nelsen (1991, 2006a).

Even though dependence relations between random variables is one of the most widely studied topics in probability theory and statistics, when reviewing the literature in this subject it is surprising to notice the lack of a formal definition for what exactly it should be understood by a *measure of dependence.* It is also common to find expressions such as *measure of association* within a context in which it is not clear if it is a synonym of *measure of dependence*, or if this last one is a particular type of a *measure of association.* Moreover, sometimes when talking about dependence, concepts such as *concordance measure* are brought into the discussion, without making clear if this last one is a particular type of *dependence measure*, or maybe a particular type of an *association measure.* It seems like for a long time there has been a "fuzzy boundary" among expressions such as *association, concordance, correlation* and *dependence.*

Blomqvist (1950) proposed what he called a *measure of dependence* between two random variables, but when defining it explicitly he uses the expression

> As a measure of correlation we define [...]

Within the context of his work the reader could hardly have a clear idea of what the author understood by *measure of dependence*, nor if *measure of correlation* is a synonym or a particular case of a measure of dependence.

Another example: Kruskal (1958) outlines the historical development of *ordinal measures of association.* Among them, the author includes well-known measures such as *Kendall's tau* and *Spearman's rho*, which are *concordance measures,* as we shall see later in this section. So, whatever it should be understood by *measure of association,* this suggests that *concordance measures* are a particular case of them. But in the introduction of his work the author started with a question:

> What is meant by the degree of association or dependence between two random variables with a joint distribution?

Should we infer that the author considers *degree of association* and *degree of dependence* as synonyms? Even by reading the entire work, one cannot answer this question because after the above quoted question the author only uses the expression *measure of association*, with no further mention to *measure of dependence* or *degree of dependence.*

As far as it was possible to investigate, a formal definition of *measure of association* was not found, just quite general and vague ideas of what is meant with this terminology. As stated by Kruskal (1958):

> There are infinitely many possible measures of association, and it sometimes seems that almost as many have been proposed at one time or another. On the other hand, it has been argued that, except in special cases, it is fatuous to attempt to represent the degree of association of a bivariate population by a single number.

Since the first edition of Nelsen (2006a), which dates from 1999, we may understand *measures of association* as a general expression which includes, among others, the following specific two types of measures:

- Measures of Concordance
  - Kendall's τ (1938)
  - Spearman's ρ (1904)
  - Gini's γ (1914)
  - Blomqvist's β (1950)
- Measures of Dependence
  - Schweizer-Wolff's σ (1981)
  - Hoeffding's Φ (1940)
  - Fernández-González Barrios's δ (2004)

We may find a different classification approach in Drouet and Kotz (2001), but we will just briefly make some comments on Nelsen's classification.

1.5.1 Measures of Concordance.

Informally, as stated by Nelsen (2006a):

> [...] a pair of random variables are concordant if "large" values of one tend to be associated with "large" values of the other and "small" values of one with "small" values of the other.

1.37. Definition. Two observations (x_i, y_i) and (x_j, y_j) from a random vector (X, Y) of continuous random variables are *concordant* if $x_i < x_j$ and $y_i < y_j$ or $x_i > x_j$ and $y_i > y_j$. Similarly, they are *discordant* if $x_i < x_j$ and $y_i > y_j$ or $x_i > x_j$ and $y_i < y_j$.

Equivalently, (x_i, y_i) and (x_j, y_j) are concordant if $(x_i - x_j)(y_i - y_j) > 0$ and discordant if $(x_i - x_j)(y_i - y_j) < 0$. For further insight into this concept see Kruskal (1958) or Lehmann (1975).

1.38. Definition. (Nelsen 2002, 2006a) Let (X_1, Y_1) and (X_2, Y_2) be random vectors of continuous random variables with (possibly) different joint distributions H_1 and H_2, but with common margins F and G. The *concordance function* Q between H_1 and H_2 is defined by

$$Q(H_1, H_2) := \mathbb{P}\big[\,(X_1 - X_2)(Y_1 - Y_2) > 0\,\big] - \mathbb{P}\big[\,(X_1 - X_2)(Y_1 - Y_2) < 0\,\big]\,. \tag{1.85}$$

Equivalently, Q is the difference between the probabilities of concordance and discordance of (X_1, Y_1) and (X_2, Y_2).

1.39. Theorem. *If in addition to Definition 1.38 the random vectors are independent, with underlying copulas C_1 and C_2, respectively, then*

$$Q(C_1, C_2) \equiv Q(H_1, H_2) = 4\iint_{[0,1]^2} C_2(u,v)\,dC_1(u,v) - 1 = Q(C_2, C_1)\,. \tag{1.86}$$

In a few words, concordance between two independent random vectors of continuous random variables may be calculated just in terms of their underlying copulas. This is crucial in defining the following three measures:

1.40. Definition. Let X and Y be continuous random variables whose copula is C. We define the following measures based on the concordance function:

a) *Kendall's tau:* $\tau_C := Q(C, C)\,,$

b) *Spearman's rho:* $\rho_C := 3Q(C,\Pi)\,,$

c) *Gini's gamma:* $\gamma_C := Q(C,M) + Q(C,W)\,,$

where Π, M and W are the product copula, and Fréchet-Hoeffding upper and lower bound copulas, respectively.

Of course, the above definitions are not the original ones, they were defined in terms of the random variables involved. For example, Kendall's tau was originally defined as

$$\tau_{XY} := \mathbb{P}\big[\,(X_1 - X_2)(Y_1 - Y_2) > 0\,\big] - \mathbb{P}\big[\,(X_1 - X_2)(Y_1 - Y_2) < 0\,\big]\,, \tag{1.87}$$

where (X_1, Y_1) and (X_2, Y_2) are independent random vectors with common joint distribution. But it has been proved, see Nelsen (2002, 2006a) and Li *et al* (2002), that the above three measures, in the case of continuous random variables, may be expressed just in terms of the underlying copula, which is not a surprise since by Sklar's theorem the dependence structure is uniquely determined by the copula.

Moreover, in defining measures of association between continuous random variables, as suggested by Drouet and Kotz (2001), we may consider the possibility of stating as a desirable property that such measures should be marginally free, that is, defined just in terms of the underlying copula. This would lead us to reject linear correlation as a measure of association, as we shall see later in this section.

Going back to Definition 1.40, we may interpret those measures as follows: Kendall's τ measures how concordant (or discordant) would be independent observations of a random vector of continuous random variables with copula C; Spearman's ρ measures how concordant (or discordant) would be independent observations of a random vector of continuous random variables with copula C compared to the product copula, or just briefly, in terms of concordance how close or far is C from the copula Π which represents independence; Gini's γ measures, in terms of concordance, how close or far is the underlying copula C from the Fréchet-Hoeefding bounds, that is, how close or far are the involved continuous random variables from being comonotonic or countermonotonic (see section 1.1).

In view of Theorem 1.39 and other results by Nelsen (2002, 2006a) and Li *et al* (2002) we may arrive to the following equivalent expressions for the discussed concordance measures:

$$\tau_C = 1 - 4\iint_{[\,0,1\,]^2} \frac{\partial}{\partial u}\,C(u,v)\,\frac{\partial}{\partial v}\,C(u,v)\,dudv\,, \tag{1.88}$$

$$\rho_C = 12\iint_{[\,0,1\,]^2} C(u,v)\,dudv \;-\; 3\,, \tag{1.89}$$

$$\gamma_C = 4\left[\int_0^1 C(u, 1-u)\, du \;-\; \int_0^1 \big[\, u - C(u,u)\,\big]\, du\right]. \tag{1.90}$$

Blomqvist (1950) proposed what he called a (sample) "measure of dependence" or (sample) "measure of correlation," which happens to be a measure based on the concept of concordance since for a random vector (X, Y) the population version of the sample version he defined would be:

$$\beta_{XY} := \mathbb{P}\big[\,(X - m_X)(Y - m_Y) > 0\,\big] - \mathbb{P}\big[\,(X - m_X)(Y - m_Y) < 0\,\big]\,, \tag{1.91}$$

where m_X and m_Y are the medians of X and Y, respectively. In words of Blomqvist (1950):

> Let $(x_1, y_1) \cdots (x_n, y_n)$ be a sample from a two-dimensional population with cdf [cumulative distribution function] $F(x, y)$, and consider the two sample medians $\tilde{x}$ and $\tilde{y}$. The cdf $F(x, y)$ is assumed to have continuous marginal cdf's $F_1(x)$ and $F_2(y)$ in order that the probability of obtaining two equal x-values or two equal y-values in the sample will be zero. Let the x, y-plane be divided into four regions by the lines $x = \tilde{x}$ and $y = \tilde{y}$. It is the clear that some information about the correlation between x and y can be obtained from the number of sample points, say n_1, belonging to the first or third quadrants compared with the number, say n_2, belonging to the second or fourth quadrants [...] As a measure of correlation we define
>
> $$q' = \frac{n_1 - n_2}{n_1 + n_2} = \frac{2n_1}{n_1 + n_2} - 1 \qquad (-1 \leq q' \leq 1)\,.$$

It is straightforward to verify, see Nelsen (2002, 2006a), that in the case of a pair of continuous random variables X and Y with underlying copula C we may calculate Blomqvist's measure just in terms of the copula:

$$\beta_C = 4\, C\left(\frac{1}{2}, \frac{1}{2}\right) \;-\; 1\,. \tag{1.92}$$

So far, the four concordance measures analyzed in this subsection have in common that they use the concept of *concordance,* each one in a different way, having therefore their own interpretation. Nelsen (2002, 2006a) adopted from Scarsini (1984) the following set of "desirable" properties for a measure of concordance:

1.41. Definition. Let X and Y be any continuous random variables with underlying copula C. A numeric measure of association between these two random variables, which will be denoted by $\kappa_{X,Y}$ or κ_C, is a *measure of concordance* if it satisfies the following properties:

1. κ is defined for every pair of continuous random variables;
2. $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,X} = 1$, $\kappa_{X,-X} = -1$;
3. $\kappa_{X,Y} = \kappa_{Y,X}$;
4. If X and Y are independent, then $\kappa_{X,Y} = \kappa_{\Pi} = 0$;
5. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;
6. If C_1 and C_2 are copulas such that $C_1(u,v) \leq C_2(u,v)$ for all (u,v) in $[0,1]^2$, then $\kappa_{C_1} \leq \kappa_{C_2}$;
7. If $\{(X_n, Y_n)\}$ is a sequence of continuous random variables with underlying copulas C_n, and if $\{C_n\}$ converges pointwise to C, then $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

As a consequence of the above definition we have the following two theorems:

1.42. Theorem. *Let κ be a measure of concordance for continuous random variables X and Y:*

1. *If Y is almost surely an increasing function of X, then $\kappa_{X,Y} = \kappa_M = 1$;*
2. *If Y is almost surely a decreasing function of X, then $\kappa_{X,Y} = \kappa_W = -1$;*
3. *If α and β are almost surely strictly monotone functions on Ran X and Ran Y, respectively, then $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$.*

1.43. Theorem. *If X and Y are continuous random variables with underlying copula C, then Kendall's τ_C, Spearman's ρ_C, Ginni's γ_C, and Blomqvist's β_C, satisfy the properties in Definition 1.41, and therefore the properties in Theorem 1.42.*

An important remark on Definition 1.41: A concordance measure equal to zero does not imply independence, so if $\kappa_{X,Y} = 0$ we will just refer to the random variables X and Y as *non-concordant.* For example, let X a continuous random variable uniformly distributed on $[-1,1]$ and define $Y := |X|$; by using (1.11) it is straightforward to verify that the four measures of concordance mentioned in Theorem 1.43 are equal to zero, in spite of the obvious dependence between X and Y.

Even though concordance measures do not characterize independence, some of them had been used to build nonparametric tests of independence with null hypothesis $\mathcal{H}_0 : \kappa = 0$ against the alternative $\mathcal{H}_1 : \kappa \neq 0$, see for example Lehmann (1975), Hollander and Wolfe (1999), Gibbons and Chakraborti (2003). We will discuss such tests in chapter 5.

We discussed in this subsection bivariate measures of concordance. For a discussion on the general case of multivariate measures of concordance see Joe (1990, 1997), Nelsen (1996a, 2002, 2006a), Taylor (2006), Dolati and Úbeda-Flores (2006).

1.5.2 Measures of Dependence.

In the same year in which Abe Sklar introduced the concept of *copula,* Rényi (1959) proposed a set of axioms for a *measure of dependence* for pairs of random variables, that is, a list of properties that certain quantities which are used to measure the strength of dependence between two random variables should satisfy. In 1959 Rényi himself showed that, among various well-known measures of dependence, the only one which satisfied all of his axioms was the *maximal correlation coefficient* for a pair of random variables X and Y defined by

$$S(X,Y) := \sup_{f\,,\,g} r\big(f(X), g(Y)\big)\,, \tag{1.93}$$

where r denotes Pearson's correlation coefficient, and the supremum is taken over all Borel functions f and g for which such correlation is defined. As mentioned by Schweizer and Wolff (1981), such measure has major drawbacks: it equals to 1 too often and is generally not effectively computable. In addition, they gave several examples which indicate that, at least for nonparametric measures, Rényi's conditions are too strong, and so they proposed a modified version of them:

1.44. Definition. Let X and Y be any continuous random variables with underlying copula C. A numeric measure of association between these two random variables, which will be denoted by $\mu_{X,Y}$ or μ_C, is a *measure of dependence* if it satisfies the following properties:

1. μ is defined for every pair of continuous random variables;
2. $0 \leq \mu_{X,Y} \leq 1$;
3. $\mu_{X,Y} = \mu_{Y,X}$;
4. X and Y are independent if and only if $\mu_{X,Y} = \mu_\Pi = 0$;
5. $\mu_{XY} = 1$ if and only if each of X and Y is almost surely a strictly monotone function of the other;
6. If α and β are strictly monotone almost surely on $Ran\,X$ and $Ran\,Y$, respectively, then $\mu_{\alpha(X),\beta(Y)} = \mu_{X,Y}$;

7. If $\{(X_n, Y_n)\}$ is a sequence of continuous random variables with underlying copulas C_n, and if $\{C_n\}$ converges pointwise to C, then $\lim_{n\to\infty} \mu_{C_n} = \mu_C$.

It is important to emphasize the fourth property since it implies that measures of dependence do characterize independence, and so concordance measures may not be considered dependence measures.

Definition 1.44 is exactly as it appears in Nelsen (2006a), in the original work by Schweizer and Wolff (1981) they also include the following property:

8. If the joint distribution of X and Y is bivariate normal, with correlation coefficient r, then $\mu_{X,Y}$ is a strictly increasing function of $|r|$.

In contrast with concordance measures where specific ways of involving the concept of concordance define each measure, Schweizer and Sklar (2005) suggested to measure how close/far is the underlying copula of continuous random variables from the only copula that represents independence:

> [...] the fact that the surface for Π (the copula of independence) lies midway between the surfaces for W and M (the copulas of extreme monotone dependence) it is natural to use any measure of distance between surfaces as a measure of dependence for pairs of random variables.

It is important to make clear in which sense Π lies "midway" between W and M. A straightforward verification shows that if C_1 and C_2 are copulas then any convex linear combination of them is also a copula, that is $C_\theta = (1-\theta)C_1 + \theta C_2$ with θ in $[\,0,1\,]$, so the surface of the copula $\frac{1}{2}(W+M)$ lies midway between W and M under the usual metric for the real line, and it is certainly not equal to Π. Schweizer and Sklar (2005) made clear that they refer to the fact that

$$M(u,v) - \Pi(u,v) \;=\; \Pi(u,1-v) - W(u,1-v) \tag{1.94}$$

for all (u,v) in $[\,0,1\,]$, whence the graph of W is a "twisted" reflection of the graph of M in the graph of Π. So, as mentioned by Schweizer and Wolff (1981):

> [...] any suitable normalized measure of distance between the surfaces $z = C(u,v)$ and $z = uv$, e.g., any L_p-distance, should yield a symmetric, nonparametric measure of dependence.

From Nelsen (2006a) we have that for any p in $[1, \infty[$ the L_p-distance between C and Π is given by

$$\mho_C(p) := \left(k_p \iint_{[0,1]^2} |C(u,v) - uv|^p dudv \right)^{1/p}, \tag{1.95}$$

where k_p is a constant chosen so that (1.95) is $\mho_C(p) = 1$ whenever C equals M or W (and so properties 2 and 5 in Definition 1.44 are satisfied).

1.45. Theorem. *For each p in $[1, \infty[$ the L_p-distance $\mho_C(p)$ is a measure of dependence, where*

$$k_p = \frac{\Gamma(2p+3)}{2\big[\Gamma(p+1)\big]^2}. \tag{1.96}$$

Particular cases are: $\mho_C(1)$ which is known as the σ of Schweizer and Wolff (1981), and $[\mho_C(2)]^2$ which is known as the *dependence index* Φ of Hoeffding (1940). In the particular case of the measure of dependence studied by Schweizer and Wolff (1981) given by

$$\sigma_C := 12 \iint_{[0,1]^2} |\, C(u,v) - uv \,|\, dudv\,, \tag{1.97}$$

we may interpret it as a measure of "average absolute distance" between the joint distribution of a pair of continuous random variables (represented by their copula C) and the independence joint distribution (represented by $\Pi(u,v) = uv$).

Schweizer and Wolff (1981) also studied the properties of the L_∞-distance between C and Π given by

$$\delta_{X,Y} = \delta_C := 4 \sup_{(u,v) \in [0,1]^2} |\, C(u,v) - uv \,|\,, \tag{1.98}$$

and they proved that this measure satisfies all except property 5 in Definition 1.44. We will discuss more about this measure later in this subsection under the work by González-Barrios (2003b) and Fernández and González-Barrios (2004).

Embrechts *et al* (2003a) proposed "desired properties of a dependence measure," which we will denote ξ, for a pair of random variables:

E0. ξ assigns a real number to any pair of random variables X and Y;

E1. $\xi(X,Y) = \xi(Y,X)$ (symmetry);

E2. $-1 \leq \xi(X,Y) \leq 1$ (normalization);

E3. $\xi(X,Y) = 1$ if and only if X and Y are comonotonic;

$\xi(X,Y) = -1$ if and only if X and Y are countermonotonic;

E4. For $T : \mathbb{R} \to \mathbb{R}$ strictly monotonic on the range of X :

$$\xi(T(X),Y) = \begin{cases} \xi(X,Y) & T \text{ increasing,} \\ -\xi(X,Y) & T \text{ decreasing.} \end{cases}$$

E5. $\xi(X,Y) = 0$ if and only if X and Y are independent;

but they also proved that there is no dependence measure satisfying E4 and E5: Let (X,Y) be uniformly distributed on the unit circle S^1 in $\mathbb{R}^2$, so that $(X,Y) = (\cos\Theta, \sin\Theta)$ with Θ uniformly distributed on $]\,0, 2\pi\,[$. Since $(-X,Y)$ has the same distribution as (X,Y) we have

$$\xi(-X,Y) = \xi(X,Y) = \xi(X,-Y)\,,$$

which implies $\xi(X,Y) = 0$ although X and Y are dependent. With the same argumentation it can be shown that the measure is zero for any spherical distribution in $\mathbb{R}^2$.

If E5 is required, Embrechts *et al* (2003a) suggested to consider measures which only assign non-negative values, and in such case they proposed the amended properties:

E2b. $0 \le \xi(X,Y) \le 1$;

E3b. $\xi(X,Y) = 1$ if and only if X and Y are comonotonic or countermonotonic;

E4b. For $T : \mathbb{R} \to \mathbb{R}$ strictly monotonic: $\xi(T(X),Y) = T(X,Y)$.

We notice that E0, E1, E2b, E3b, E4b, and E5 are equivalent and/or weaker properties than some of those included in Definition 1.44. In opinion of Embrechts *et al* (2003a):

> If we restrict ourselves to the case of continuous random variables there are dependence measures that fulfill all of E1, E2b, E3b, E4b and E5, although they are in general measures of theoretical rather than practical interest [...] The disadvantage of all of these measures is that they are constrained to give non-negative values and as such cannot differentiate between positive and negative dependence and that it is often not clear how to estimate them.

Drouet and Kotz (2001) proposed “desirable properties of a measure of dependence”:

(D1) Standardization: The values of an index are between 0 and 1.

(D2) Independence: If X and Y are independent, the index should be zero.

(D3) Functional dependence: If one variable is a function of the other, the index must be equal to 1.

(D4) Increasing property: The index should increase as the dependence increases.

(D5) Invariance: The index should be invariant with respect to a linear transformation of the variables. A stronger condition would be that the index is marginally free, i.e. the measure of the dependence is the same as the corresponding measure on the copula.

(D6) Symmetry: If the variables are exchangeable, then the index should be symmetric.

(D7) Relationship with measures for ordinal variables: If the index is defined for both ordinal and continuous variables, there should be a close relationship between the two measures.

(D8) Interpretability. This is a very delicate and intangible property. Roughly speaking, it means that the numerical value of this index can be translated into a qualitative meaningful measure.

A few comments on some of the above proposed properties, in the case of continuous random variables: (D2) is not an if-and-only-if condition, so this would be compatible with the definition of a concordance measure rather than those we have discussed for dependence measures which do imply independence whenever they are equal to zero; (D3) may be too strong, not only because we would have to discard L_p-distances, also because with no restrictions on the functional relationship it is possible to build dependence relationships between two continuous random variables for which it would be difficult to reach the maximum value 1, as we will see later in this subsection; it is not clear what should it be understood by “increasing dependence” in (D4); the invariance property as stated in (D5) may be considered as extremely weak since it is a common and desirable required property that the dependence measure remains unchanged at least under strictly increasing transformations of the variables (not only linear) due to the fact that the underlying copula is the same (see Theorem 1.34).

Wolff (1980) worked on n-dimensional extensions of the bivariate dependence measures (1.97) and (1.98), and for such purpose he proposed higher dimensional analogues of modified Rényi’s axioms, designed as requisites for a symmetric nonparametric measure of dependence:

1.46. Definition. Let $(X_1,\ldots,X_n)$ be a random vector of continuously distributed random variables. A number R is a *multivariate dependence measure* if it satisfies the following:

W1. R is defined for any $(X_1,\ldots,X_n)$.

W2. $R(X_1,\ldots,X_n) = R(\wp(X_1,\ldots,X_n))$ for all permutations $\wp$ of $(X_1,\ldots,X_n)$.

W3. $0 \leq R(X_1,\ldots,X_n) \leq 1$.

W4. $R(X_1,\ldots,X_n) = 0$ if and only if $X_1,\ldots,X_n$ are independent.

W5. $R(X_1,\ldots,X_n) = 1$ if and only if each of $X_1,\ldots,X_n$ is an increasing function almost surely of the others.

W6. If $f_1,\ldots,f_n$ are all strictly increasing, then $R(X_1,\ldots,X_n) = R(f_1(X_1),\ldots,f_n(X_n))$.

W7. Let the joint distribution of $(X_1,\ldots,X_n)$ be multivariate normal and let r_{ij} be the correlation coefficient of X_i and X_j. If the r_{ij} are either all nonnegative or all nonpositive, then R is a strictly increasing function of each of the $|r_{ij}|$.

W8. If the sequence $\{(X_{1m},\ldots,X_{nm})\}$ converges in law to $(X_1,\ldots,X_n)$, then

$$\lim_{m\to\infty} R(X_{1m},\ldots,X_{nm}) = R(X_1,\ldots,X_n).$$

Let $f : [\,0,1\,]^n \to [\,0,1\,]$ be any integrable function, and let $\mathcal{I}$ denote the operator

$$\mathcal{I}(f) := \int\!\!\cdots\!\!\int_{[\,0,1\,]^n} f(u_1,\ldots,u_n)\,du_1\cdots du_n\,. \tag{1.99}$$

Recalling definitions in (1.17), we have that

$$\mathcal{I}(W^{(n)}) = \frac{1}{(n+1)!}\,, \qquad \mathcal{I}(\Pi^{(n)}) = \frac{1}{2^{\,n}}\,, \qquad \mathcal{I}(M^{(n)}) = \frac{1}{n+1}\,. \tag{1.100}$$

Wolff (1980) noticed that $\mathcal{I}(M^{(2)} - \Pi^{(2)}) = \mathcal{I}(\Pi^{(2)} - W^{(2)})$ (which is a consequence of (1.94)), but that this symmetry breaks down in higher dimensions. Moreover, by defining

$$a_n := \mathcal{I}(M^{(n)} - \Pi^{(n)})\,, \qquad b_n := \mathcal{I}(\Pi^{(n)} - W^{(n)})\,, \tag{1.101}$$

we have that $\lim_{n\to\infty}(b_n/a_n) = 0$, that is, as n increases, the graphs of $z = W^{(n)}(u_1,\ldots,u_n)$ and $z = \Pi^{(n)}(u_1,\ldots,u_n)$ are much closer to one another than the graphs of $z = M^{(n)}(u_1,\ldots,u_n)$ and $z = \Pi^{(n)}(u_1,\ldots,u_n)$.

Let $n \geq 2$ and let $(X_1, \ldots, X_n)$ have continuous marginal distributions and copula C. By defining the multivariate versions of (1.97) and (1.98):

$$\sigma_n := a_n^{-1} \mathcal{I}(\,|\,C - \Pi^{(n)}\,|\,)\,, \tag{1.102}$$

$$\delta_n := t_n^{-1} \mathcal{S}(\,|\,C - \Pi^{(n)}\,|\,)\,, \tag{1.103}$$

where the operator $\mathcal{S}$ is

$$\mathcal{S}(f) := \sup_{(u_1,\ldots,u_n)\,\in\,[\,0,1\,]^{\,n}} |\,f(u_1, \ldots, u_n)\,|\,, \tag{1.104}$$

and

$$t_n := \mathcal{S}(\,|\,M^{(n)} - \Pi^{(n)}\,|\,) = \frac{n-1}{n^{(n/(n-1))}}\,, \tag{1.105}$$

Wolff (1980) proved the following:

1.47. Theorem. *For any $n > 2$, the quantity σ_n satisfies all the conditions in Definition 1.46, and the quantity δ_n all except W5.*

González-Barrios (2003b) and Fernández and González-Barrios (2004) proposed what they called a *multidimensional dependency measure:*

$$\delta_{X_1,\ldots,X_n} := \sup_{(x_1,\ldots,x_n)\,\in\,\mathbb{R}^{\,n}} \left| F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) - \prod_{i\,=\,1}^{n} F_{X_i}(x_i) \right|, \tag{1.106}$$

where $F_{X_1,\ldots,X_n}$ is the joint distribution function of the X_i's and F_{X_i} is the distribution function of X_i. They proved the following:

1.48. Theorem. *For any random variables $X_1, \ldots, X_n$:*

FG1. $\delta_{X_1,\ldots,X_n} = \delta_{X_{\omega(1)},\ldots,X_{\omega(n)}}$ *for every permutation ω of* $\{1, 2, \ldots, n\}$.

FG2. $\delta_{X_1,\ldots,X_n} = 0$ *if and only if $X_1, \ldots, X_n$ are independent.*

FG3. For every $x_i \in \mathbb{R}$ and $i \in \{1, 2, \ldots, n\}$

$$-\left(\frac{n-1}{n}\right)^n \leq F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) - \prod_{i\,=\,1}^{n} F_{X_i}(x_i) \leq \left(\frac{1}{n}\right)^{\frac{1}{n-1}} \left(1 - \frac{1}{n}\right) < 1\,.$$

Hence $0 \leq \delta_{X_1,\ldots,X_n} \leq 1$. Besides the bounds above can be attained.

FG4. $0 \leq \delta_{X_1,X_2} \leq \delta_{X_1,X_2,X_3} \leq \cdots \leq \delta_{X_1,\ldots,X_{n-1}} \leq \delta_{X_1,\ldots,X_n}$.

An important remark on the above theorem is the fact that the random variables are not required to be continuous. In case they are, if their underlying copula is C then by virtue of Sklar's theorem we have that (1.106) becomes

$$\delta_C := \sup_{(u_1,\ldots,u_n)\in[0,1]^n} \Big| C(u_1,\ldots,u_n) - \prod_{i=1}^{n} u_i \Big| . \tag{1.107}$$

So far, what we find in the literature is that there is no consensus on what a definition of a *measure of concordance* or a *measure of dependence* should be. Some authors have proposed what they think are useful measures, analyzed their properties, and then proposed such properties as a set of "desirable properties" for any concordance or dependence measure. It is beyond the purpose of this thesis to provide final definitions in these matters, but we will make some additional remarks:

A1. In the case of continuous random variables, for all the measures of concordance and dependence analyzed in this section it had been always possible to calculate them in terms of the underlying copula, which should not be a surprise since Sklar's theorem suggests that the dependence structure is uniquely determined by such copula, therefore (as suggested by Drouet and Kotz in D5) we may consider to require dependence or concordance measures to be computable just in terms of the corresponding copula.

A2. Drouet and Kotz (2001) proposed, among others, the following "desirable property" for a measure of dependence:

> (D4) Increasing property: The index should increase as the dependence increases.

Although it was not clarified by the authors in which sense dependence is considered to be "increased", for the multivariate case we may consider including in this category property FG4 in Theorem 1.48.

A3. We began section 1.4 with quotations from Drouet and Kotz (2001) and Embrechts (1999, 2003a) which strongly question the general usefulness of the linear correlation coefficient. Among the complaints, Embrechts (1999) includes:

> Correlation is not invariant under transformations of the risks. For example, $\log X$ and $\log Y$ generally do not have the same correlation as X and Y.

Recalling Theorem 1.34, we have that the underlying copula for a vector (X,Y) of continuous random variables is the same as for $(\alpha(X),\beta(Y))$ with α and β strictly increasing functions, and this suggests to require concordance and dependence measures to be invariant, at least, under strictly increasing transformations of the random variables. It is in this sense why we already mentioned that we consider the "desirable property" D5 in Drouet and Kotz (2001) extremely weak because it only requires invariance under linear transformations of the variables.

A4. Another complaint against the linear correlation coefficient was the following by Embrechts (2003a):

> [...] linear correlation is not a copula-based measure of dependence, it can often be quite misleading and should not be taken as the canonical dependence measure.

The following lemma is due to Hoeffding (1940), but it did not become widely known until it was quoted by Lehmann (1966):

1.49. Lemma. *If H denotes the joint and F and G the marginal distributions of random variables X and Y, then*

$$\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \iint_{\mathbb{R}^2} \big[\, H(x,y) - F(x)G(y) \,\big]\, dxdy\,, \tag{1.108}$$

provided the expectations on the left hand side exist.

In the particular case X and Y are continuous, if the corresponding copula is C, then by making the substitution $u = F(x)$ and $v = G(y)$ we obtain the following expression for the linear correlation coefficient via Corollary 1.8:

$$r(X,Y) = \frac{1}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} \iint_{[0,1]^2} \big[\, C(u,v) - uv \,\big]\, dF^{-1}(u) dG^{-1}(v)\,. \tag{1.109}$$

The linear correlation coefficient is not a function of C alone, it is not a copula-based measure since it also depends on the marginal behavior of X and Y, and this might be misleading since, without changes on the dependence structure uniquely determined by the copula, the linear correlation coefficient may report different values just by changing the marginal distributions.

A5. In Lancaster (1963) and Kimeldorf and Sampson (1978) a random variable Y is defined to be *completely dependent* on a random variable X if there exists a function g such that

$$\mathbb{P}\,[\, Y \,=\, g(X)\,] \,=\, 1\,. \tag{1.110}$$

As mentioned by Kimeldorf and Sampson (1978):

> Y is completely dependent on X if Y is perfectly predictable from X. The random variables X and Y are defined (see Lancaster (1963)) to be mutually completely dependent (MCD) if Y is completely dependent on X and X is completely dependent

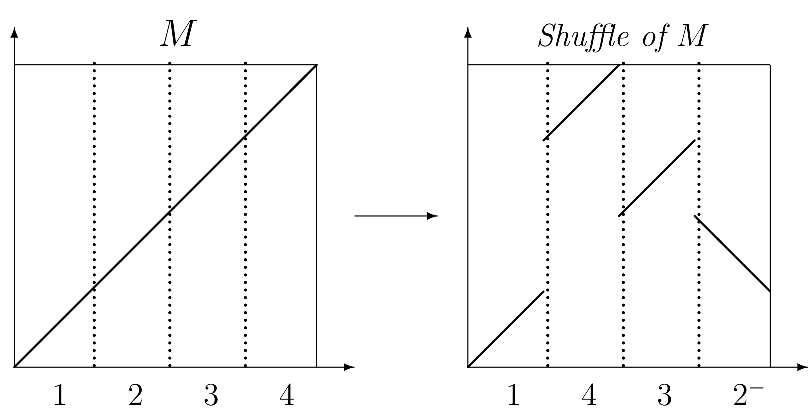
> on Y. Equivalently, X and Y are MCD if (1.110) holds for some one-to-one function g. The concept of mutual complete dependence is, in a real sense, directly opposite to that of stochastic independence, in that mutual complete independence entails complete predictability of either random variable from the other, while stochastic independence entails complete unpredictability.

There is a result which proves that there exist mutually completely dependent random variables whose joint distribution function approximates as much as desired the distribution function of two independent random variables with the same marginals, a result that could make questionable to consider, for example, strict monotone relationships between random variables as the opposite extreme of independence in Definition 1.44(5), or E3b in Embrechts (2003a), or D3 in Drouet and Kotz (2001), or W5 in Wolff's Definition 1.46.

By Definition 1.9 it is straightforward to verify that the Fréchet-Hoeffding upper bound copula $M(u,v) = \min(u,v)$ is a singular copula whose support is the main diagonal of $[0,1]^2$, that is the graph of $v = u$ for u in $[0,1]$. Mikusiński *et al* (1991, 1992) constructed a particular type of singular copulas that they called *Shuffles of Min* as follows:

> The mass distribution for a shuffle of M can be obtained by (1) placing the mass for M on $[0,1]^2$, (2) cutting $[0,1]^2$ vertically into a finite number of strips, (3) shuffling the strips with perhaps some of them flipped around their vertical axes of symmetry, and then (4) reassembling them to form the square again. The resulting mass distribution will correspond to a copula called a shuffle of M.

For example:



Formally (Nelsen, 2006a), a shuffle of M is determined by a positive integer n, a finite partition $\{J_i\} = \{J_1, J_2, \cdots, J_n\}$ of $[0,1]$ into n closed subintervals, a permutation γ on the set $S_n := \{1, 2, \ldots, n\}$, and a function $\omega : S_n \to \{-1, 1\}$ where $\omega(i)$ is -1 or 1 according to whether or not the strip $J_i \times [0, 1]$ is flipped. We denote permutations by the vector of images $(\gamma(1), \gamma(2), \ldots, \gamma(n))$. The resulting shuffle of M may then be unambiguously denoted by $M(n, \{J_i\}, \gamma, \omega)$, where n is the number of connected components in its support.

The probabilistic interpretation of a *Shuffle of Min* given by Mikusiński *et al* (1991):

> [...] we say that a function f is strongly piecewise strictly monotone if and only if [the extended real plane] $\overline{\mathbb{R}}^2$ can be partitioned into a finite number of rectangles such that in each column and in each row of rectangles there is exactly one rectangle having a nonempty intersection with the graph of f and that portion of the graph of f is strictly monotone [...] The copula for (X, Y) is a shuffle of Min if and only if X and Y are strongly piecewise strictly monotone functions of each other.

Equivalently, if the copula of (X, Y) is a shuffle of Min then, because the support of any shuffle is the graph of a one-to-one function, it follows that X and Y are mutually completely dependent (MCD) in the sense of Lancaster (1963).

1.50. Theorem. ***(Mikusiński et al, 1991)*** *The shuffles of Min are dense in the set of all copulas endowed with the* sup *norm. That is, for any given copula C and for all $\varepsilon > 0$ there exists a shuffle of M, which we denote C_ε, such that*

$$\sup_{u,v \in [0,1]} |\, C(u, v) - C_\varepsilon(u, v)\,| < \varepsilon\,. \tag{1.111}$$

If we choose $C = \Pi$, as remarked by Nelsen (2006a), the above theorem implies that there are MCD random variables whose joint distribution functions are arbitrarily close to the joint distribution function of independent random variables with the same marginals. As noted in Mikusiński *et al* (1991), this implies that in practice, the behavior of any pair of independent continuous random variables can be approximated so closely by a pair of MCD continuous random variables that it would be impossible, experimentally, to distinguish one pair from the other. Moreover, under these ideas Kimeldorf and Sampson (1978) proved that it is possible to construct a sequence $\{(X_n, Y_n)\}$ of pairs of MCD random variables which converge in law to a pair (X, Y) of independent random variables. This is the reason why the "desired property" D3 in Drouet and Kotz (2001) might be too strong.

This last result recalls the classical discussion on *Determinism,*[1] a philosophical point of view defended by Laplace[2] in the $18^{\underline{th}}$ century which states that all events are completely

[1]The New Encyclopædia Britannica, Ready Reference (1998), Vol.4, p.39.

[2]Encyclopedia of Statistical Sciences (1982), Vol.1, pp.405-411.

determined, and that “randomness” is just an euphemism for the impossibility (possibly temporary) to explain a certain behavior in a deterministic way. The shuffles of Min might suggest, from a deterministic point of view, that what we pretend to state as independence could be sometimes lack of ability in finding a relationship. So it seems that there is still a lot of work to be done on the problem of dependence among random variables, how to define it and measure it.

Chapter 2

Empirical diagonal and properties

We have seen so far that, in the case of Archimedean bivariate copulas, the diagonal section contains all the information we need to build the copula in case Frank's condition $\delta'(1-) = 2$ is satisfied, and in such case this leads us to concentrate in studying and estimating the diagonal. The main benefit of this fact is a reduction in the dimension of the estimation, from 2 to 1 in the case of bivariate copulas, and from m to 1 in the case of m-variate copulas. The results in this chapter for the bivariate case are included in Erdely and González-Barrios (2006b).

2.1 Bivariate case.

Let $S := \{(x_1, y_1), \ldots, (x_n, y_n)\}$ denote a sample of size n from a random vector of continuous random variables (X, Y). As defined by Nelsen (2006a), the *bivariate empirical copula* is the function C_n given by

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n} \sum_{(x\,,\,y)\,\in\,S} \mathbf{1}_{]\,-\,\infty\,,\,x_{(i)}\,]\times]\,-\,\infty\,,\,y_{(j)}\,]}(x, y)\,,$$

where $x_{(i)}$ and $y_{(j)}$ denote the order statistics of the sample, for i and j in $\{1, \ldots, n\}$, and $C_n(\frac{i}{n}, 0) = 0 = C_n(0, \frac{j}{n})$.

Remark. A bivariate empirical copula is not a copula, but a (two-dimensional) *subcopula*, for details of subcopulas see Nelsen (2006a).

2.1. Definition. The *bivariate empirical diagonal* is the function δ_n given by

$$\delta_n\left(\frac{j}{n}\right) := C_n\left(\frac{j}{n}, \frac{j}{n}\right) \qquad j = 0, 1, \ldots, n\,.$$

Without loss of generality we may assume that the x_k values in S are ordered, then

$$\delta_n\left(\frac{j}{n}\right) = \frac{1}{n}\sum_{k=1}^{j} \mathbf{1}_{]-\infty\,,\,y_{(j)}\,]}(y_k)\,, \quad j = 1,\ldots,n-1\,, \tag{2.1}$$

and $\delta_n(0) = 0$, $\delta_n(1) = 1$. It is clear from above that δ_n is a nondecreasing function of j. Moreover, by Fréchet-Hoeffding bounds :

$$\max\left(\frac{2j}{n} - 1, 0\right) \leq \delta_n\left(\frac{j}{n}\right) \leq \frac{j}{n}\,. \tag{2.2}$$

2.2. Proposition.

$$\delta_n\left(\frac{j+1}{n}\right) - \delta_n\left(\frac{j}{n}\right) \in \left\{0, \frac{1}{n}, \frac{2}{n}\right\}. \tag{2.3}$$

Proof:

$$\begin{aligned}
\delta_n\left(\frac{j+1}{n}\right) - \delta_n\left(\frac{j}{n}\right) &= \frac{1}{n}\Big[\mathbf{1}_{]-\infty\,,\,y_{(j+1)}\,]}(y_{j+1}) + \sum_{k=1}^{j}\Big(\mathbf{1}_{]-\infty\,,\,y_{(j+1)}\,]}(y_k) - \mathbf{1}_{]-\infty\,,\,y_{(j)}\,]}(y_k)\Big)\Big]\,, \\
&= \frac{1}{n}\Big[\mathbf{1}_{]-\infty\,,\,y_{(j+1)}\,]}(y_{j+1}) + \sum_{k=1}^{j}\mathbf{1}_{]\,y_{(j)}\,,\,y_{(j+1)}\,]}(y_k)\Big]\,, \\
&= \frac{1}{n}\Big[\mathbf{1}_{]-\infty\,,\,y_{(j+1)}\,]}(y_{j+1}) + \sum_{k=1}^{j}\mathbf{1}_{\{y_{(j+1)}\}}(y_k)\Big]\,,
\end{aligned}$$

and since any of the last two indicator functions may independently take the value 0 or 1 the result follows. □

This means that all the possible paths $\{\delta_n(\frac{j}{n}) \,:\, j = 0,1,\ldots,n\}$ are between the paths $\{\max(\frac{2j}{n} - 1, 0) \,:\, j = 0,1,\ldots,n\}$ and $\{\frac{j}{n} \,:\, j = 0,1,\ldots,n\}$ with jumps of size 0, $\frac{1}{n}$, or $\frac{2}{n}$ between consecutive steps.

Let X be a continuous uniform random variable in $]\,0,1\,[$ and define the random variable $Y := X$. Then the corresponding copula for (X,Y) is the Fréchet-Hoeffding upper bound copula $M(u,v) := \min(u,v)$. In this case, a size n sample of observations of (X,Y) would be $S = \{(x_1,x_1),\ldots,(x_n,x_n)\}$, and applying formula (2.1) we get $\delta_n(\frac{j}{n}) = \frac{j}{n}$, which is the Fréchet-Hoeffding upper bound in (2.2). If, instead, we define $Y := 1 - X$, the corresponding copula for (X,Y) is the Fréchet-Hoeffding lower bound copula $W(u,v) := \max(u+v-1,0)$, and $\delta_n(\frac{j}{n})$ equals the lower bound in (2.2).

Extensive work has been done to study convergence of the empirical copula C_n to the true copula C : Deheuvels (1979, 1981), van der Vaart and Wellner (1996), Fermanian *et al* (2004). These results automatically guarantee the convergence of the empirical diagonal δ_n to δ_C, the true diagonal section of the copula.

2.3. Definition. Let us define the *bivariate diagonal random path* by the vector

$$\mathbf{T} = \left(\Delta\left(\frac{0}{n}\right), \Delta\left(\frac{1}{n}\right), \ldots, \Delta\left(\frac{n}{n}\right)\right)$$

where

$$\Delta\left(\frac{j}{n}\right) := \frac{1}{n}\sum_{k=1}^{j} \mathbf{1}_{]-\infty, Y_{(j)}]}(Y_k)\,, \quad j = 1, \ldots, n-1\,, \tag{2.4}$$

with $\Delta(0) = 0\,,\ \Delta(1) = 1\,.$

Associated to the diagonal random path we may define the *diagonal random increments* by the vector $\mathbf{I} := \left(\Delta\left(\frac{1}{n}\right) - \Delta\left(\frac{0}{n}\right), \Delta\left(\frac{2}{n}\right) - \Delta\left(\frac{1}{n}\right), \ldots, \Delta\left(\frac{n}{n}\right) - \Delta\left(\frac{n-1}{n}\right)\right)$ so that knowledge of $\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)$ is equivalent to knowledge of $\mathbf{I} = (i_1, \ldots, i_n)\,,$ where $i_j = t_j - t_{j-1}$ and $t_j = \sum_{k=1}^{j} i_k\,.$

Alternatively, we may write $\mathbf{I} = (i_1, \ldots, i_n) = \frac{1}{n}(b_1, \ldots, b_n)$ where the $b_j \in \{0, 1, 2\}\,.$ Moreover, $i_n = 1 - t_{n-1} = 1 - \sum_{k=1}^{n-1} i_k$ so knowledge of $\mathbf{B} := (b_1, \ldots, b_{n-1})$ completely specifies any path. Different values of $\mathbf{B}$ can be labeled as vectors of ternary numbers. For example, with $n = 7$ the Fréchet-Hoeffding lower bound path $\{\max(\frac{2j}{n} - 1, 0) : j = 0, 1, \ldots, n\}$ is specified by the vector $(0,0,0,1,2,2)$, while the Fréchet-Hoeffding upper bound path $\{\frac{j}{n} : j = 0, 1, \ldots, n\}$ is specified by the vector $(1,1,1,1,1,1)$.

In general, the upper bound path is specified by the ternary representation of

$$\sum_{k=0}^{n-2} 3^k = \frac{3^{n-1} - 1}{2}\,, \tag{2.5}$$

while for the lower bound path we need the ternary representation of

$$2 \cdot \sum_{k=0}^{n/2-2} 3^k = 3^{n/2-1} - 1\,, \quad \text{if } n \text{ even}, \tag{2.6}$$

$$2 \cdot \sum_{k=0}^{(n-1)/2-2} 3^k = 2 \cdot 3^{(n-1)/2-1} - 1\,, \quad \text{if } n \text{ odd}. \tag{2.7}$$

But not every ternary representation between (2.6) or (2.7) and (2.5) will generate a valid path. For example, for $n = 6$ we have that the ternary representation $(0,2,2,1,0)$ is a number between $(0,0,0,2,2)$ and $(1,1,1,1,1)$ but the generated path is out of Fréchet-Hoeffding bounds. In general, we just have to check which of the ternary representations satisfy

$$\max\left(\frac{2j}{n} - 1, 0\right) \le t_j \le \frac{j}{n} \quad \equiv \quad \max(2j - n, 0) \le \sum_{k=1}^{j} b_k \le j\,, \quad j = 1, \ldots, n-1\,. \tag{2.8}$$

We will call an *admissible diagonal path* any vector of ternary numbers satisfying the Fréchet-Hoeffding conditions (2.2), or equivalently, conditions (2.8). We wish to count the number of admissible paths for $n \geq 2$. For this purpose, we have to recall the concept of *Catalan numbers,* see for example Barcucci and Verri (1992):

> The sequence of Catalan numbers, defined by
> $$E_m := \frac{1}{m+1}\binom{2m}{m}, \tag{2.9}$$
> has been widely studied and it has been proved that it is the sequence that enumerates a lot of classes of combinatorial objects, such as the partitioning of a polygon into triangles, the bracketing of the non-associative product of $m+1$ terms, the binary trees with n nodes, and some walks on the integral lattice.

For the case of walks on the integral lattice, we look at walks of $m+k$ unit steps into upward and rightward directions, starting at the origin $(0,0)$ and ending at (m,k). The number of such paths without further restrictions is $\binom{m+k}{k}$, as exactly k of the $m+k$ steps are upward steps. Now consider just those upward-rightward paths with $k \leq m$, that is, paths remaining on or under the diagonal. For this to happen it is necessary to have at any step of the path an accumulated number of rightward steps equal o larger that the number of upward steps: "a rightward step before any upward step."

For the case $k = m$ it is proved in Theorem 3.1 in Barcucci and Verri (1992) that E_m is the number of the under-diagonal rightward-upward one-step walks on the integral lattice. An equivalent result is the classical *Chung-Feller Theorem*, see Chung and Feller (1949), or Feller (1968) in its chapter *On fluctuations in coin-tossing and random walks.* For the general case, by a result in Bailey (1996), we may calculate the number of under-diagonal rightward-upward one-step walks on the integral lattice, starting at $(0,0)$ and ending at (m,k), by

$$\frac{m+1-k}{m+1}\binom{m+k}{m}, \quad k \leq m\,. \tag{2.10}$$

In its original version, Bailey (1996) obtains (2.10) by counting the number of sequences with non-negative partial sums that consist of m positive 1's and k negative 1's. An equivalent result is found in Engleberg (1965).

In the the following result we find the exact number of admissible paths for any $n \geq 2$.

2.4. Proposition. *Let P_n denote the number of admissible paths for the bivariate empirical diagonal δ_n of n points in $[0,1]^2$ and $n \geq 2$. Then*

$$P_n = \sum_{r_0=0}^{[\![n/2]\!]} \frac{\binom{n}{r_0}\binom{n-r_0}{r_0}}{r_0+1}$$

where $[\![x]\!]$ *denotes the greatest integer less than or equal to* x.

Proof: Let $(x_1, y_1), \ldots, (x_n, y_n) \in [0,1]^2$ be n points and let δ_n be their empirical diagonal. Let r_0, r_1 and r_2 denote the number of zeros, ones and twos, respectively, in the empirical diagonal $n\delta_n$. Then $r_0 + r_1 + r_2 = n$ and $r_1 + 2r_2 = n$, so $r_0 = r_2$. Hence

$$r_1 + 2r_0 = n, \tag{2.11}$$

any nonnegative integers r_0 and r_1 satisfying (2.11), could provide an admissible path for the empirical diagonal whenever the Fréchet-Hoeffding bounds are satisfied. Observe that $r_0 \leq [\![n/2]\!]$, by (2.8).

If $\delta_n = \frac{1}{n}(b_1, b_2, \ldots, b_n)$ is an admissible path, then $b_1 = 0$ or $b_1 = 1$, and for $2 \leq i \leq n$ $b_i = 0, 1$ or 2. We observe that the restrictions $\sum_{i=1}^{n} b_i = n$ and $\sum_{i=1}^{j} b_i \leq j$ must be fulfilled. The *basic rule* to find admissible paths is "zero before two." That is if some $b_i = 2$, then there exists $1 \leq j < i$ such that $b_j = 0$. For example if $n = 5$ and $r_0 = 1$, then $(0, 1, 1, 2, 1)$ is an admissible path, but $(2, 1, 1, 0, 1)$ is not admissible, since for example $\sum_{i=1}^{3} b_i = 4 \not\leq 3$. Therefore, given the number of zeros among the b_i's, we only have to see where can they be located in the vector $(b_1, \ldots, b_n)$, following the basic rule. Observe that the ones can be located any place.

So first assume that $r_0 = 0$ and $r_1 = n$, then the only path, which by the way is admissible, is $(1, 1, \ldots, 1)$, that is, Fréchet-Hoeffding upper bound path. Now fix some r_0 such that $1 \leq r_0 \leq [\![n/2]\!]$. We have to count all the admissible paths that follow the basic rule "zero before two." Since the ones can be located any place, we just have to count the different ways in which we can locate the zeros and the twos. First, we have to choose $r_0 + r_2 = 2r_0$ places for the zeros an twos out of the n places available, which can be made in $\binom{n}{2r_0}$ different ways. This last number has to be multiplied by the number of different ways in which we can locate the r_0 zeros and the $r_2 = r_0$ twos in the $2r_0$ chosen places, but always following the basic rule. We may relate the zeros to rightward unit steps and the twos to upward unit steps in (2.10) with $m = k = r_0$, so the number of admissible paths with r_0 zeros is given by

$$\binom{n}{2r_0}\binom{2r_0}{r_0}\frac{1}{r_0+1} = \frac{1}{r_0+1}\binom{n}{r_0}\binom{n-r_0}{r_0}. \tag{2.12}$$

The result now follows summing over all possible values of r_0. $\square$

Remark. (2.12) also simplifies in terms of a multinomial coefficient to

$$\frac{1}{r_0+1}\binom{n}{r_0,\, r_1,\, r_2}, \tag{2.13}$$

where $r_2 = r_0$ and $r_0 + r_1 + r_2 = n$. This may be understood as follows: the multinomial coefficient along with the restriction $r_0+r_1+r_2 = n$ represents the number of permutations of repeated elements (r_0 zeros, r_1 ones, and r_2 twos, in this case), but some of these permutations do not follow the basic rule "zero before two," so the role of the factor $\frac{1}{r_0+1}$ is to leave just those permutations that follow the rule. To understand how this factor is obtained, we need some combinatorics concepts as in, for example, Callan (2006), considering the problem of building sequences of 1's and -1's, and their partial sums:

> A *balanced* n-*path* is a sequence of n [1's] and n [-1's], represented as a path of unit upsteps $(1,1)$ and downsteps $(1,-1)$ from $(0,0)$ to $(2n,0)$ [...] A *Dick* n-*path* is a balanced n-path that never drops below the de x-axis (ground level) [...] the parameter X on balanced n-paths defined by $X =$ "number of upsteps above ground level" is uniformly distributed over $\{0,1,2,\ldots,n\}$ and hence divides the $\binom{2n}{n}$ balanced n-paths into $n+1$ equal-size classes, one of which consists of the Dick n-paths (the one with $X = n$). Indeed, for $1 \leq i \leq n$, a bijection from balanced n-paths with $X = 0$ (inverted Dick paths) to those with $X = i$ is as follows. Number the upsteps from right to left and top to bottom, starting with the last upstep. Then remove the first downstep $[-1]$ encountered directly west of upstep i to obtain the subpaths P and Q, and reassemble as $Q\,[-1]\,P$.

So the problem of counting the different ways of allocating the zeros and twos following the basic rule is the same as counting the number of Dick n-paths, using the 1's to represent zeros, and the -1's to represent twos (remember that the ones may be allocated any place, so they will be allocated in the remaining places), and the result (2.13) follows.

Now we will calculate the probability of any given (admissible) path, under the hypothesis of independence.

2.5. Theorem. *Let $S = \{(X_1,Y_1),\ldots,(X_n,Y_n)\}$ be a random sample from the random vector of continuous random variables (X,Y). If X and Y are independent and if $\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)$ is an admissible diagonal path, then*

$$P\big[\,\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)\,\big] = \frac{1}{n!}\prod_{j=1}^{n} f(j)\,,$$

where $f(j)$ is obtained in terms of the following formula, for $j = 1,\ldots,n$:

$$f(j) = \begin{cases} 1 & \text{if} \quad n(t_j - t_{j-1}) = 0\,, \\ 2(j - n\,t_{j-1}) - 1 & \text{if} \quad n(t_j - t_{j-1}) = 1\,, \\ (j - 1 - n\,t_{j-1})^2 & \text{if} \quad n(t_j - t_{j-1}) = 2\,. \end{cases}$$

Proof: From the continuity assumption we know that, with probability one, there are no ties among the $X_i{}'s$ or the $Y_i{}'s$. Without loss of generality we may assume that the X_k $(k = 1, \ldots, n)$ are ordered, so we have, by independence of X and Y, that the probability of the random sample $\{(X_1, Y_1), \ldots, (X_n, Y_n))\}$ equals that of $\{(X_1, Y_{\sigma(1)}), \ldots, (X_n, Y_{\sigma(n)})\}$ where $\sigma(1), \ldots, \sigma(n)$ is any permutation of $(1, \ldots, n)$, and every permutation has probability $(n!)^{-1}$.
By rescaling we can assume that $X_i = i$, for $i = 1, 2, \ldots, n$ and $Y_{\sigma(i)} = \sigma(i)$, for $i = 1, \ldots, n$. Hence the sample S is a subset of the grid $\{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\} := I_n^2$. In fact, for every $i \in \{1, 2, \ldots, n\}$ there exists a unique $j = \sigma(i) \in \{1, 2, \ldots, n\}$ such that $(i, \sigma(i)) \in S$. That is, for any horizontal or vertical segment in the grid I_n^2 there is exactly one point that belongs to the sample S.
In order to calculate $P\big[\mathbf{T} = (0, t_1, \ldots, t_{n-1}, 1)\big]$ we just need to count the number of orderings of $\{Y_1, \ldots, Y_n\}$ that would lead to the admissible path $(t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)$. We will show that this probability is given by $(n!)^{-1} \prod_{j=1}^{n} f(j)$.
Let $\mathbf{B} = (b_1, b_2, \ldots, b_n)$ be the vector of ternary numbers which is equivalent to the admissible diagonal path $\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)$, that is $b_i = n(t_i - t_{i-1})$ for $i = 1, 2, \ldots, n$. Define

$$K := \min\{i \in \{1, \ldots, n\} \,|\, b_i > 0\}.$$

Since $\mathbf{B}$ is an admissible diagonal path, we have that all the $b_i{}'s$ are equal to $0, 1$ or 2, except for b_1 which equals 0 or 1, and $\sum_{i=1}^{n} b_i = n$. Then $K \leq [\![n/2]\!]$ by (2.8). Therefore, if $K = 1$ it means that $b_1 = 1$, and then $(1, 1) \in S$, and there is only one possibility for $\sigma(1)$, that is $\sigma(1) = 1$.
So, assume that $K > 1$, from the definition of the bivariate empirical copula it is clear that

$$nt_i = nC_n\left(\frac{i}{n}, \frac{i}{n}\right) = \operatorname{card}(S \cap (\{1, \ldots, i\} \times \{1, \ldots, i\})) \quad \text{for} \quad i = 1, 2, \ldots, n,$$

where $\operatorname{card}(\cdot)$ stands for cardinality of a set. So nt_i gives us the number of sample points in the sample S that belong to the sub-grid $\{1, \ldots, i\}^2$. Since $K > 1$, then $b_1 = \cdots = b_{K-1} = 0$, which is equivalent to $t_1 = \cdots = t_{K-1} = 0$.
Now, first assume that $b_K = 1$, or equivalently $nt_K = 1$. Then we observe that the intersection of the sub-grid $\{1, \ldots, K-1\} \times \{1, \ldots, K-1\}$ and the sample S is empty, but the intersection of the sub-grid $\{1, \ldots, K\} \times \{1, \ldots, K\}$ and the sample S contains a unique point. By noticing that $(\{1, \ldots, K\} \times \{1, \ldots, K\}) \backslash (\{1, \ldots, K-1\} \times \{1, \ldots, K-1\}) = \{(1, K), (2, K), \ldots, (K, K), (K, K-1), \ldots, (K, 1)\}$ we can select the point of the sample in $2K - 1 = 2(K - nt_{K-1}) - 1$ forms, if for example we select the point $(2, K)$ then we know that any point of the form $(2, j)$ for $j \neq K$, and any point of the form (l, K) with $l \neq 2$ do

not belong to the sample, that is, we cancel one column and one row and the remaining points that were not selected.
The other possibility is $b_K = 2$. Then we observe that the intersection of the sub-grid $\{1, \ldots, K-1\} \times \{1, \ldots, K-1\}$ and the sample S is empty, but the intersection of the sub-grid $\{1, \ldots, K\} \times \{1, \ldots, K\}$ and the sample S contains exactly two points. Just as above we know that, $(\{1, \ldots, K\} \times \{1, \ldots, K\}) \backslash (\{1, \ldots, K-1\} \times \{1, \ldots, K-1\}) = \{(1, K), (2, K), \ldots, (K, K), (K, K-1), \ldots, (K, 1)\}$ contains two points of the sample S. Observe that (K, K) can not be a sample point, since in that case, none of the points $(1, K), \ldots (K-1, K), (K, K-1), \ldots, (K, 1)$ can belong to the sample. Therefore we can select one point from $(1, K), \ldots, (K-1, K)$ and another from $(K, K-1), \ldots, (K, 1)$, that is we have $(K-1)^2 = (K-1-nt_{K-1})^2$ possible choices. After selecting these two points we can not repeat the same indexes for columns or rows, so we cancel two columns and two rows and the remaining points which were not selected.
Now we define

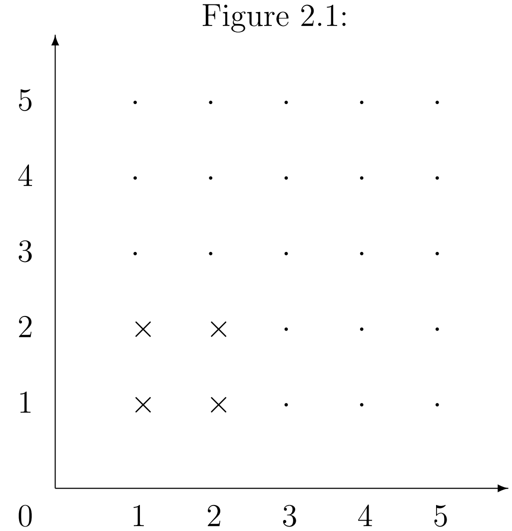
$$K_1 := \min\{i \in \{K+1, \ldots, n\} \mid b_i > 0\},$$

and we proceed inductively by reducing the dimension of the grid. As an example consider that $n = 5$, and the admissible path is given by $\mathbf{T} = (0 = t_0, t_1 = 0, t_2 = 0, t_3 = 1/5, t_4 = 3/5, t_5 = 5/5 = 1)$, or equivalently $\mathbf{B} = (b_1 = 0, b_2 = 0, b_3 = 1, b_4 = 2, b_5 = 2)$, in this case

$$K = \min\{i \in \{1, \ldots, 5\} \mid b_i > 0\} = 3.$$

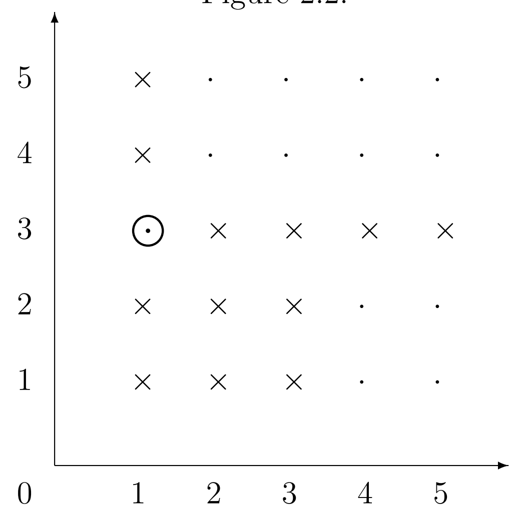
We first notice that $(1,1), (1,2), (2,1)$ and $(2,2)$ are not sample points, since $K = 3$, see Figure 2.1.

Figure 2.1:



Now, $b_3 = 1$, so we have to select only one point in the set $\{(1,3), (2,3), (3,3), (3,2), (3,1)\}$, that is we have $5 = 2(3 - 5t_2) - 1$ choices. Assume we select $(1, 3)$, then we cancel the remaining elements not selected and those of the first column and the third row, see Figure 2.2.

Figure 2.2:



Now

$$K_1 = \min\{i \in \{4, 5\} \,|\, b_i > 0\} = 4,$$

and $b_4 = 2$, in this case we have to select one point between $(2, 4)$ and $(3, 4)$ and another between $(4, 1)$ and $(4, 2)$, that is $2^2 = (4 - 1 - 5t_3)^2 = (3 - 5(1/5))^2$ choices, recall that $(4, 4)$ can not be selected. Assume we select $(3, 4)$ and $(4, 1)$, so we cancel the third and fourth columns and the first and fourth row, and the points that were not selected, see Figure 2.3.

Figure 2.3:

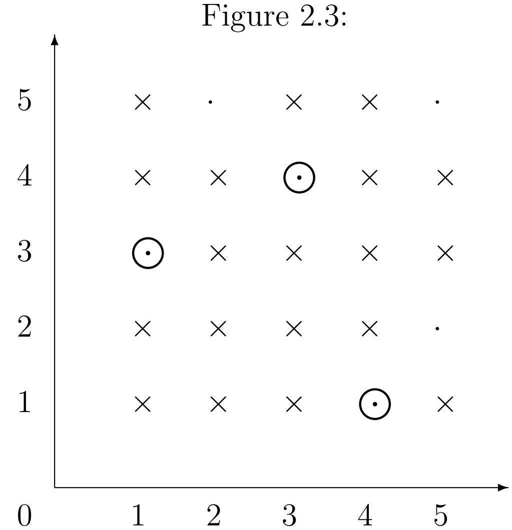
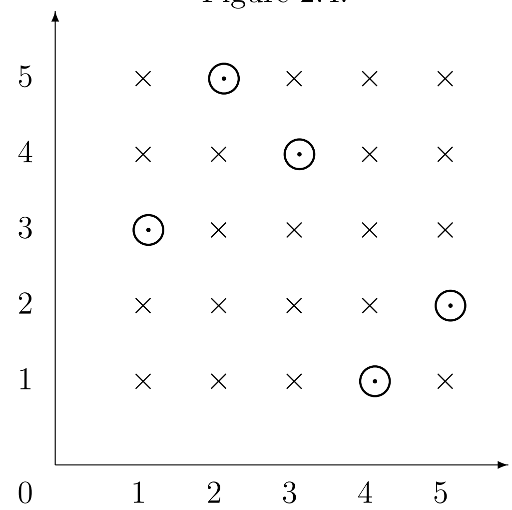


Figure 2.4:



Finally,

$$K_2 = \min\{i \in \{5\} \,|\, b_i > 0\} = 5,$$

since $b_5 = 2$ we have two select two points between $(2,5)$ and $(5,2)$ recall that $(5,5)$ can not be selected, this can be done only in $1 = 1^2 = (5-1-5t_4)^2 = (5-1-5(3/5))^2$ way, see Figure 2.4.
Therefore the number of permutations that lead to the diagonal path T is $1{\cdot}1{\cdot}5{\cdot}2^2{\cdot}1^2 = 20$, and hence the probability of T is $20/5! = 1/6$. $\square$

2.2 Trivariate case and further.

We recall from (1.17) and (1.18) that if C is any m-copula, then for every $(u_1, \ldots, u_m)$ in $[\,0,1\,]^m$ we have that

$$\max(u_1 + \cdots + u_m - m + 1\,,\, 0) \;\le\; C(u_1, \ldots, u_m) \;\le\; \min(u_1, \ldots, u_m)\,, \tag{2.14}$$

but the Fréchet-Hoeffding lower bound is never a copula for $m > 2$, and the above inequality cannot be improved, see Theorem 1.10. According to (2.14) we have that the diagonal section of an m-copula satisfies

$$\max(mu - m + 1, 0) \;\le\; \delta(u) \;\le\; u\,, \qquad u \in [\,0,1\,]\,. \tag{2.15}$$

Particularly, the product (or independence) m-copula $\Pi^{(m)}(u_1, \ldots, u_m) = u_1 u_2 \cdots u_m$ has a diagonal section $\delta_\Pi(u) = u^m$. For an Archimedean m-copula, from Kimberling (1974) we have that its generator must be strict and completely monotonic, see Definition 1.30 and Theorem 1.31, and in such case we have the following expression for its diagonal section:

$$\delta(u) \;=\; \varphi^{-1}\big(m\varphi(u)\big)\,, \quad u \in [\,0,1\,]^{m}\,, \tag{2.16}$$

or equivalently

$$\varphi\big(\delta(u)\big) \;=\; m\varphi(u)\,, \tag{2.17}$$

which again leads us to Schröder's functional equation, see (1.59). As a particular case of Theorem 6.6 in Kuczma (1968) (or Theorem 2.3.12 in Kuczma *et al* (1990)), let the function $\gamma : [\,0,1\,] \to [\,0,1\,]$ be such that $0 < \gamma(u) < u$ for all $u \in \,]\,0,1\,[\,$, and $\gamma\,'(0) = \frac{1}{m}$. If $s(u)$ is a solution of the functional equation

$$s\big(\gamma(u)\big) \;=\; \frac{1}{m}\,s(u) \tag{2.18}$$

such that the function $s(u)/u$ is monotonic in $]\,0,1\,[$, then

$$s(u) \;=\; k \lim_{r\to\infty} m^{\,r}\gamma^{\,r}(u)\,, \tag{2.19}$$

where $\gamma^{\,r}$ is the r-*th iteration* of γ, that is the composition of γ with itself r times, and k any constant. And by a similar argument as in Frank's Theorem, see Theorem 1.33, we have that if C is an Archimedean m-copula whose diagonal δ satisfies

$$\delta\,'(1-) \;=\; m \tag{2.20}$$

then it is uniquely determined by its diagonal. That is the case, for example, of the product m-copula, which in the context of a m-dimensional random vector of continuous random variables represents independence.

We will now analyze analogous properties of the empirical diagonal as done in the previous section, for the case $m = 3$, hoping that this suffices to convince the reader that analogous results may be obtained for higher dimensions.

Let $S := \{(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)\}$ denote a sample of size n from a random vector of continuous random variables (X, Y, Z). Analogously as defined by Nelsen (2006a), we may define the *trivariate empirical copula* as the function C_n given by

$$C_n\left(\frac{i}{n}, \frac{j}{n}, \frac{k}{n}\right) = \frac{1}{n} \sum_{(x\,,\,y\,,\,z)\,\in S} \mathbf{1}_{]\,-\infty\,,\,x_{(i)}\,]\times]\,-\infty\,,\,y_{(j)}\,]\times]\,-\infty\,,\,z_{(k)}\,]}(x, y, z)\,,$$

where $x_{(i)}$, $y_{(j)}$ and $z_{(k)}$ denote the order statistics of the sample, for i, j and k in $\{1, \ldots, n\}$, and $C_n(x, y, z) = 0$, whenever any of x, y or z equals 0.

2.6. Definition. The *trivariate empirical diagonal* is the function δ_n given by

$$\delta_n\left(\frac{j}{n}\right) := C_n\left(\frac{j}{n},\frac{j}{n},\frac{j}{n}\right) \qquad j = 0, 1, \ldots, n\,.$$

Without loss of generality we may assume that the x_k values in S are ordered, then

$$\delta_n\left(\frac{j}{n}\right) = \frac{1}{n}\sum_{k=1}^{j} \mathbf{1}_{]-\infty\,,\,y_{(j)}\,]\times]-\infty\,,\,z_{(j)}\,]}(y_k, z_k)\,, \quad j = 1, \ldots, n-1\,, \tag{2.21}$$

and $\delta_n(0) = 0$, $\delta_n(1) = 1$. It is clear from above that δ_n is a nondecreasing function of j. Moreover, by Fréchet-Hoeffding bounds :

$$\max\left(\frac{3j}{n} - 2, 0\right) \leq \delta_n\left(\frac{j}{n}\right) \leq \frac{j}{n}\,. \tag{2.22}$$

2.7. Proposition.

$$\delta_n\left(\frac{j+1}{n}\right) - \delta_n\left(\frac{j}{n}\right) \in \left\{0, \frac{1}{n}, \frac{2}{n}, \frac{3}{n}\right\}. \tag{2.23}$$

Proof: Define the set $A(j) :=]-\infty, y_{(j)}\,]\times]-\infty, z_{(j)}\,]$. Then

$$\begin{aligned} n\left[\delta_n\left(\frac{j+1}{n}\right) - \delta_n\left(\frac{j}{n}\right)\right] &= \sum_{k=1}^{j+1}\mathbf{1}_{A(j+1)}(y_k, z_k) - \sum_{k=1}^{j}\mathbf{1}_{A(j)}(y_k, z_k)\,, \\ &= \mathbf{1}_{A(j+1)}(y_{j+1}, z_{j+1}) + \sum_{k=1}^{j}\Big[\mathbf{1}_{A(j+1)}(y_k, z_k) - \mathbf{1}_{A(j)}(y_k, z_k)\Big]\,, \\ &= \mathbf{1}_{A(j+1)}(y_{j+1}, z_{j+1}) + \sum_{k=1}^{j}\mathbf{1}_{A(j+1)\setminus A(j)}(y_k, z_k)\,. \end{aligned}$$

Since the last two indicator functions may take values 0 or 1 independently, and since the set $A(j+1)\setminus A(j)$ may contain 0, 1 or 2 points, the result follows. $\square$

This means that all the possible paths $\{\delta_n(\frac{j}{n}) \,:\, j = 0, 1, \ldots, n\}$ are between the paths $\{\max(\frac{3j}{n} - 2, 0) \,:\, j = 0, 1, \ldots, n\}$ and $\{\frac{j}{n} \,:\, j = 0, 1, \ldots, n\}$ with jumps of size $0, \frac{1}{n}, \frac{2}{n}$ or $\frac{3}{n}$ between consecutive steps.

2.8. Definition. Let us define the *trivariate diagonal random path* by the vector

$$\mathbf{T} = \left(\Delta\left(\frac{0}{n}\right), \Delta\left(\frac{1}{n}\right), \ldots, \Delta\left(\frac{n}{n}\right)\right)$$

where

$$\Delta\left(\frac{j}{n}\right) := \frac{1}{n}\sum_{k=1}^{j}\mathbf{1}_{A(j)}(Y_k, Z_k)\,, \quad j = 1, \ldots, n-1\,, \tag{2.24}$$

with $\Delta(0) = 0$, $\Delta(1) = 1$, and $A(j) :=]-\infty, Y_{(j)}]\times]-\infty, Z_{(j)}]$.

Associated to the diagonal random path we may define the *diagonal random increments* by the vector $\mathbf{I} := \left(\Delta\left(\frac{1}{n}\right) - \Delta\left(\frac{0}{n}\right), \Delta\left(\frac{2}{n}\right) - \Delta\left(\frac{1}{n}\right), \ldots, \Delta\left(\frac{n}{n}\right) - \Delta\left(\frac{n-1}{n}\right)\right)$ so that knowledge of $\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)$ is equivalent to knowledge of $\mathbf{I} = (i_1, \ldots, i_n)$, where $i_j = t_j - t_{j-1}$ and $t_j = \sum_{k=1}^{j} i_k$.

Alternatively, we may write $\mathbf{I} = (i_1, \ldots, i_n) = \frac{1}{n}(b_1, \ldots, b_n)$ where the $b_j \in \{0, 1, 2, 3\}$. Moreover, $i_n = 1 - t_{n-1} = 1 - \sum_{k=1}^{n-1} i_k$ so knowledge of $\mathbf{B} := (b_1, \ldots, b_{n-1})$ completely specifies any path. Different values of $\mathbf{B}$ can be labeled as vectors of base-4 numbers. For example, with $n = 7$ the Fréchet-Hoeffding lower bound path $\{\max(\frac{3j}{n} - 2, 0) : j = 0, 1, \ldots, n\}$ is specified by the vector $(0, 0, 0, 0, 1, 3)$, while the Fréchet-Hoeffding upper bound path $\{\frac{j}{n} : j = 0, 1, \ldots, n\}$ is specified by the vector $(1, 1, 1, 1, 1, 1)$.

But not every base-4 representation will generate a valid path. For example, for $n = 7$ we have that $(0, 3, 0, 0, 2, 0)$ is a base-4 number between $(0, 0, 0, 0, 1, 3)$ and $(1, 1, 1, 1, 1, 1)$, but it represents a path that is out of Fréchet-Hoeffding bounds. In general, we just have to check which of the base-4 representations satisfy

$$\max\left(\frac{3j}{n} - 2, 0\right) \leq t_j \leq \frac{j}{n}, \quad j = 1, \ldots, n-1, \tag{2.25}$$

which is equivalent to satisfy

$$\max(3j - 2n, 0) \leq \sum_{k=1}^{j} b_k \leq j, \quad j = 1, \ldots, n-1. \tag{2.26}$$

We will call an *admissible diagonal path* any vector of base-4 numbers satisfying the Fréchet-Hoeffding conditions (2.22), or equivalently, conditions (2.26). In the the following result we find the exact number of admissible paths for any $n \geq 3$.

2.9. Proposition. *Let P_n denote the number of admissible paths for the trivariate empirical diagonal δ_n of n points in $[0,1]^3$ and $n \geq 3$. Then*

$$P_n = \sum_{r_0=0}^{[\![2n/3]\!]} \sum_{r_2 + 2r_3 = r_0} \frac{\binom{n}{r_0}\binom{n-r_0}{r_3}\binom{n-r_0-r_3}{r_2}}{r_0+1} = \sum_{r_0=0}^{[\![2n/3]\!]} \sum_{r_3=0}^{[\![r_0/2]\!]} \frac{\binom{n}{r_0}\binom{n-r_0}{r_3}\binom{n-r_0-r_3}{r_0-2r_3}}{r_0+1},$$

where $[\![x]\!]$ denotes the greatest integer less than or equal to x.

Proof: Let $(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n) \in [0,1]^3$ be n points and let δ_n be their empirical diagonal. Let r_0, r_1, r_2 and r_3 denote the number of zeros, ones, twos, and threes, respectively, in the empirical diagonal $n\delta_n$. Then $r_0 + r_1 + r_2 + r_3 = n$ and $r_1 + 2r_2 + 3r_3 = n$, so $r_2 + 2r_3 = r_0$. Hence, any nonnegative integers r_1, r_0, r_3 satisfying

$$r_1 + 2r_0 - r_3 = n, \tag{2.27}$$

$$r_0 \geq 2r_3\,, \tag{2.28}$$

could provide an admissible path for the empirical diagonal whenever the Fréchet-Hoeffding bounds are satisfied. Observe that $r_0 \leq [\![\,2n/3\,]\!]$, by (2.26), and that (2.28) is a consequence of the fact that $0 \leq r_2 = r_0 - 2r_3$.

If $\delta_n = \frac{1}{n}(b_1, b_2, \ldots, b_n)$ is an admissible path, then $b_1 \in \{0,1\}$, $b_2 \in \{0,1,2\}$, and $b_i \in \{0,1,2,3\}$ for $3 \leq i \leq n$. We observe that the restrictions $\sum_{i=1}^{n} b_i = n$ and $\sum_{i=1}^{j} b_i \leq j$ must be fulfilled. Since $r_0 = r_2 + 2r_3$ now the basic rule to find admissible paths is "one zero before each two, two zeros before each three". That is if some $b_i = 2$, then there exists $1 \leq j < i$ such that $b_j = 0$, and if some $b_i = 3$, then there exist $1 \leq j < k < i$ such that $b_j = 0 = b_k$. Therefore, given the number of zeros and threes among the $b_i's$, we only have to see where can they be located in the vector $(b_1, \ldots, b_n)$, following the basic rule, with $r_2 = r_0 - 2r_3$. Observe that the ones can be located any place.

First assume that $r_0 = 0$, then (2.28) implies $r_3 = 0$, and $r_2 = r_0 - 2r_3 = 0$, so the only possibility is $r_1 = n$, that is, the admissible path $(1, 1, \ldots, 1)$, which is Fréchet-Hoeffding upper bound path. Then, by an analogous argument as in Proposition 2.4, for any r_0 positive integer such that $r_0 \leq [\![\,2n/3\,]\!]$, and any nonnegative integer r_3 such that $r_2 + 2r_3 = r_0$, the number of admissible paths with r_0 zeros and r_3 threes is given by

$$\frac{\binom{n}{r_0}\binom{n-r_0}{r_3}\binom{n-r_0-r_3}{r_2}}{r_0+1} = \frac{1}{r_0+1}\binom{n}{r_0,\, r_1,\, r_2,\, r_3}\,, \tag{2.29}$$

where the right side of this last equation is justified by analogous arguments as in the Remark of Proposition 2.4. The result now follows summing over all possible values of r_0 and r_3, subject to the constraint $r_2 + 2r_3 = r_0$, which is equivalent to sum

$$\frac{\binom{n}{r_0}\binom{n-r_0}{r_3}\binom{n-r_0-r_3}{r_0-2r_3}}{r_0+1}\,, \tag{2.30}$$

over all possible values of $r_0 \leq [\![\,2n/3\,]\!]$, and $r_3 \leq [\![\,r_0/2\,]\!]$, by (2.28). □

Remark. A thorough justification for (2.29) may be given in terms of Riordan Group enumeration techniques, as in Cameron (2002).

Now we will calculate the probability of any given (admissible) path, under the hypothesis of independence.

2.10. Theorem. *Let $S = \{(X_1, Y_1, Z_1), \ldots, (X_n, Y_n, Z_n)\}$ be a random sample from the random vector of continuous random variables (X, Y, Z). If X, Y and Z are independent and if $\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)$ is an admissible diagonal path, then*

$$P\big[\,\mathbf{T} = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1)\,\big] = \frac{1}{(n!)^2}\prod_{j=1}^{n} f(j)\,,$$

where $f(j)$ is obtained in terms of the following formula, for $j = 1, \dots, n$:

$$f(j) = \begin{cases} 1 & \text{if} \quad n(t_j - t_{j-1}) = 0\,, \\ 3(j - nt_{j-1})(j - 1 - nt_{j-1}) + 1 & \text{if} \quad n(t_j - t_{j-1}) = 1\,, \\ 3(j - 1 - nt_{j-1})^4 & \text{if} \quad n(t_j - t_{j-1}) = 2\,, \\ (j - 1 - nt_{j-1})^3 (j - 2 - nt_{j-1})^3 & \text{if} \quad n(t_j - t_{j-1}) = 3\,. \end{cases}$$

Proof: From the continuity assumption we know that, with probability one, there are no ties among the $X_i{}'s$, the $Y_i{}'s$ or the $Z_i{}'s$. By independence of X , Y and Z, the probability of the random sample $\{(X_1, Y_1, Z_1), \dots, (X_n, Y_n, Z_n))\}$ equals that of $\{(X_{(1)}, Y_{\sigma(1)}, Z_{\tau(1)}), \dots, (X_{(n)}, Y_{\sigma(n)}, Z_{\tau(n)})\}$ where $(\sigma(1), \tau(1)), \dots, (\sigma(n), \tau(n))$ is any bivariate permutation of $I_n := \{1, \dots, n\}$, and every permutation has probability $(n!)^{-2}$.

By rescaling we can assume that $X_{(i)} = i$, $Y_{\sigma(i)} = \sigma(i)$, and $Z_{\tau(i)} = \tau(i)$, for $i \in I_n$, that is to consider the one-to-one rank-mapping $(X_i, Y_j, Z_k) \mapsto (rank(X_i), \sigma(i), \tau(i))$. Hence the rank-mapped sample S becomes a subset of the (three-dimensional) grid $\{1, 2, \dots, n\}^3 := I_n^3$. In fact, for every $i \in I_n$ there exists a unique $(j, k) = (\sigma(i), \tau(i)) \in I_n^2$ such that $(i, j, k) = (i, \sigma(i), \tau(i)) \in S$. That is, for any horizontal or vertical segment in the three bivariate grids $\{i\} \times I_n^2$, $I_n \times \{j\} \times I_n$, and $I_n^2 \times \{k\}$, there is exactly one point that belongs to the sample S.

In order to calculate $P\big[\mathbf{T} = (0, t_1, \dots, t_{n-1}, 1)\big]$ we just need to count the number of orderings of $\{Y_1, \dots, Y_n\}$ and $\{Z_1, \dots, Z_n\}$ that would lead to the admissible path $(t_0 = 0, t_1, \dots, t_{n-1}, t_n = 1)$. We will show that this probability is given by $(n!)^{-2} \prod_{j=1}^n f(j)$.

Let $\mathbf{B} = (b_1, b_2, \dots, b_n)$ be the vector of base-4 numbers which is equivalent to the admissible diagonal path $\mathbf{T} = (t_0 = 0, t_1, \dots, t_{n-1}, t_n = 1)$, that is $b_i = n(t_i - t_{i-1})$ for $i = 1, 2, \dots, n$.

If $b_k = 0 = n(t_k - t_{k-1})$ then there is only one possibility: no sample point of S is rank-mapped to $\{1, \dots, k\}^3 \setminus \{1, \dots, k-1\}^3$.

Define

$$K := \min\{i \in \{1, \dots, n\} \,|\, b_i > 0\}.$$

Since $\mathbf{B}$ represents an admissible diagonal path, we have that all the $b_i{}'s$ are equal to $0, 1, 2$ or 3, except for b_1 which equals 0 or 1 , and b_2 which equals $0, 1$ or 2, and $\sum_{i=1}^n b_i = n$. Then $K \le [\![\, 2n/3 \,]\!]$ by (2.25).

If $K = 1$ it means that $b_1 = 1$, and then $(1, 1, 1) \in S$, and there is only one possibility for $\sigma(1)$ and $\tau(1)$, that is $(\sigma(1), \tau(1)) = (1, 1)$.

Define $\mathcal{D}_1 := \{(1,1,1)\}$ and for $r = 2,3,\ldots$ let $\mathcal{D}_r := \{1,2,\ldots,r\}^3 \setminus \bigcup_{w=1}^{r-1} \mathcal{D}_w$. Then $\text{card}(\mathcal{D}_r) = r^3 - (r-1)^3 = 3\,r(r-1)+1$, where $\text{card}(\,\cdot\,)$ stands for cardinality of a set. Geometrically, $\mathcal{D}_r$ may be interpreted as a grid on three faces of a cube of volume r^3, with one *vertex,* $3(r-1)$ points on the three edges (excluding the vertex), which we will call *edge points*, and therefore

$$\frac{r^3-(r-1)^3-3(r-1)-1}{3} = (r-1)^2$$

points on each face (without edges), which we will call *face points*. So in $\mathcal{D}_r$ we have $3(r-1)^2$ face points, $3(r-1)$ edge points, and exactly 1 vertex. All points $(i,j,k) \in \mathcal{D}_r$ must have at least one entry equal to r. If a point in $\mathcal{D}_r$ has only one entry equal to r then it is a face point; if it has 2 entries equal to r and the other one different from r then it is an edge point; and if it has its 3 entries equal to r it is obviously the vertex of $\mathcal{D}_r$.

If $K = 2$ it means that $b_1 = 0 = nt_1$, that is $(1,1,1) \notin S$, and $b_2 \in \{1,2\}$. In case $b_2 = 1$, there is only one point $(X_i, Y_j, Z_k) \in S$ which is rank-mapped to one of the elements of $\mathcal{D}_2$, that is, there are $\text{card}(\mathcal{D}_2) = 2^3 - 1^3 = 7$ possibilities for such point. In case $b_2 = 2$, there are exactly two points $(X_i, Y_j, Z_k) \in S$ which are rank-mapped to 2 different elements of $\mathcal{D}_2$, and the number of possibilities depends on whether one of the points belongs to one of the $3(2-1)+1 = 4$ points that lie on the 3 edges of $\mathcal{D}_2$. First of all, we have to discard the vertex $(2,2,2)$ since this point belongs to the three edges of $\mathcal{D}_2$, and this would eliminate the possibility of using any other point in the three faces of $\mathcal{D}_2$, and we need to allocate two points. If one of the points is an edge point, then it automatically eliminates the possibility of choosing the other point from 2 faces and the 3 edges of $\mathcal{D}_2$, that is, the other point has to be a face point, so at least one of the two points has to be a face point. So first we count the number of ways in which we can choose the face point, which is 3 : $(1,1,2)$, $(1,2,1)$, $(2,1,1)$; then, its selection eliminates 2^2 points on the face where it is located (the vertex included) plus $(2-1) = 1$ face points on each of the other two faces, and so there is left $7 - 2^2 - 2(1) = 1$ possibility for the other point, for a total of $3(1) = 3$ different ways of choosing the two points.

Now assume that $K \geq 3$, and therefore $b_K \in \{1,2,3\}$. This implies that $b_1 = \cdots = b_{K-1} = 0$, which is equivalent to $t_1 = \cdots = t_{K-1} = 0$, that is, there are no points in the sample S which are rank-mapped to the set $\bigcup_{w=1}^{K-1} \mathcal{D}_w = \{1,\ldots,K-1\}^3$. From the definition of the trivariate empirical copula it is clear that

$$nt_i = nC_n\!\left(\frac{i}{n},\frac{i}{n},\frac{i}{n}\right) = \text{card}(S \cap \{1,\ldots,i\}^3)\,,$$

that is, nt_i is the number of (rank-mapped) sample points in S that belong to $\{1,\ldots,i\}^3$. If $b_K = 1$, or equivalently $nt_K = 1$, and since $nt_{K-1} = 0$ implies that there are not any points in S rank-mapped to $\{1,\ldots,K-1\}^3$, we have that there is exactly one point of S which is rank-mapped to $\mathcal{D}_K$, and there are $\mathrm{card}(\mathcal{D}_K) = K^3 - (K-1)^3 = 3K(K-1)+1$ different possibilities to choose this point. It is important to mention that the corresponding rank-mapped point, say (i^*, j^*, k^*), automatically cancels the possibility that any other point of the sample S is rank-mapped to a point $(i,j,k) \in \{1,\ldots,n\}^3$ such that $i = i^*$ or $j = j^*$ or $k = k^*$.

If $b_K = 2$ then there are exactly 2 sample points of S which are rank-mapped to 2 different elements of $\mathcal{D}_K$, and the number of possibilities depends on whether one of the points belongs to one of the $3(K-1)+1$ points that lie on the 3 edges of $\mathcal{D}_K$. First of all, we have to discard the vertex (K,K,K) since this point belongs to the three edges of $\mathcal{D}_K$, and this would eliminate the possibility of using any other point in the three faces of $\mathcal{D}_K$, and we need to allocate two points. If one of the points is an edge point then it automatically eliminates the possibility of choosing the other point from 2 faces and the 3 edges of $\mathcal{D}_K$, that is, the other point has to be a face point, so at least one of the two points has to be a face point. So first we count the number of ways in which we can choose the face point, which is $\binom{3}{1}(K-1)^2$; then, its selection eliminates K^2 points on the face where it is located, and so it eliminates $2K-1$ points of the K^2 points on each of the other two faces, that is, $K^2 - (2K-1) = (K-1)^2$ points are left on each of the other two faces; we may choose one out of the two faces left and so we have $\binom{2}{1}(K-1)^2$ different possibilities for the second point, and so we have a total of

$$\frac{\binom{3}{1}(K-1)^2\binom{2}{1}(K-1)^2}{2!} = 3(K-1)^4$$

different ways of choosing the two points (we divided by 2! since the order of the two points chosen is not important).

If $b_K = 3$ then there are exactly 3 sample points of S which are rank-mapped to 3 different elements of $\mathcal{D}_K$, which necessarily have to be face points (one on each of the three faces of $\mathcal{D}_K$) since the presence of one edge point would just leave one (or zero in the case of the vertex) faces for choosing the other two points, which is impossible since it is only possible to have one point per face. Then, we have $\binom{3}{1}(K-1)^2$ different ways of choosing the first point, which just leaves available $(K-1)^2 - (K-1) = (K-1)(K-2)$ points on each of the other two faces, so we may choose the second point in $\binom{2}{1}(K-1)(K-2)$ different ways, which in turn will eliminate $(K-2)$ points of the remaining face, leaving $\binom{1}{1}[(K-1)(K-2) - (K-2)] = \binom{1}{1}(K-2)^2$ ways of choosing the third point, for a total of

$$\frac{\binom{3}{1}(K-1)^2\binom{2}{1}(K-1)(K-2)\binom{1}{1}(K-2)^2}{3!} = [(K-1)(K-2)]^3$$

different ways for choosing the 3 points (we divided by 3! since the order of the three points chosen is not important).

For $b_J{}'s$ with $J > K$ we have that $b_J \in \{0,1,2,3\}$ and we proceed in an analogous way, but eliminating the points $(i,j,k) \in \mathcal{D}_J$ for which there exists $(i^*,j^*,k^*) \in \bigcup_{w=1}^{J-1} \mathcal{D}_w$ $= \{1,\ldots,J-1\}^3$ such that $i = i^*$ or $j = j^*$ or $k = k^*$. For the calculations we proceed in an analogous ways as for $j = K$ by just eliminating for each of the three dimensions nt_{J-1} points since nt_{J-1} is the number of (rank-mapped) sample points in S that belong to $\{1,\ldots,J-1\}^3$, and so we arrive to the same formulas by just replacing K with $J - nt_{J-1}$ and the result follows. $\square$

In this chapter, for $m = 2$ and $m = 3$, we have proved that it is possible to:

- label the different paths an m-variate empirical diagonal may follow by using base-$(m+1)$ number representation,
- count the number of admissible diagonal paths, given a sample size n (Propositions 2.4 and 2.9),
- obtain the exact distribution of the empirical diagonal under the hypothesis of a vector of continuous independent random variables (Theorems 2.5 and 2.10),

with the possibility of obtaining analogous results for higher dimensions, with an increasing difficulty in the combinatorics involved.

Within the Archimedean family of copulas, since copula Π characterizes independence and satisfies Frank's condition $\delta_\Pi{}'(1-) = m$ (and so it is the only Archimedean copula with diagonal section $\delta(u) = u^m$), we may use some appropriate transformation of the empirical diagonal to build a test statistic for a nonparametric test for independence, and since we know the exact distribution of the empirical diagonal under the hypothesis of a bivariate (or trivariate) vector of continuous independent random variables, then we will be able to obtain the exact distribution of such test statistic. This will be done in chapter 4.

Before we proceed with the above idea, a natural question would be if these results would be useful outside the Archimedean world. To be more precise, we restate the question as follows: Does there exist a bivariate copula C such that $\delta_C(u) = \delta_\Pi(u) = u^2$ but $C \neq \Pi$?

The answer is yes, as shown by Fredricks and Nelsen (1997a, 1997b, 2002), see Theorem 1.18 and Proposition 1.19 in this thesis. They built singular and symmetric copulas that have the same diagonal section as, for example, the independence copula, but that are not the independence copula. In practice, we would not worry too much about singular copulas, since in such case we would be able to detect them immediately from a scatter plot. What we would worry about is if it is possible to have absolutely continuous non-Archimedean copulas such that $\delta_C(u) = \delta_\Pi(u) = u^2$ but $C \neq \Pi$. The answer is in the positive sense, as proved by Erdely and González-Barrios (2006a) who have shown how to build a broad family of absolutely continuous copulas, not necessarily symmetric, which have the same diagonal section as the independence copula (see next chapter 3), and there are also results by Nelsen, Quesada-Molina *et al* (2006), and Nelsen (2006b). So it may happen that, outside the Archimedean world, a nonparametric test for independence based on the empirical diagonal may have a low power under alternative hypotheses such as those constructed by Fredricks and Nelsen (1997a, 1997b, 2002), Erdely and González-Barrios (2006a), Nelsen, Quesada-Molina *et al* (2006), and Nelsen (2006b).

Chapter 3

Absolutely continuous copulas with given restrictions

The results of the present chapter have already been published by Erdely and González-Barrios (2006a). We construct a broad family of copulas using a fixed absolutely continuous copula $D(u,v)$. The main idea is to construct families of copulas with given restrictions, such as values of the copula on diagonal sections, or horizontal and vertical sections, even including restrictions on closed subsets, in a few words, to construct copulas with a given *agreement region* with $D(u,v)$.

Let us consider the function

$$f(x,y) = \sin(x)\sin(y) \quad \text{for} \quad x,y \in [\,0,2\pi\,].$$

Then it is clear that if $0 \le a < b \le 2\pi$, we have that

$$\int_a^b \int_0^{2\pi} f(x,y)dxdy = 0 \quad \text{and} \quad \int_0^{2\pi} \int_a^b f(x,y)dxdy = 0,$$

that is, integrals of f along vertical or horizontal segments are always zero. We also observe that $f(x,y) = 0$ on the border of $[\,0,2\pi\,]^2$. We will use appropriate rescalings of the function $f(x,y)$ in order to construct families of absolutely continuous copulas with given diagonals. In fact we will prove the following:

3.1. Theorem. *Let $D(u,v)$ be an absolutely continuous copula with density $\frac{\partial^2}{\partial u \partial v}D(u,v) = d(u,v)$ which will be assumed to be continuous and positive on $[\,0,1\,]^2$. Let $\delta(u) = D(u,u)$ be the diagonal section of $D(u,v)$. Then there exists a family of absolutely continuous copulas $\mathcal{C}$, not necessarily symmetric even when D is symmetric, such that for every $C \in \mathcal{C}$*

if $\delta_C(u) = C(u,u)$*, then* $\delta_C(u) = \delta(u)$*, and for almost every* $(u,v) \in (0,1)^2$ *[*λ*], where* λ *is the Lebesgue measure on* $[0,1]^2$*,* $C(u,v) \neq D(u,v)$*.*

Proof: Let $D(u,v)$ be an absolutely continuous copula with density $d(u,v) = \frac{\partial^2}{\partial u \partial v} D(u,v)$ which is continuous and positive on $[0,1]^2$. Since $d(u,v)$ is continuous and positive on $[0,1]^2$, which is compact, then there exists $M > 0$, such that $d(u,v) \geq M$ for every $(u,v) \in [0,1]^2$.
Let $\Delta_u := \{(u,v) \in [0,1]^2 \,|\, u \leq v\}$ and $\Delta_l := \{(u,v) \in [0,1]^2 \,|\, u \geq v\}$, be the upper and lower triangles above and below the diagonal of $[0,1]^2$. Let $0 \leq u_1 < u_2 \leq v_1 < v_2 \leq 1$, then the rectangle $[u_1,u_2] \times [v_1,v_2] \subset \Delta_u$. Similarly, if $0 \leq v_1 < v_2 \leq u_1 < u_2 \leq 1$, then the rectangle $[u_1,u_2] \times [v_1,v_2] \subset \Delta_l$.
Now we rescale the function f given just before this Theorem to the rectangle $J = [u_1,u_2] \times [v_1,v_2]$. That is, we consider

$$f_J(u,v) = \sin\left(\frac{2\pi(u-u_1)}{u_2-u_1}\right) \sin\left(\frac{2\pi(v-v_1)}{v_2-v_1}\right) 1_J(u,v)\,, \tag{3.1}$$

where 1_A denotes the indicator function of the set A. If $u_1 \leq u \leq u_2$ and $v_1 \leq v \leq v_2$, then

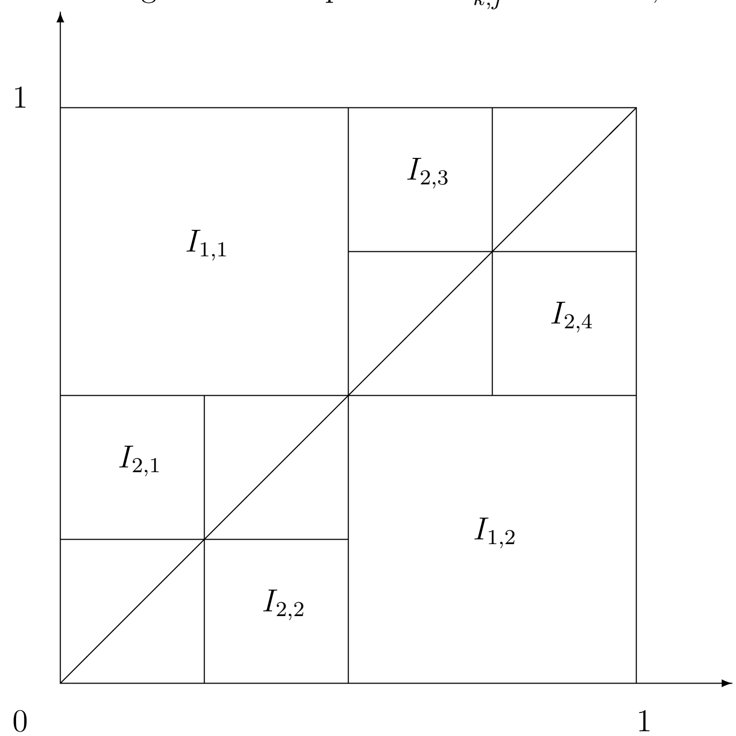
$$\begin{aligned} \int_{v_1}^{v}\int_{u_1}^{u} f_J(s,t)dsdt &= \int_{v_1}^{v}\int_{u_1}^{u} \sin\left(\frac{2\pi(s-u_1)}{u_2-u_1}\right) \sin\left(\frac{2\pi(t-v_1)}{v_2-v_1}\right) dsdt \\ &= \frac{(u_2-u_1)(v_2-v_1)}{4\pi^2} \int_0^{\frac{2\pi(u-u_1)}{u_2-u_1}} \int_0^{\frac{2\pi(v-v_1)}{v_2-v_1}} \sin(w)\sin(z)dwdz \\ &= \frac{(u_2-u_1)(v_2-v_1)}{4\pi^2}\left(1-\cos\left(\frac{2\pi(u-u_1)}{u_2-u_1}\right)\right)\left(1-\cos\left(\frac{2\pi(v-v_1)}{v_2-v_1}\right)\right). \end{aligned}$$

Of course this integral is always nonnegative, and if $u = u_2$ or $v = v_2$, then $\int_{v_1}^{v}\int_{u_1}^{u} f_J(s,t)dsdt = 0$. On the other hand, as can be easily verified, if $u = (u_2+u_1)/2$ and $v = (v_2+v_1)/2$, then $\int_{v_1}^{v}\int_{u_1}^{u} f_J(s,t)dsdt = (u_2-u_1)(v_2-v_1)/\pi^2$, which is the maximum value for this integral.
Let $I_{1,1} = [0,1/2] \times [1/2,1]$ and $I_{1,2} = [1/2,1] \times [0,1/2]$, then $I_{1,1} \subset \Delta_u$ and $I_{1,2} \subset \Delta_l$. For $k \geq 2$ and $j = 1,2,\ldots,2^k$, define

$$I_{k,j} := \begin{cases} \left[\frac{j-1}{2^k}, \frac{j}{2^k}\right] \times \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right] & \text{if} \quad j = 1,3,\ldots,2^k-1 \\ \\ \left[\frac{j-1}{2^k}, \frac{j}{2^k}\right] \times \left[\frac{j-2}{2^k}, \frac{j-1}{2^k}\right] & \text{if} \quad j = 2,4,\ldots,2^k. \end{cases}$$

Then for every $k \geq 2$ and $j = 1,2,\ldots,2^k$, $I_{k,j} \subset \Delta_u$ if j is odd and $I_{k,j} \subset \Delta_l$ if j is even, see Figure 1, for the cases $k=1$ and $k=2$. If we denote the interior of a set A by $\mathrm{int}(A)$, then it is also clear that for any k and every $j_1, j_2 \in \{1,2,\ldots,2^k\}$, with $j_1 \neq j_2$, we have

Figure 3.1: Graph of the $I'_{k,j}s$ for $k = 1, 2$



that $\mathrm{int}(I_{k,j_1}) \cap \mathrm{int}(I_{k,j_2}) = \emptyset$. We also have that for any $k < l$ and any $j_1 \in \{1, 2, \ldots, 2^k\}$ and $j_2 \in \{1, 2, \ldots, 2^l\}$, $\mathrm{int}(I_{k,j_1}) \cap \mathrm{int}(I_{l,j_2}) = \emptyset$. We finally observe that

$$\Delta_u = \bigcup_{k=1}^{\infty} \bigcup_{j=1,3,\ldots,2^k-1} I_{k,j} \quad \text{and} \quad \Delta_l = \bigcup_{k=1}^{\infty} \bigcup_{j=2,4,\ldots,2^k} I_{k,j}\,.$$

Now, for every $k \geq 1$ and every $(u, v) \in [\,0, 1\,]^2$ we define using (3.1)

$$f_{(k;\alpha_{k,1},\ldots,\alpha_{k,2^k})}(u, v) = \sum_{j=1}^{2^k} \alpha_{k,j} f_{I_{k,j}}(u, v) \quad \text{where} \quad |\alpha_{k,j}| \leq M \quad \text{for every} \quad j = 1, 2, \ldots, 2^k.$$

It is important to observe that for any $0 \leq a < b \leq 1$ and any selection of $\alpha_{k,1}, \ldots, \alpha_{k,2^k}$,

$$\int_0^1 \int_a^b f_{(k;\alpha_{k,1},\ldots,\alpha_{k,2^k})}(u, v)dudv = \int_a^b \int_0^1 f_{(k;\alpha_{k,1},\ldots,\alpha_{k,2^k})}(u, v)dudv = 0\,. \qquad (3.2)$$

This property follows from the definition of $f_{(k,\alpha_{k,1},\ldots,\alpha_{k,2^k})}$ and the definition of the subsets $I_{k,j}$, see Figure 3.1. Now for $n \geq 1$ define

$$f_n(u, v) = d(u, v) + \sum_{k=1}^{n} f_{(k;\alpha_{k,1},\ldots,\alpha_{k,2^k})}(u, v),$$

then f_n depends on $\alpha_{1,1}, \alpha_{1,2}, \alpha_{2,1}, \cdots, \alpha_{2,4}, \cdots, \alpha_{n,1}, \cdots \alpha_{n,2^n}$, that is $2 + 2^2 + \cdots + 2^n = 2^{n+1} - 2$ parameters. From the selection of the $\alpha_{k,j}$, for $k = 1, \ldots, n$ and $j = 1, \ldots, 2^k$, we have that every $f_n(u, v)$ is a nonnegative continuous function, since $d(u, v)$ is continuous as well as every $f_{(k,\alpha_{k,1},\ldots,\alpha_{k,2^k})}$. From equation (3.2) we also obtain that

$$\int_0^1 \int_0^1 f_n(u, v) du dv = \int_0^1 \int_0^1 d(u, v) = 1.$$

Hence, for every $n \geq 1$ and selection of parameters $|\alpha_{k,j}| \leq M$, for $k = 1, \ldots, n$ and $j = 1, \ldots 2^k$, $f_n(u, v)$ is a continuous density. In fact, it is the density of an absolutely continuous copula whose formula is given by

$$C_n(u, v) = D(u, v) +$$

$$\sum_{k=1}^{n} \sum_{j=1}^{2^k} \frac{\alpha_{k,j}}{2^{2k+2}\pi^2} \left\{1 - \cos\left(2^{k+1}\pi(u - ((j-1)/2^k))\right)\right\} \left\{1 - \cos\left(2^{k+1}\pi(u - (j^*/2^k))\right)\right\} 1_{I_{k,j}}(u, v),$$

where $j^* = j$ if j is odd and $j^* = j - 2$ if j is even. Here we observe that $C_n(u, v)$ is close to $D(u, v)$ except for perturbations on every $I_{k,j}$. In fact, it is easy to see that

$$\sup_{(u,v) \in [0,1]^2} |C_n(u, v) - D(u, v)| \leq \frac{M}{4\pi^2},$$

for every selection of the parameters $\alpha_{k,j}$. Of course, if every $\alpha_{k,j} \neq 0$, then $C_n(u, v) \neq D(u, v)$ for almost every $(u, v) \in I_{k,j}$, except only if $u = (2j-1)/2^{k+1}$, or if $v = (2j+1)/2^{k+1}$ with j odd, or $v = (2j - 3)/2^{k+1}$ with j even.

If we define $C(u, v) = \lim_{n\to\infty} C_n(u, v)$, we obtain that $C(u, v)$ is a copula depending on an infinite number of parameters, if all of them are non zero, then $C(u, v) \neq D(u, v)$ almost surely for the Lebesgue measure on $[0, 1]^2$. We finally observe that if $\alpha_{k,j} \neq \alpha_{k,j+1}$ for any $k \geq 1$ and any $j = 1, 3, \ldots 2^k - 1$, then $C_n(u, v)$ is a non symmetric copula for any $n \geq k$. In fact, we can construct, using the methodology above, an almost everywhere asymmetric copula with respect to Lebesgue measure. □

3.2. Example. *Let $D(u, v) = uv$ for $(u, v) \in [0, 1]^2$. Then D is an absolutely continuous copula with density $d(u, v) = 1$ for every $(u, v) \in [0, 1]^2$. Let $J = [0, 1/2] \times [1/2, 1]$, if we define*

$$C(u, v) = \begin{cases} uv + \frac{1}{16\pi^2}\{\cos(4\pi u) - 1\}\{\cos(4\pi(v - 1/2)) - 1\} & \text{if} \quad (u, v) \in J \\ uv & \text{if} \quad (u, v) \in [0, 1]^2 \setminus J. \end{cases}$$

Then $C(u, v)$ is an asymmetric copula which coincides with $D(u, v)$ on the diagonal.

We do have easy extensions of Theorem 3.1, such as

3.3. Corollary. *Let $D(u,v)$ be an absolutely continuous copula with density $\frac{\partial^2}{\partial u \partial v} D(u,v) = d(u,v)$ which we will be assumed to be continuous and positive on $[\,0,1\,]^2$. Let $n, m \geq 1$ and let $0 = u_0 < u_1 < u_2 < \cdots < u_{n-1} < u_n = 1$ and $0 = v_0 < v_1 < v_2 < \cdots < v_{m-1} < v_m = 1$ any points. Consider the vertical sections $V_k = \{(u,v) \in [\,0,1\,]^2 \,|\, u = u_k\}$, for $k = 0, \ldots, n$ and the horizontal sections $H_j = \{(u,v) \in [0,1]^2 \,|\, v = v_j\}$ for $j = 0, \ldots, m$. Then there exists an infinite family of copulas $\mathcal{C}$, such that for every $C \in \mathcal{C}$, C is absolutely continuous and $C(u,v) = D(u,v)$ for every $(u,v) \in V_k$, $k = 0, 1, \ldots, n$ and every $(u,v) \in H_j$, $j = 0, 1, \ldots, m$.*

The proof of this Corollary follows the same steps as Theorem 3.1, by defining the density of C on every $I_{k,j} = [\,u_k, u_{k+1}\,] \times [\,v_j, v_{j+1}\,]$, for every $k = 0, 1, \ldots n$ and $j = 0, 1, \ldots, m$.

The function $f(u,v) = \sin(u)\sin(v)$ on $[0, 2\pi]^2$ can be substituted by any function of the form $f(u,v) = g(u)g(v)$ on $[a,b]^2$, as long as $a < b$, g is continuous with $g(a) = g(b) = 0$, and $\int_a^b g(x)dx = 0$.

In the above Corollary the term “infinite family of copulas” is due to the different ways of defining the density of C on every $I_{k,j} = [\,u_k, u_{k+1}\,] \times [\,v_j, v_{j+1}\,]$ and the use of all the functions of the form $f(u,v) = g(u)g(v)$ mentioned above.

Another way of proving Theorem 3.1, is to rescale the function

$$f(u,v) = \sin(2\pi v)\sin(2\pi u/v) 1_{\{(u,v)\in[0,1]^2 \,|\, u \leq v\}}(u,v),$$

and its symmetric version.

We can also find families of copulas that agree on closed sets with a given absolutely continuous copula. For example, copulas that agree with the absolute continuous copula $D(u,v)$ on $[\,1/4, 3/4\,] \times [\,1/4, 3/4\,]$, or even on circles such as $\{(u,v) \in [\,0,1\,]^2 \,|\, (u-1/2)^2 + (v-1/2)^2 \leq 1/4\}$, simply by noticing that the usual topology on $[\,0,1\,]^2$ with the usual metric $d\big(\,(x,y)\,,\,(u,v)\,\big) := \big((x-u)^2 + (y-v)^2\big)^{1/2}$ is the same as the topology metrized by $\rho\big(\,(x,y)\,,\,(u,v)\,\big) := \max\{\,|x-u|, |y-v|\,\}$ where the open balls are (open) rectangles. In fact, from the remark above we have the following

3.4. Corollary. *Let $D(u,v)$ be an absolutely continuous copula with density $\frac{\partial^2}{\partial u \partial v} D(u,v) = d(u,v)$ which we will be assumed to be continuous and positive on $[\,0,1\,]^2$. Let $C(u,v)$ another copula such that $C(u,v) = D(u,v)$ on a largest closed subset $A \subset [0,1]^2$. Then $C(u,v) = D(u,v)$ for every $(u,v) \in [\,0,1\,]^2$ if and only if $A = [\,0,1\,]^2$.*

Proof: Assume that $C(u,v) = D(u,v)$ on a closed subset $A \subset [\,0,1\,]^2$, but $A \neq [\,0,1\,]^2$. Then A^c, the complement of A is an open nonempty set. Hence, we can find $0 < u_1 < u_2 < 1$ and $0 < v_1 < v_2 < 1$, such that $J = [\,u_1, u_2\,] \times [\,v_1, v_2\,] \subset A^c$. By defining f on this rectangle as in Theorem 3.1, we obtain a copula which coincides with $D(u,v)$ on A but it is different from D on J. □

From Corollary 3.4 the only way to determine uniquely an absolutely continuous copula is by giving its values on a dense subset of $[\,0,1\,]^2$.

The hypothesis of a copula having a positive density on $[\,0,1\,]^2$, can be also weakened, obtaining similar results. For example the density can be zero on the border of $[\,0,1\,]^2$, or even the density can be zero on a closed region, and still the construction will work outside this region.

It is important to notice that in all previous results, we can construct asymmetric copulas that follow the given restrictions.

In the present chapter we focused on bivariate copulas but it is possible to extend these results for dimensions higher than two without dealing with the *compatibility problem*, because the starting point for building, say, an n-dimensional asymmetric and absolutely continuous copula $C(u_1,\ldots,u_n)$ would be a given n-dimensional absolutely continuous copula $D(u_1,\ldots,u_n)$ (see Theorem 3.1) and so the proposed methodology does not deal with the compatibility with $(n-m)$-dimensional marginal copulas, where $2 \leq m < n$. For example, with the analogous ideas used in Example 3.2, define

$$C(u,v,w) = \begin{cases} uvw + \frac{1}{64\pi^3}\big[\cos(4\pi u) - 1\big]\big[\cos(4\pi(v-\frac{1}{2})) - 1\big]\big[\cos(4\pi w) - 1\big] & \text{if } (u,v,w) \in J\,, \\ uvw & \text{if } (u,v,w) \in [0,1]^3 \setminus J\,, \end{cases}$$

where $J = [\,0,\frac{1}{2}\,] \times [\,\frac{1}{2},1\,] \times [\,0,\frac{1}{2}\,]$. Then $C(u,v,w)$ is an asymmetric copula which coincides with the independence copula $\Pi(u,v,w) = uvw$ on the diagonal.

We can also find families of n-dimensional copulas that agree on closed sets with a given n-dimensional absolutely continuous copula, even on n-dimensional spheres simply by noticing that the usual topology on $[\,0,1\,]^n$ with the usual metric $d(\,\mathbf{x}\,,\,\mathbf{u}\,) := \big((x_1-u_1)^2+\cdots+(x_n-u_n)^2\big)^{1/2}$ is the same as the topology metrized by $\rho(\,\mathbf{x}\,,\,\mathbf{u}\,) := \max\{\,|x_1-u_1|,\ldots,|x_n-u_n|\,\}$ where the open balls are (open) n-cubes.

The results of this chapter were published by Erdely and González-Barrios (2006a) in April of 2006. During the *10th International Congress on Insurance: Mathematics and Economics* held in Leuven, Belgium, in July of 2006, Quesada-Molina *et al* (2006) presented another way of constructing absolutely continuous non-Archimedean copulas via quasi-copulas, which is part of a paper in preparation by Nelsen, Quesada-Molina *et al* (2006). In August 2006, Nelsen (2006b) kindly sent to the author of this thesis this nice example of a family of absolutely continuous non-Archimedean copulas, non-symmetric when $\theta \neq 0$, with the same diagonal as $\Pi(u, v) = uv$:

$$C_\theta(u, v) \;=\; uv + \theta uv(1-u)(1-v)(u-v)\,, \qquad \theta \in [-1, 1]\,. \tag{3.3}$$

Chapter 4

A new nonparametric test for independence

Alsina, Frank and Schweizer (2003) published a collection of open problems connected with the functional equation of associativity (1.35), and for the present thesis, we got interested in Problem 16, which again appeared as an open problem in Alsina, Frank and Schweizer (2006) :

> Can one design a test of statistical independence based on the assumptions that the copula in question is Archimedean and that its diagonal section is $\delta(u) = u^2$?

The results in this chapter are in Erdely and González-Barrios (2006b) and pretend to give an answer to the above open problem.

4.1 Proposed test statistic.

An immediate consequence of Sklar's theorem (Theorem 1.6) for a random vector (X, Y) of continuous random variables is that the product copula $\Pi(u, v) = uv$ is the copula of (X, Y) if and only if X and Y are independent. The product copula is Archimedean and satisfies Frank's condition $\delta_\Pi{}'(1-) = 2$ so it is characterized by the diagonal section $\delta_\Pi(u) = u^2$. If we are interested in analyzing independence of two random variables, these results suggest to measure some kind of closeness between the empirical diagonal and the diagonal section of the product copula. Moreover, a nonparametric test of independence may be carried out. Let (X, Y) be a random vector of continuous random variables with Archimedean copula C , then the following hypothesis are equivalent:

$$H_0 : X \text{ and } Y \text{ are independent} \quad \Leftrightarrow \quad H_0^* : C = \Pi \quad \Leftrightarrow \quad H_0^{**} : \delta_C(u) = u^2 \,. \tag{4.1}$$

Using the results of chapter 2, we will propose a statistical test based on the empirical diagonal because under H_0 we know the exact distribution of the empirical diagonal (Theorem 2.5) and so we could theoretically obtain the exact distribution of any test statistic based on it.

It is straightforward to verify that under H_0 the expectation $E\,[\,\Delta(\frac{j}{n})\,] = \delta_\Pi(\frac{j}{n}) = \frac{j^2}{n^2}$ so we define

$$\xi\Big(\frac{j}{n}\Big) := \frac{|\Delta(\frac{j}{n}) - \frac{j^2}{n^2}|}{\frac{j}{n} - \max(\frac{2j}{n} - 1, 0)}\,, \quad j = 1, \ldots, n-1\,, \tag{4.2}$$

as a way of measuring pointwise closeness to independence, noticing that the denominator just standardizes dividing by the distance between the Fréchet-Hoeffding bounds at point $\frac{j}{n}$. It is straightforward to verify that $0 \leq \xi(\frac{j}{n}) \leq \max\{\frac{j}{n}\,,\,1 - \frac{j}{n}\} \leq 1 - \frac{1}{n}$. We propose as a test statistic

$$S_n := \frac{1}{n-1}\sum_{j=1}^{n-1} \xi\Big(\frac{j}{n}\Big)\,, \tag{4.3}$$

rejecting H_0 whenever $S_n \geq k_1(\alpha)$, for α a given test size. Before we proceed, from inequalities (1.32) let us denote by $\delta_M(u) = u$ and $\delta_W(u) = \max(2u-1, 0)$ the upper and lower Fréchet-Hoeffding diagonal bounds, respectively. For u in $[\,0, 1\,]$, the average distance between $\delta_\Pi(u)$ and $\delta_M(u)$ is $\frac{1}{6}$, while the average distance between $\delta_\Pi(u)$ and $\delta_W(u)$ is $\frac{1}{12}$, this means that the diagonal that represents independence is, in average, twice closer to the lower than to the upper Fréchet-Hoeffding diagonal bound, thus independence is far from being in the middle of such bounds, and so we should consider the possibility of taking this into account in defining a test statistic. We define

$$h\Big(\frac{j}{n}\Big) := \frac{\frac{j}{n} - \frac{j^2}{n^2}}{\frac{j^2}{n^2} - \max(\frac{2j}{n} - 1, 0)} = \begin{cases} \frac{n-j}{j} & \text{if } j \leq \frac{n}{2}\,, \\ \frac{j}{n-j} & \text{if } j > \frac{n}{2}\,, \end{cases} \tag{4.4}$$

as a factor to be multiplied by $\xi(\frac{j}{n})$ for those observations for which $\Delta(\frac{j}{n}) < \frac{j^2}{n^2}$ in order to compensate somehow the non-equal closeness of the independence diagonal to the Fréchet-Hoeffding bounds. In other words, let us define

$$\nu\Big(\frac{j}{n}\Big) := \begin{cases} h\Big(\frac{j}{n}\Big)\xi\Big(\frac{j}{n}\Big) & \text{if } \Delta\Big(\frac{j}{n}\Big) < \frac{j^2}{n^2}\,, \\ \xi\Big(\frac{j}{n}\Big) & \text{if } \Delta\Big(\frac{j}{n}\Big) \geq \frac{j^2}{n^2}\,, \end{cases} \tag{4.5}$$

It is straightforward to verify that $h(\frac{j}{n})$ is symmetric respect to $\frac{1}{2}$ and that $1 \leq h(\frac{j}{n}) \leq h(\frac{1}{n}) = h(1 - \frac{1}{n}) = n - 1$. We now propose the following test statistic

$$A_n := \frac{1}{n-1}\sum_{j=1}^{n-1} \nu\Big(\frac{j}{n}\Big)\,, \tag{4.6}$$

rejecting H_0 whenever $A_n \geq k_2(\alpha)$, for α a given test size. As we will see in the following section, the test statistics (4.3) and (4.6) alone sometimes lead to biased tests of independence, but an appropriate combination of both leads to an approximately unbiased independence test, under an Archimedean copula, by rejecting H_0 whenever $S_n \geq k_1$ or $A_n \geq k_2$ with k_1 and k_2 chosen appropriately such that

$$P\big(\{S_n \geq k_1\} \cup \{A_n \geq k_2\} \,|\, H_0\big) \leq \alpha\,. \tag{4.7}$$

It is important to make some remarks regarding S_n and A_n. From their definitions it is immediate to verify that $S_n \leq A_n$, and that these two statistics are bounded: $0 < S_n \leq A_n \leq \frac{3}{4} - \frac{1}{4n}$. Moreover, A_n and S_n are discrete random variables whose exact distribution under H_0 may be obtained due to the fact that these statistics are defined in terms of the empirical diagonal whose exact distribution under H_0 has been obtained in chapter 2. For this reason, we have to use less or equal instead of equality in (4.7), but with the understanding that we choose k_1 and k_2 as small as possible such that $P\big(\{S_n \geq k_1\}\cup\{A_n \geq k_2\} \,|\, H_0\big)$ gets as close as possible to α, that is $P\big(\{S_n \geq k_1\} \cup \{A_n \geq k_2\} \,|\, H_0\big) \approx \alpha$.

If we analyze a test with the following rejection rule:

$$\text{reject } H_0 \text{ whenever } S_n \geq k_1 \text{ or } A_n \geq k_2 \tag{4.8}$$

for appropriately chosen k_1 and k_2 accordingly to a desired test size α, see (4.7), we may choose, say, k_1 such that $P(S_n \geq k_1 \,|\, H_0) = 0$ and in this particular case we will have that $P(\{S_n \geq k_1\} \cup \{A_n \geq k_2\} \,|\, H_0) = P(A_n \geq k_2 \,|\, H_0) \leq \alpha$, and we get a test based only in A_n, and the same for S_n. In any other case we get a test based on both statistics.

By elementary probability we rewrite (4.7) as

$$\begin{aligned} \alpha &\geq P(S_n \geq k_1 \,|\, H_0) + P(A_n \geq k_2 \,|\, H_0) - P(S_n \geq k_1\,,\, A_n \geq k_2 \,|\, H_0) \\ &\geq \alpha_1 + \alpha_2 - \alpha_{12}\,. \end{aligned} \tag{4.9}$$

So we may first find k_1 such that $P(S_n \geq k_1 \,|\, H_0) \leq \alpha_1 \leq \alpha$ and then we search for k_2 such that $\alpha_1 + \alpha_2 - \alpha_{12} \approx \alpha$ without exceeding α. The election of (k_1, k_2) is not unique but we can compute all the different possibilities since we are able to obtain the exact distribution under H_0 of the discrete random vector (S_n, A_n) using Theorem 2.5.

For example, setting $\alpha = 0.05$, with sample size $n = 15$ we get 54 different possibilities displayed in Figure 4.1: the first graph plots the different pairs of values (k_1, k_2) such that $P(\{S_{15} \geq k_1\} \cup \{A_{15} \geq k_2\} \,|\, H_0) \approx 0.05$ without exceeding 0.05. The second graph plots the exact value of $P(\{S_{15} \geq k_1\} \cup \{A_{15} \geq k_2\} \,|\, H_0)$ for each pair (k_1, k_2), from these values we obtain that $0.04968 \leq P(\{S_{15} \geq k_1\} \cup \{A_{15} \geq k_2\} \,|\, H_0) \leq 0.05$. The third graph plots the marginal probabilities $P(S_n \geq k_1 \,|\, H_0)$ and $P(A_n \geq k_2 \,|\, H_0)$ for each (k_1, k_2), and we identify the particular case $(k_1, k_2) = (0.27210884, 0.61717687)$ where $P(S_n \geq k_1 \,|\, H_0)$ and $P(A_n \geq k_2 \,|\, H_0)$ are very much alike, with probabilities 0.02500343 and 0.02527600, respectively.

Even though the election of (k_1, k_2) is not unique, we will see in the following section that, in order to obtain an approximately unbiased test, a good choice for the alternatives we analyze is (k_1, k_2) such that $\alpha_1 = P(S_n \geq k_1 \,|\, H_0) \approx P(A_n \geq k_2 \,|\, H_0) = \alpha_2$. We cannot prove this in general for all possible alternatives since the power of the test for $\theta \neq \theta_0$ depends on the distribution under the alternative hypothesis, but it seems to work adequately in all the following simulations.

4.2 Simulation study.

To check the performance of the independence test based on S_n and/or A_n, we performed a simulation study, and we made a comparison to some well-known independence tests based on the following statistics:

- Spearman's Rank correlation coefficient, see Lehmann (1975) :

$$R = \frac{12}{n(n^2-1)} \sum_{j=1}^{n} \text{rank}(X_j)\text{rank}(Y_j) - \frac{3(n+1)}{n-1}\,, \tag{4.10}$$

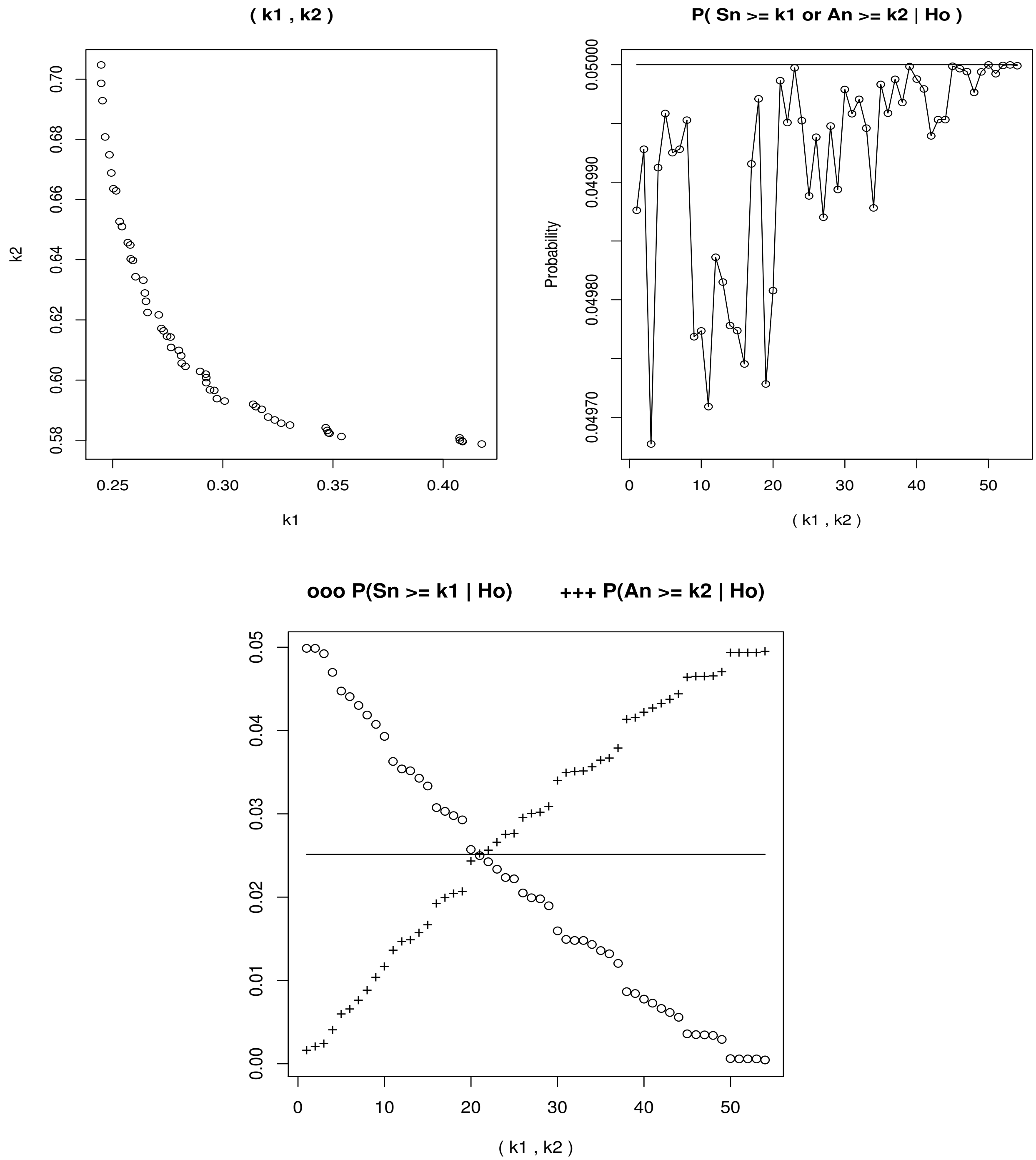
  rejecting H_0 for large values of $|R|$.

- the modified Hoeffding test as introduced by Blum, Kiefer, and Rosenblatt (1961):

$$B = \int\!\!\int \big[\, H_n(x,y) - F_n(x)G_n(y) \,\big]^2 \, dH_n(x,y)\,, \tag{4.11}$$

  where $H_n(x,y) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbf{1}_{\{\, X_k \leq x\,,\, Y_k \leq y \,\}}$, $F_n(x) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbf{1}_{\{\, X_k \leq x \,\}}$, $G_n(y) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbf{1}_{\{\, Y_k \leq y \,\}}$, rejecting H_0 for large values of B.

Figure 4.1:


(k1 , k2)
k2
k1
0.70
0.68
0.66
0.64
0.62
0.60
0.58
0.25
0.30
0.35
0.40
P(Sn >= k1 or An >= k2 | Ho)
Probability
0.05000
0.04990
0.04980
0.04970
0
10
20
30
40
50
(k1 , k2)
ooo P(Sn >= k1 | Ho) +++ P(An >= k2 | Ho)
0.05
0.04
0.03
0.02
0.01
0.00
0
10
20
30
40
50
(k1 , k2)

- Kallenberg and Ledwina (1999) test statistic:

$$V = \begin{cases} V(1,1) \quad \text{if } \max\{V(1,2), V(2,1), V(2,2)\} < \log n\,, \\ V(1,1) + \max\{V(1,2), V(2,1), V(2,2)\} \quad \text{otherwise,} \end{cases} \tag{4.12}$$

  with $V(r,s) = \left\{ \dfrac{1}{\sqrt{n}} \displaystyle\sum_{i=1}^{n} b_r\left(\frac{R_i - \frac{1}{2}}{n}\right) b_s\left(\frac{S_i - \frac{1}{2}}{n}\right) \right\}^2$, where $\{b_j\}$ denote the orthonormal Legendre polynomials on $[0,1]$, R_i and S_i are the ranks of X_i and Y_i, and we reject H_0 for large values of V.

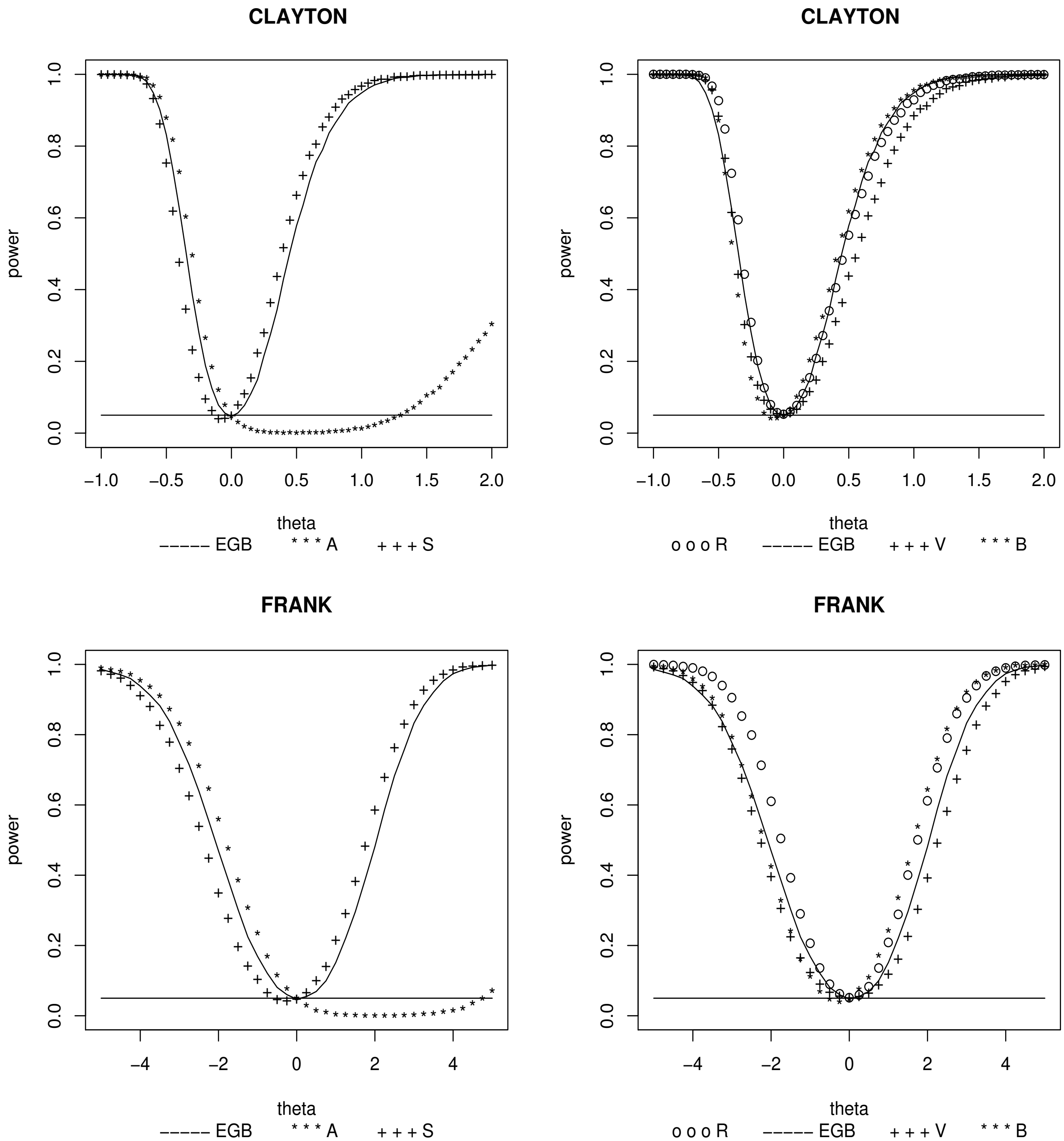
The simulated power comparisons presented here were obtained with $n = 50$ and $\alpha = 0.05$. Every Monte Carlo experiment reported here has been simulated 10,000 times, using some one-parameter Archimedean copulas as alternatives, divided into two groups. Let C_θ be an Archimedean copula with (one-dimensional) parameter θ :

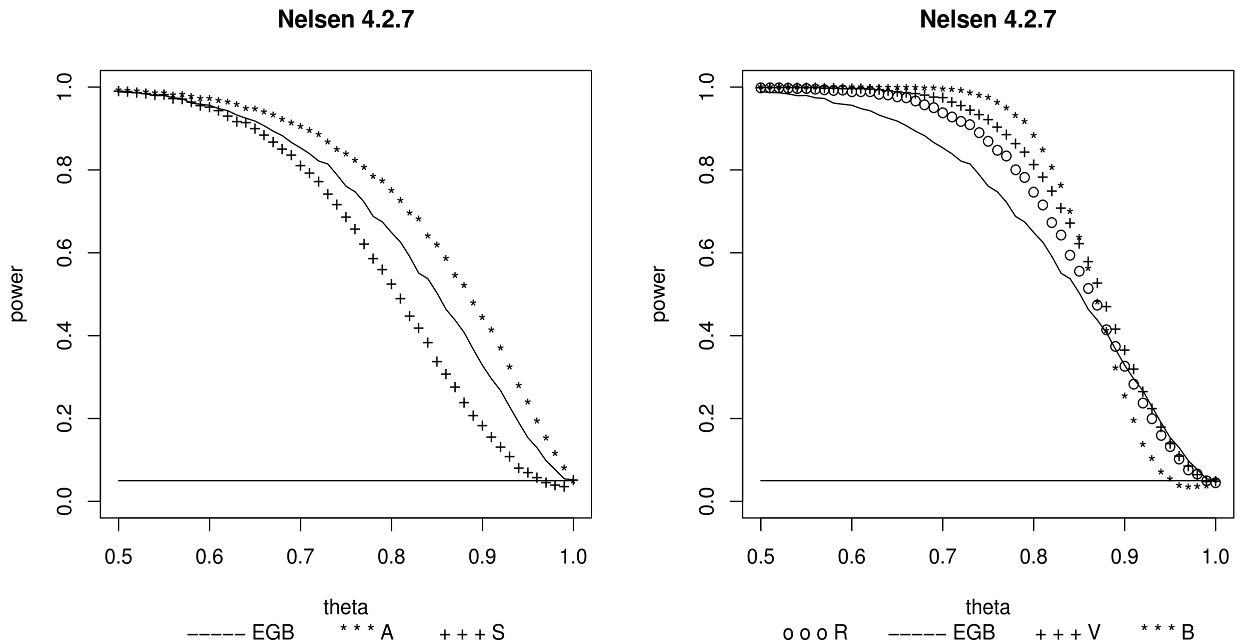
Group 1: There exists a unique θ_0 such that $C_{\theta_0} = \Pi$.

Group 2: $C_\theta \neq \Pi$ for all θ.

Let $\Psi(\theta)$ denote the power of the test. For Group 1, hypothesis (4.1) becomes $H_0 : \theta = \theta_0$ versus the alternative $H_1 : \theta \neq \theta_0$ and so an unbiased test should satisfy $\inf_{\theta \neq \theta_0} \Psi(\theta) \geq \Psi(\theta_0) = \alpha$. For Group 2 we expect $\Psi(\theta)$ to be as close as possible to 1 for all θ. We will denote by EGB the test proposed by Erdely and González-Barrios (2006b) in (4.8), by A the test based on the statistic A_n (4.6), and by S the test based on the statistic S_n (4.3). For each copula we present two graphs: one comparing the power of EGB, A and S; a second one comparing the power of EGB versus R, B and V, as defined in (4.10), (4.11), and (4.12).

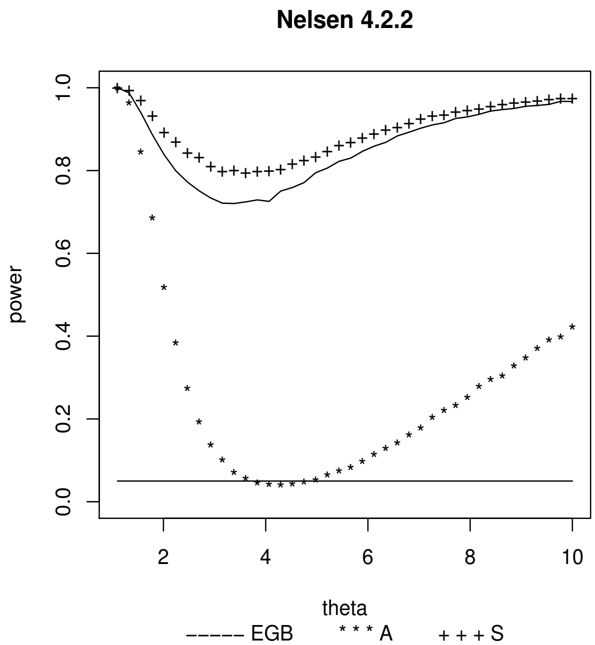
Group 1. We compare the test powers for $H_0 : \theta = 0$ against $H_1 : \theta \neq 0$ with the following alternative Archimedean copulas, for details see Nelsen (2006a): Clayton, Frank, and Nelsen's catalog number 4.2.7. In all cases these copulas satisfy $C_\theta = \Pi$ if and only if $\theta = 0$, or $\lim_{\theta \to 0} C_\theta = \Pi$, and satisfy Frank's condition $\delta'(1-) = 2$.

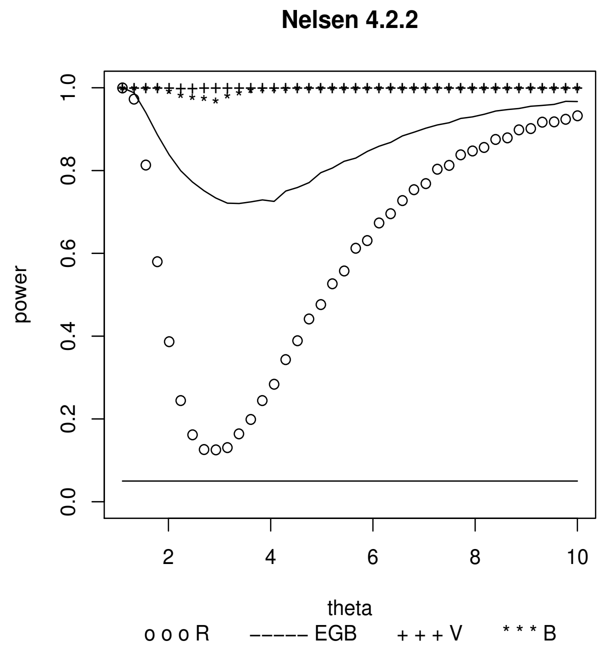

CLAYTON
power
theta
----- EGB
* * * A
+ + + S
CLAYTON
power
theta
o o o R
----- EGB
+ + + V
* * * B
FRANK
power
theta
----- EGB
* * * A
+ + + S
FRANK
power
theta
o o o R
----- EGB
+ + + V
* * * B

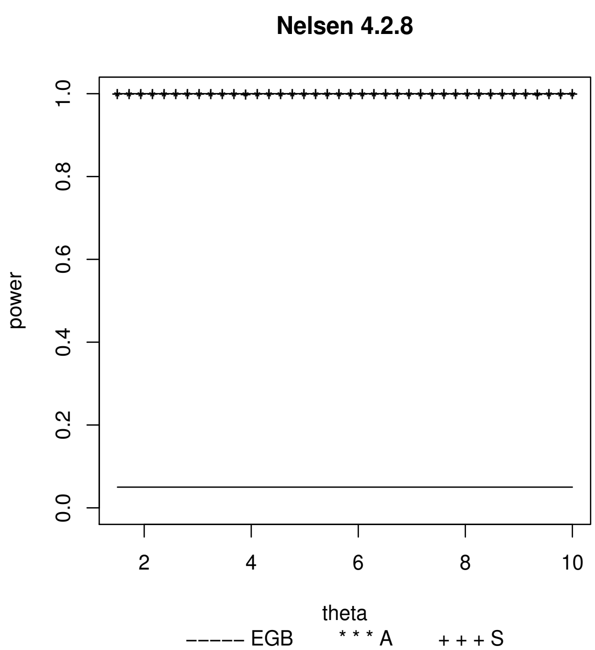


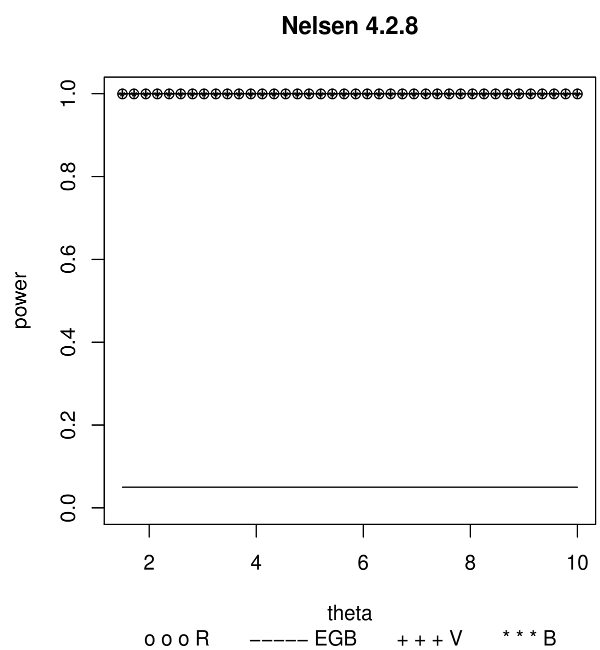
We notice that in most cases, A and S are biased. Test A has a good performance if $\theta \leq 0$ but a poor performance if $\theta \geq 0$, while the test based on S alone has a good performance if $\theta \geq 0$ and a not so good performance whenever $\theta \leq 0$. Test EGB is an attempt to combine the good part of tests A and S. Comparing tests EGB, R, V and B we notice that none of them is uniformly better than the others. At the end of this section we have a summary table of comparisons. In the case of the Clayton copula, EGB and R seem to have a better relative performance than V and B. Under Frank copula, R performs roughly better than the others. In the case of copula Nelsen 4.2.7 we notice the crossings of powers, so that, for example, EGB is good for values of θ close to 1 or less than 0.5, but not so good for other values.

Group 2. We compare the test powers for $H_0 : C_\theta = \Pi$ where $C_\theta \neq \Pi$ for all θ so in this case we expect the test powers to be as close a possible to 1 for all θ . The analyzed Archimedean copulas were the following, for details see Nelsen(2006a): Nelsen's catalog numbers 4.2.2 and 4.2.8. Copula 4.2.8 satisfies Frank's condition, but copula 4.2.2 does not, and for this case and under the EGB test what we obtained should be considered as a rough estimation of the power, since there may be other Archimedean copulas with the same diagonal.





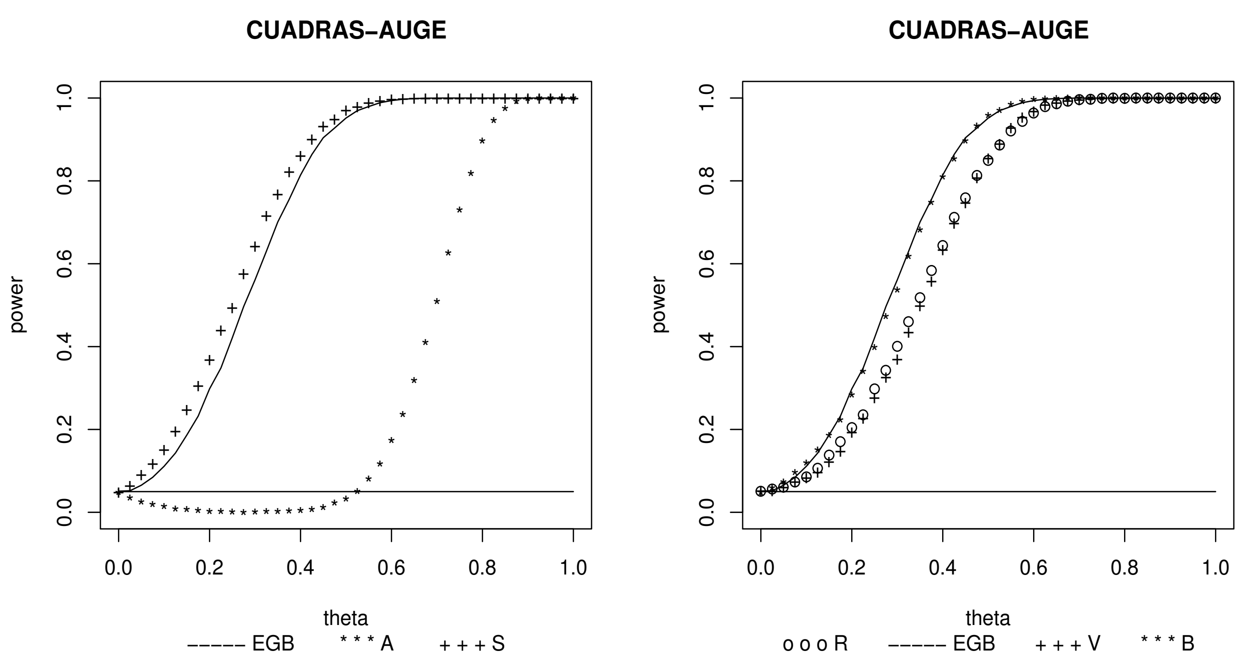


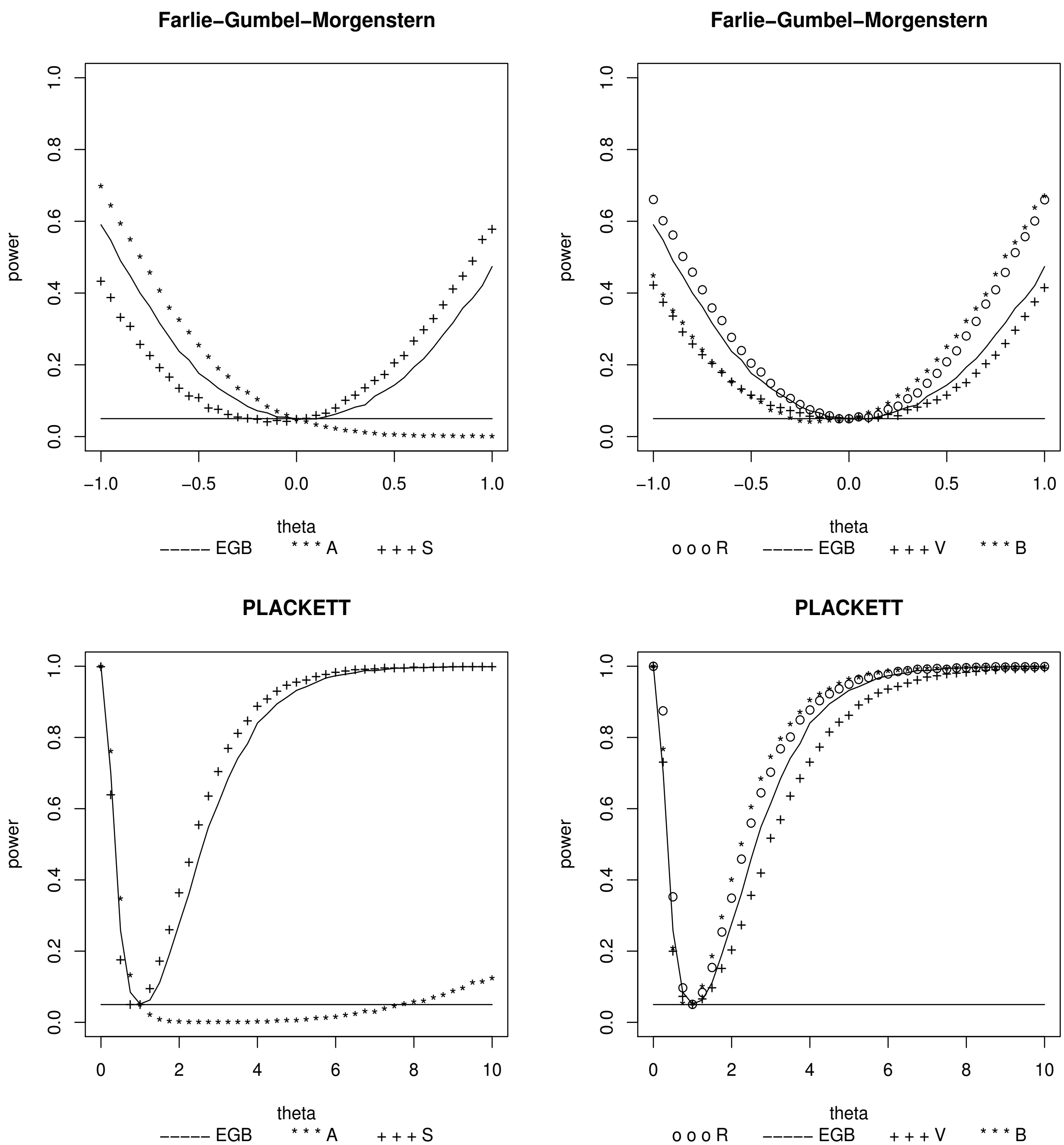


In the case of copula Nelsen 4.2.2 tests V and B seem to be uniformly better than EGB and R, but test R might be quite misleading for certain values of θ. The dependence structure of copula Nelsen 4.2.8 is very well detected by all tests.

So far, we are taking advantage of the fact that whenever the underlying copula is of the Archimedean type and satisfies Frank's condition, all the information is contained in the diagonal section, and we ended with an independence test under this assumption. Of course an obvious question is what happens with the proposed EGB test outside the Archimedean world. As proved by Fredricks and Nelsen (1997a, 1997b, 2002) it is possible to build copulas different from the product (or independence) copula $\Pi(u,v) = uv$ but with the same diagonal as Π, but they are singular, so in practice, even though the proposed EGB test would not detect them, because of its singular nature it would be quite easy to detect them by simple inspection of a scatter plot, besides the fact that singular copulas rarely appear in real problems. What is really an issue for the proposed EGB test is the fact that there are absolutely continuous non-Archimedean copulas which have the same diagonal as Π, as proved by Erdely and González-Barrios (2006a), see chapter 3, and also proved by Nelsen, Quesada-Molina *et al* (2006), and Nelsen (2006b), see (3.3), so outside the Archimedean world the proposed EGB test may face dependence structures that it will not be able to detect. Anyway, we performed similar simulation studies for some well-known non-Archimedean families of copulas, with surprising results:

Group 3. We compare the test powers for $H_0 : \theta = 0$ against $H_1 : \theta \neq 0$ with the following alternative non-Archimedean copulas: Cuadras-Augé, Farlie-Gumbel-Morgenstern, and Plackett. In all cases these copulas satisfy $C_\theta = \Pi$ if and only if $\theta = 0$, or $\lim_{\theta \to 0} C_\theta = \Pi$. For details see Nelsen (2006a).





What we notice in the case of these three non-Archimedean copulas is that, even though they are not characterized by their diagonal section, if they are taken as alternative hypotheses the proposed EGB still has an interesting performance. This suggests that even outside the Archimedean families, the diagonal section has valuable information (yet not all) about some non-Archimedean families of copulas.

Table 4.1:

Group number	Alternative Copula	*EGB*	*R*	*V*	*B*
1	Clayton	27	24	44	56
1	Frank	42	24	52	50
1	Nelsen 4.2.7	28	22	15	70
2	Nelsen 4.2.2	28	87	0	3
2	Nelsen 4.2.8	0	0	0	0
3	Cuadras-Augé	12	32	37	8
3	Farlie-Gumbel-Morgenstern	44	25	53	51
3	Plackett	40	18	49	43

We made a summary of the above graphs in the format suggested by Kallenberg and Ledwina (1999): "for each test statistic, we have calculated the difference between the power of the test and the maximal power of the tests under consideration at the given alternative. For each graph this difference is maximized over the alternatives in the graph. This number can be seen as a summary for the behavior of the test in that graph, although of course some information of the graph is lost." In Table 4.1 we present percentage differences in maximal power of the four tests under comparison and the power of the given test at various alternatives, so the lower the difference the better is the relative performance of the test.

In the above comparison table we may notice, for example, that the R test is the best choice under certain copulas, but the worst, by far, under copula Nelsen 4.2.2. Analogous comment applies for the B test comparing its performance under copulas Nelsen 4.2.2 and Nelsen 4.2.7. In practice, when using a nonparametric test for independence we usually do not know what alternative we are dealing with, so what is valuable about a nonparametric test is its ability to maintain a good performance under different alternatives, rather than being the best one under specific ones. Even though the proposed EGB test does not appear to be the best one under a particular copula in the above comparison, it never breaks down as bad the R test in case of copula Nelsen 4.2.2 or as bad as the B test in case of copula Nelsen 4.2.7, we may say that the EGB test has some kind of "stability," at least under the above set of copulas.

4.3 Comparison against the locally most powerful rank test.

The results of the previous two sections were presented at the *10th International Congress on Insurance: Mathematics and Economics* held in Leuven, Belgium, in July of 2006, and in that occasion Dr. Christian Genest suggested to make a comparison of the proposed test against the locally most powerful test, as in, for example, Genest (2005), using a result from Garralda-Guillén (1997):

> **Proposition (Garralda-Guillén).** Let $(R_1, S_1), \ldots, (R_n, S_n)$ denote the ranks associated with a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from a distribution whose underlying copula belongs to a class (C_θ) satisfying the following conditions:
>
> (i) the parameter space Θ is a closed interval and there exists $\theta_0 \in \Theta$ such that $C_{\theta_0} = \Pi$ is the copula corresponding to independence;
>
> (ii) the family (C_θ) is ordered by positive quadrant dependence, that is, the implication $\theta < \theta' \Rightarrow C_\theta(u,v) \leq C_{\theta'}(u,v)$ is valid for all $u, v \in (0,1)$;
>
> (iii) for every $\theta \in \Theta$, C_θ is absolutely continuous and its associated density $c_\theta(u,v)$ is absolutely continuous as a function of θ for every $u, v \in (0,1)$;
>
> (iv) $\dot{c}_\theta\,(u,v) = \partial c_\theta(u,v)/\partial\theta$ is continuous in θ in a neighborhood of θ_0, and
>
> $$\lim_{\theta \to \theta_0} \int_{(0,1)^2} |\dot{c}_\theta\,(u,v)|\, dvdu = \int_{(0,1)^2} |\dot{c}_{\theta_0}\,(u,v)|\, dvdu < \infty\,.$$
>
> The locally most powerful rank test of level α rejects $\mathcal{H}_0 : \theta = \theta_0$ of independence against $\mathcal{H}_1 : \theta > \theta_0$ for large enough values of
>
> $$T_n^* \;=\; \frac{1}{n}\sum_{i=1}^{n} T(R_i, S_i)\,, \tag{4.13}$$
>
> where
>
> $$T(r,s) \;=\; \mathbb{E}\left\{\frac{\partial}{\partial\theta}\log\, c_\theta(B_r, B_s^*)\big|_{\theta=\theta_0}\right\} \tag{4.14}$$
>
> with B_r and B_s^* representing two independent random variables respectively distributed as Beta$(r, n-r+1)$ and Beta$(s, n-s+1)$.

We may not apply the above result to copulas Nelsen 4.2.7 and Cuadras-Augé since they are not absolutely continuous, neither to copulas Nelsen 4.2.2 and Nelsen 4.2.8 since these two copulas are such that $C_\theta \neq \Pi$ for all $\theta \in \Theta$. But as shown in Genest (2005) or

Garralda-Guillén (1997) the above proposition does apply to the following families of copulas: Clayton, Frank, Farlie-Gumbel-Morgenstern, and Plackett.

It is important to notice that Garralda-Guillén's Proposition provides the locally most powerful *one-sided* test

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{versus} \quad \mathcal{H}_1 : \theta > \theta_0 \,, \tag{4.15}$$

in contrast with the *two-sided* test $\mathcal{H}_0 : \theta = \theta_0$ versus $\mathcal{H}_1 : \theta \neq \theta_0$ analyzed in the previous section (Group 1). Since Garralda-Guillén's Proposition requires the family of copulas used as alternative to be positively ordered, and since our proposed test EGB uses the bivariate statistic (S_n, A_n), see (4.8), where S_n has a better performance alone in comparison with (S_n, A_n) when $\theta \geq \theta_0$, the comparison to the locally most powerful rank test has been done against the following rejection rule:

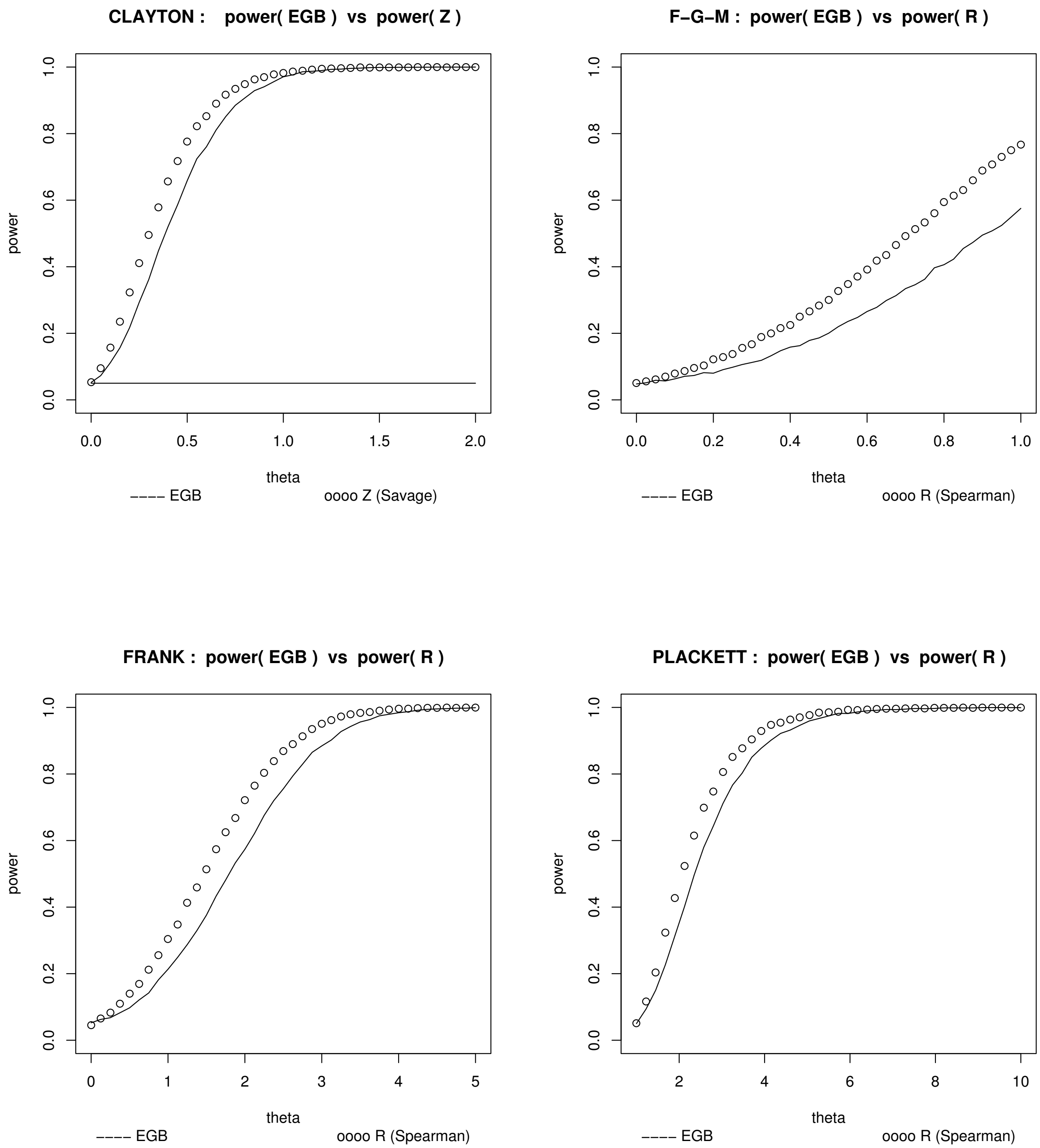
$$\text{reject } H_0 \text{ whenever } S_n \geq k_\alpha \tag{4.16}$$

for k_α appropriately chosen accordingly to a desired test size α , that is, such that $\mathbb{P}(S_n \geq k_\alpha \,|\, \mathcal{H}_0) \leq \alpha$, with S_n defined as in (4.3).

As shown in Garralda-Guillén (1997), or Genest (2005), the locally most powerful rank test for (4.15) is Spearman's, see (4.10), in case the family of copulas used as alternative is Frank, Farlie-Gumbel-Morgenstern, or Plackett; and in the case of the Clayton family of copulas used as alternative in (4.15), the locally most powerful rank test is the one based on the *Savage scores statistic:*

$$Z_n := \frac{1}{n} \sum_{i=1}^{n} \Bigg(\sum_{j=R_i}^{n} \frac{1}{j} \Bigg) \Bigg(\sum_{k=S_i}^{n} \frac{1}{k} \Bigg) - 1 \,. \tag{4.17}$$

The simulated power comparisons presented here were obtained with $n = 50$ and $\alpha = 0.05$. Every Monte Carlo experiment reported here has been simulated 10,000 times. We use the notation EGB to denote the proposed test in (4.16), Z for the test based on Savage's statistic (4.17), and R for Spearman's test (4.10):



CLAYTON : power(EGB) vs power(Z)
power
theta
---- EGB
oooo Z (Savage)
F−G−M : power(EGB) vs power(R)
power
theta
---- EGB
oooo R (Spearman)
FRANK : power(EGB) vs power(R)
power
theta
---- EGB
oooo R (Spearman)
PLACKETT : power(EGB) vs power(R)
power
theta
---- EGB
oooo R (Spearman)

In practice, of course, we usually don't know under which family of alternatives we are, so the above comparisons are in a certain sense "unfair" since many locally most powerful rank tests perform not so good under other alternatives, as shown, for example, by Genest *et al* (2004); we may just consider them as reference benchmarks. From the previous graphs, we observe that for the alternatives Clayton, Frank, and Plackett, the performance of our test is competitive against the locally most powerful rank test.

4.4 A dependence structure hard to detect.

A theorem by Mikusiński *et al* (1991), see Theorem 1.50 in chapter 1, implies that in practice, the behavior of any pair of independent continuous random variables can be approximated so closely by a pair of MCD (mutually completely dependent) continuous random variables that it would be impossible, experimentally, to distinguish one pair from the other, and this is done by constructing a copula C_ε such that for any given copula C and for all $\varepsilon > 0$

$$\sup_{u,v\,\in\,[\,0,1\,]} |\, C(u,v) - C_\varepsilon(u,v)\,| < \varepsilon\,.$$

Therefore, we may simulate observations from a copula C_ε , and for ε sufficiently small, we may defeat any nonparametric test for independence. Such copula C_ε is a *Shuffle of Min,* see Mikusiński *et al* (1991, 1992) or subsection 1.5.2 in this thesis, which is a copula of the singular type. A pragmatic statistician may argue that this interesting theoretical example is not of practical interest because of the singularity. So we will then proceed to analyze an example provided by Nelsen (2006b), see (3.3) at the end of chapter 3 :

$$C_\theta(u,v) \;=\; uv + \theta uv(1-u)(1-v)(u-v)\,, \qquad \theta \in \,[-1,1\,]\,, \tag{4.18}$$

which happens to be an absolutely continuous non-Archimedean copula, asymmetric when $\theta \neq 0$, with the same diagonal as the independence copula, but $C_\theta \neq \Pi$ whenever $\theta \neq 0$.

We should first notice that this parametric family of copulas is not ordered, neither positively or negatively, and so we may not apply Garralda-Guillén's proposition (see previous section) in order to obtain a locally most power rank test for independence. We will just then compare the power of the proposed test EGB, see (4.8), versus Spearman's, Blum-Kiefer-Rosenblatt's, and Kallenberg-Ledwina's (see section 4.2).

When $\theta = 0$, from (4.18) we have that $C_0 = \Pi$, and therefore we will analyze the performance of the above mentioned tests for

$$\mathcal{H}_0 : \theta = 0 \qquad \text{versus} \qquad \mathcal{H}_1 : \theta \neq 0. \tag{4.19}$$

Since the diagonal section of (4.18) is the same as Π, $\delta_\theta(u) = u^2$ for all $\theta \in [-1,1]$, there is no hope for our proposed test EGB to reject $\mathcal{H}_0 : \theta = 0$ (independence), even for the extreme values $\theta = -1$ or $\theta = 1$. In the case of Spearman's test, by using (1.89) we get that Spearman's rho $\rho_\theta = 0$ for all $\theta \in [-1,1]$, and there is no hope for this test, neither. This is an example to illustrate that concordance measures do not characterize independence. If instead we use a *dependence measure* as defined in subsection 1.5.2, such as, for example, the one proposed by Schweizer and Wolff (1981), see (1.97) :

$$\sigma_\theta = 12 \iint_{[0,1]^2} |\, C_\theta(u,v) - uv \,|\, dudv \,, \tag{4.20}$$

where $0 \leq \sigma_\theta \leq 1$, for the particular case of (4.18), we find that

$$\sigma_\theta \;=\; \frac{6}{70}\,|\,\theta\,| \;\leq\; \frac{6}{70} \;\approx\; 0.08571429\,,$$

and since we may think of (4.20) as a standardized average absolute distance between copula C_θ and Π, we have that even for the extreme values of the parameter θ both copulas are too close, and so we may have no big expectations on the performance of the tests by Blum-Kiefer-Rosenblatt and Kallenberg-Ledwina.

The simulated power comparisons presented here were obtained with $n = 50$ and $\alpha = 0.05$. Every Monte Carlo experiment reported here has been simulated 10,000 times, and since copulas C_θ and Π are too close even for the extreme values of the parameter θ, we just summarize the power of the tests for $\theta = -1, 0, 1$:

Table 4.2:

	Power		
Test	$\theta = -1$	$\theta = 0$	$\theta = 1$
proposed EGB	0.0558	0.0506	0.0548
Spearman	0.0518	0.0500	0.0502
Blum-Kiefer-Rosenblatt	0.0667	0.0497	0.0709
Kallenberg-Ledwina	0.1210	0.0519	0.1170

From the above table we have that the four tests' performance ranges from poorly to useless, since in the best case, under Kallenberg-Ledwina's, we get values for the power of the test under $\theta = -1$ or $\theta = 1$ very far away from 1, when for a good test it would be required power values as close as possible to 1.

Conclusions and open problems

In Chapter 1 we made an extensive review of known results about copulas, and we included a special section for *Schröder's Functional Equation,* which is a key result for the result announced by Frank (1996), where bivariate Archimedean copulas are characterized by their diagonal whenever *Frank's condition* $\delta\,'(1-) = 2$ is satisfied. In Alsina, Frank and Schweizer (2006) a counterexample is given to illustrate that if Frank's condition is not satisfied, it is possible to build a parametric family of generators of Archimedean copulas $\{\varphi_\beta : 0 < \beta \leq 1/(1+8\pi)\}$ such that their diagonal section $\delta_\beta = \delta_C$ but $C_{\beta_1} \neq C_{\beta_2}$ for $\beta_1 \neq \beta_2$. We proved that the upper bound given for β in this example is not sharp, and we obtained a sharp one, see section 1.4.

In Chapter 2 we obtained the exact distribution of the empirical diagonal under the hypothesis of independence, for two and three dimensions. It is possible to obtain analogous results for higher dimensions. There is also a computational problem to be tackled, since for the three-dimensional case we have a problem of order $(n!)^2$, and so for large values of n it takes too long for a computer to obtain the exact distribution, unless a very efficient algorithm is developed. Further research may be done on the possibility of using the empirical diagonal for nonparametric estimation of Archimedean copulas, by finding an appropriate way of smoothing the empirical diagonal in such a way that the limit (1.62) exists and equals a function that fulfills the properties of an Archimedean copula generator.

In Chapter 3 we investigated the possibility of building absolutely continuous copulas with the same diagonal as the independence copula, but different from it outside the diagonal, and we arrived to a more general result about a broad family of absolutely continuous copulas using a fixed absolutely continuous copula $D(u, v)$. We proved that it is possible to construct families of copulas with given restrictions, such as values of the copula on diagonal sections, or horizontal and vertical sections, even including restrictions on closed subsets, in a few words, to construct copulas with a given *agreement region* with $D(u, v)$.

In Chapter 4 we solved the open problem 16 in Alsina, Frank and Schweizer (2003), designing a test of statistical independence based on the assumptions that the copula in question is Archimedean and that its diagonal section is $\delta(u) = u^2$. Moreover, the proposed test seems to be competitive when comparing its power against some other nonparametric tests for independence, under different Archimedean alternatives. An important feature of the proposed test is the fact that it is based on a statistic whose exact distribution is known, in contrast with many other tests based on the asymptotic behavior of their test statistics. The results from Chapter 3 set a warning for the use of the proposed test outside the Archimedean family. Surprisingly, under some well-known non-Archimedean families of copulas, the proposed test has an interesting performance. This suggests that even outside the Archimedean family, the diagonal section has valuable information (yet not all) about some non-Archimedean families of copulas, and this may deserve further research in order to determine under which conditions the diagonal is still quite informative, since we presented an example in section 4.4 where knowledge of the diagonal is useless in testing for independence. In any case, but specially within the Archimedean family, we may investigate the possibility of obtaining point estimators of the parameter(s) of a copula, based on the empirical diagonal.

From Alsina, Frank and Schweizer (2003), it remains unsolved open problem 15:

> Are there any statistical properties of two random variables which assure that their copula is Archimedean or, more generally, associative?

Of course that the answer to this question could possibly lead to design a nonparametric test for Archimedeaness. As we did in the case of the empirical diagonal, we may try to solve this problem by studying combinatorial properties of the empirical copula that may distinguish between Archimedean and non-Archimedean copulas, or more generally, between associative and non-associative copulas.

Bibliography

Aczél, J. (1966) *Lectures on functional equations and their applications*, Academic Press (New York).

Alsina, C., Frank, M.J., Schweizer, B. (2003) Problems on associative functions. *Aequationes Math.* **66**, 128-140.

Alsina, C., Frank, M.J., Schweizer, B. (2006) *Associative Functions: Triangular Norms and Copulas*, World Scientific Publishing Co. (Singapore).

Amblard, C. and Girard, S. (2002) Symmetry and Dependence properties within a semiparametric family of bivariate copulas. *J. Nonparametr. Statist.* **14**, (6), 715-727.

Bailey, D.F. (1996) Counting arrangements of 1's and -1's. *Math. Mag.* **69**, (2), 128-131.

Barcucci, E., Verri, M.C. (1992) Some more properties of Catalan numbers. *Discrete Mathematics* **102**, 229-237.

Block, H. and Ting, M.L. (1981) Some concepts of multivariate dependence. *Comm. Statist.-Theor. Meth.* **A10**, (8), 749-762.

Block, H., Savits, T., Shaked, M. (1982) Some concepts of negative dependence. *Ann. Probab.* **10**, 765-772.

Blomqvist, N. (1950) On a measure of dependence between two random variables. *Ann. Math. Statist.* **21**, 593-600.

Blum, J.R., Kiefer, J., Rosenblatt, M. (1961) Distribution-free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32**, 485-498.

Callan, D. (2006) Some bijections and identities for the Catalan and Fine numbers. *Séminaire Lotharingen de Combinatoire* **53**, B53e, 16 pp.

Cameron, N.T. (2002) *Random Walks, Trees and Extensions of Riordan Group Techniques*, Ph.D. Dissertation, Howard University (Dr. Louis Shapiro, advisor).

Carley, H. and Taylor, M.D. (2002) A new proof of Sklar's theorem. *Distributions with Given Marginals and Statistical Modelling; Carles M. Cuadras, Josep Fortiana and José A. Rodríguez-Lallena editors (Kluwer)* , 29-34.

Castillo, E. and Ruiz, R. (1993) *Ecuaciones funcionales y modelización en Ciencia, Ingeniería y Economía*, Reverté (Barcelona).

Chung, K.L., Feller, W. (1949) On fluctuations in coin-tossing. *Proc. Nat. Acad. Sci. USA* **35**, 605-608.

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**, 141-151.

Dall'Aglio, C. (1959) Sulla compatibilità delle funzioni di ripartizione doppia. *Rend. Mat.* **18**, (5), 385-413.

Dall'Aglio, C. (1960) Les fonctions extrêmes de la clase de Fréchet à 3 dimensions. *Publ. Inst. Statist. Paris* **9**, 175-188.

Dall'Aglio, C. (1972) Fréchet classes and compatibility of distribution functions. *Symposia Math.* **9**, 131-150.

Darsow, W.F. and Frank, M.J. (1983) Associative functions and Abel-Schröder systems. *Publ. Math. Debrecen* **30**, 253-272.

Deheuvels, P. (1979) La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci. (5)* **65**, 274-292.

Deheuvels, P. (1981) A non parametric test for independence. *Publ. Inst. Statist. Univ. Paris* **26**, 29-50.

De Matteis, R. (2001) *Fitting copulas to data*, Diploma Thesis ETH Zürich (Prof. Alexander McNeil, supervisor).

Dolati, A., Úbeda-Flores, M. (2006) On Measures of Multivariate Concordance. *J. Prob. Statist. Sci.* **4**, (2), 147-163.

Drouet Mari, D. and Kotz, S. (2001) *Correlation and dependence*, Imperial College Press (London).

Dudley, R.M. (2002) *Real Analysis and Probability*, Cambridge University Press (New York).

Durante, F., Quesada-Molina, J.J., Sempi, C. (2006) A generalization of the Archimedean class of bivariate copulas. *Ann. Inst. Statist. Math.* (to appear).

Embrechts, P., McNeil, A.J. and Straumann, D. (1999) Correlation: pitfalls and alternatives. *Risk Magazine* **5**, 69-71.

Embrechts, P., Lindskog, F. and McNeil, A.J. (2003a) Modeling dependence with copulas and applications to risk management. *Handbook of Heavy-Tailed Distributions in Finance, Rachev, S. (editor) Elsevier, chapter 8,* 329-384.

Engleberg, E.O. (1965) On some problems concerning a restricted random walk. *J. Appl. Prob.* **2**, 396-404.

Erdely, A. and González-Barrios, J.M. (2006a) On the construction of families of absolutely continuous copulas with given restrictions. *Comm. Statist.- Theory and Methods* **35**, 649-659.

Erdely, A. and González-Barrios, J.M. (2006b) Distributional properties of the diagonal of the empirical copula and a nonparametric test for independence for the Archimedean family. *IIMAS-UNAM Preimpreso* **139**, 1-24.

Esary, J.D. and Proschan, F. (1972) Relationships among some concepts of bivariate dependence. *Ann. Math. Statist.* **43**, (2), 651-655.

Eyraud (1938) Les principles de la mesure des correlations. *Ann. Univ. Lyon Series A* **1**, 30-47.

Farlie, D.J.G. (1960) The performance of some correlation coefficients for a general bivariate distribution. *Biometrika* **47**, 307-323.

Feller, W. (1968) *An introduction to probability theory and its applications, Vol. 1 (3rd edition)*, Wiley (New York).

Fermanian, J.D., Radulović, D., and Wegcamp, M. (2004) Weak convergence of the empirical copula process. *Bernoulli* **10**, (5), 847-860.

Fernández, B. and González-Barrios, J.M. (2004) Multidimensional dependency measures. *J. Multivariate Anal.* **89**, 351-370.

Féron, R. (1956) Sur les tableaux de corrélation dont les marges sont données. *Publ. Inst. Statist. Univ. Paris* **5**, 3-12.

Frank, M.J. (1996) Diagonals of copulas and Schröder's equation. *Aequationes Math.* **51**, 150.

Fréchet, M. (1951) Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon* **14**, (Sect. A Ser. 3), 53-77.

Fréchet, M. (1957) Les tableaux de corrélation et les programmes linéaires. *Revue Inst. Int. Statist.* **25**, 23-40.

Fredricks, G.A. and Nelsen, R.B. (1997a) Diagonal copulas. *Distributions with Given Marginals and Moment Problems, V. Beneš and J. Štěpán, editors (Kluwer Academic Publishers, Dordrecht)* 121-128.

Fredricks, G.A. and Nelsen, R.B. (1997b) Copulas constructed from diagonal sections. *Distributions with Given Marginals and Moment Problems, V. Beneš and J. Štěpán, editors (Kluwer Academic Publishers, Dordrecht)* 129-136.

Fredricks, G.A. and Nelsen, R.B. (2002) The Bertino family of copulas. *Distributions with Given Marginals and Statistical Modelling. C.M. Cuadras, J. Fortiana and J.A. Rodríguez-Lallena, editors (Kluwer Academic Publishers, Dordrecht)* 81-91.

Garralda-Guillén, A.I. (1997) *Dependencia y tests de rangos para leyes bidimensionales*, Ph.D. Dissertation, Universidad de Granada (Spain).

Genest, C. and Verret, F. (2005) Locally most powerful rank tests of independence for copula models. *J. Nonparametr. Statist.* **17**, (5), 521-539.

Genest, C., Quessy, J.F., Rémillard, B. (2006) Local efficiency of a Cramér-von Mises test of independence. *J. Multivariate Anal.* **97**, (1), 274-294.

Gibbons, J.D. and Chakraborti, S. (2003) *Nonparametric Statistical Inference*, Dekker (New York).

Gini, C. (1914) Di una misura delle relazioni tra le graduatorie di due caratteri. *Appendix to: Mancini, A., "Le Elezioni Generali Politiche del 1913 nel Comune di Roma," Ludovico Cecchini (Rome).*

González-Barrios, J.M. (2003b) Some Notes on a Dependency Measure. *Contemporary Mathematics* **336**, 171-179.

Gumbel, E.J. (1958) Distributions à plusieures variables dont les marges sont donnés. *Comptes Rendus de l'Academie des Sciences Paris* **246**, 2717.

Gumbel, E.J. (1960) Bivariate exponential distributions. *J. Amer. Statist. Assoc.* **55**, 698-707.

Harris, R. (1970) A Multivariate Definition for Increasing Hazard Rate. *Ann. Math. Statist.* **41**, 713-717.

Hoeffding, W. (1940) Masstabinvariante Korrelationstheorie. *Schriften des Matematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin* **5**, 179-223.

Hollander, M. and Wolfe, D.A. (1999) *Nonparametric statistical methods*, Wiley-Interscience (New York).

Itô, K., editor (1996) *Encyclopedic Dictionary of Mathematics*, The MIT Press (Cambridge).

Joe, H. (1990) Multivariate concordance. *J. Multivariate Anal.* **35**, 12-30.

Joe, H. (1996) Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer and Taylor, editors* **28**, 120-141.

Joe, H. (1997) *Multivariate models and dependence concepts*, Ed. Chapman & Hall (New York).

Kallenberg, W.C.M. and Ledwina, T. (1999) Data-Driven Rank Tests for Independence. *J. Amer. Stat. Assoc.* **94**, (445), 285-301. [4]

Kendall, M.G. (1938) A new measure of rank correlation. *Biometrika* **30**, 81-93.

Kimberling, C.H. (1974) A probabilistic interpretation of complete monotonicity. *Aequationes Math.* **10**, 152-164.

Kimeldorf, G., Sampson, A. (1978) Monotone dependence. *Ann. Statist.* **6**, 895-903.

Kimeldorf, G., Sampson, A. (1987) Positive dependence orderings. *Ann. Inst. Statist. Math.* **39**, 113-128.

Kimeldorf, G., Sampson, A. (1989) A framework for positive dependence. *Ann. Inst. Statist. Math.* **41**, 31-45.

Klement, E.P., Mesiar, R., Pap, E. (2000) *Triangular Norms*, Kluwer (Dordrecht).

Klement, E.P., Mesiar, R., editors (2005) *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, Elsevier (Amsterdam).

Koenigs, G. (1884) Recherches sur les intégrales de certaines équations fonctionnelles. *Ann. Sci. Ecole Norm. Sup.* **1**, (3), Supplément, 3-41.

Kotz, S. and Seeger, J.P. (1991) A new approach to dependence in multivariate distributions. *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, and Salinetti editors (Kluwer Academic Publishers, Dordrecht),* 113-127.

Krause, G. (1981) *A strengthened form of Ling's theorem on associative functions*, Ph.D. thesis (Illinois Institute of Technology).

Kruskal, W.H. (1958) Ordinal measures of association. *Amer. Statist.* **53**, 103-106.

Kuczma, M. (1968) *Functional equations in a single variable*, Polish Scientific Publishers (Warsaw).

Kuczma, M., Choczewski, B., and Ger, R. (1990) *Iterative functional equations*, Cambridge University Press (New York).

Lancaster, H.O. (1963) Correlation and complete dependence of random variables. *Ann. Math. Statist.* **34**, 1315-1321.

Lehmann, E.L. (1966) Some concepts of dependence. *Ann. Math. Statist.* **37**, 1137-1153.

Lehmann, E.L. (1975) *Nonparametrics: Statistical Methods Based on Ranks*, McGraw-Hill (San Francisco).

Li, X., Mikusiński, P., and Taylor, M.D. (2002) Some Integration-by-Parts Formulas involving 2-Copulas. *Distributions with Given Marginals and Statistical Modelling; Carles M. Cuadras, Josep Fortiana and José A. Rodríguez-Lallena editors (Kluwer)* , 153-159.

Ling, C.H. (1965) Representation of associative functions. *Publ. Math. Debrecen* **12**, 189-212.

Marshall, A.W. and Olkin, I. (1967) A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62**, 30-44.

Marshall, A.W. (1996) Copulas, Marginals, and Joint Distributions. *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer and Taylor, editors* **28**, 213-222.

Mesiar, R. (2006) Personal communication with J.M. González-Barrios.

Mikusiński, P., Sherwood, H., Taylor, M.D. (1991) Probabilistic interpretations of copulas and their convex sums. *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, and Salinetti editors (Kluwer Academic Publishers, Dordrecht),* 95-112.

Mikusiński, P., Sherwood, H., Taylor, M.D. (1992) Shuffles of Min. *Stochastica* **13**, 61-74.

Mikusiński, P. and Taylor, M.D. (1999) A remark on associative copulas. *Comment. Math. Univ. Carolinae* **40**, (4), 789-793.

Morgenstern, D. (1956) Einfache Beispiele Zweidimensionaler Verteilungen. *Mitteilingsblatt für Mathematische Statistik* **8**, 234-235.

Nelsen, R.B. (1991) Copulas and association. *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, and Salinetti editors (Kluwer Academic Publishers, Dordrecht),* 51-74.

Nelsen, R.B. (1995) Copulas, Characterization, Correlation, and Counterexamples. *Mathematics Magazine* **68**, (3), 193-198.

Nelsen, R.B. (1996a) Nonparametric Measures of Multivariate Association. *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer and Taylor, editors* **28**, 223-232.

Nelsen, R.B. (2002) Concordance and copulas: A survey. *Distributions with Given Marginals and Statistical Modelling; Carles M. Cuadras, Josep Fortiana and José A. Rodríguez-Lallena editors (Kluwer)* , 169-178.

Nelsen, R.B. (2006a) *An introduction to copulas*, Ed. Springer-Verlag (New York) $2^{\underline{nd}}$ edition (first edition dates from 1999).

Nelsen, R.B. (2006b) Personal communication.

Nelsen, R.B., Quesada-Molina, J.J., Rodríguez-Lallena, J.A., Úbeda-Flores, M. (2006) On the construction of asymmetric copulas with a given diagonal section (in preparation).

Pečarič, J.E., Proschan, F., Tong, Y.L. (1992) *Convex Functions, Partial Orderings and Statistical Applications*, Ser. Math. in Science and Engin. **187**, Academic Press (Boston).

Quesada-Molina, J.J. and Rodríguez-Lallena, J.A. (1994) Some advances in the study of the compatibility of three bivariate copulas. *J. Ital. Statist. Soc.* **3**, 397-417.

Quesada-Molina, J.J., Rodríguez-Lallena, J.A., Úbeda-Flores, M. (2006) Some New Characterizations and Results on Quasi-copulas. *10th International Congress on Insurance: Mathematics and Economics (Leuven, Belgium).*

Rényi, A. (1959) On measures of dependence. *Acta. Math. Acad. Sci. Hungar.* **10**, 441-451.

Rodríguez-Lallena, J.A. and Úbeda-Flores, M. (2004) A new class of bivariate copulas. *Statist. Probab. Lett.* **66**, (3), 315-325.

Scarsini, M. (1984) On measures of concordance. *Stochastica* **8**, 201-218.

Schröder, E. (1871) Über iterierte Funktionen. *Math. Ann.* **3**, 296-322.

Schweizer, B., Wolff, E.F. (1981) On nonparametric measures of dependence for random variables. *Ann. Statist.* **9**, (4), 870-885.

Schweizer, B. (1991) Thirty years of Copulas. *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, and Salinetti editors (Kluwer Academic Publishers, Dordrecht),* 13-50.

Schweizer, B., Sklar, A. (2005) *Probabilistic Metric Spaces*, Dover (New York). Republication of the 1983 edition, with supplementary historical notes and references.

Schweizer, B. (2006) Personal communication.

Shaked, M. (1977) A family of concepts of dependence for bivariate distributions. *J. Amer. Statist. Assoc.* **72**, 642-650.

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

Sklar, A. (1996a) Random Variables, Distribution Functions, and Copulas - A Personal Look Backward and Forward. *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer and Taylor, editors* **28**, 1-14.

Spearman, C. (1904) The proof and measurement of association between two things. *Amer. J. Psychol.* **15**, 72-101.

Sungur, E.A. and Yang, Y. (1996) Diagonal copulas of archimedean class. *Commun. Statist.-Theory and Methods* **25**, (7), 1659-1676.

Taylor, M.D. (2006) Multivariate measures of concordance. *Ann. Inst. Statist. Math.* (to appear).

van der Vaart, A.W. and Wellner, J.A. (1996) *Weak Convergence and Empirical Processes*, Springer-Verlag (New York).

Widder, D.V. (1946) *The Laplace Transform*, Princeton University Press (Princeton).

Wolff, E.F. (1980) N-dimensional measures of dependence. *Stochastica* **4**, (3), 175-188.