An introduction to Copulas

Carlo Sempi

Dipartimento di Matematica “Ennio De Giorgi”
Università del Salento
Lecce, Italy
carlo.sempi@unisalento.it

The 33rd Finnish Summer School on Probability Theory and Statistics, June 6th–10th, 2011

Outline

The beginning of the story

The history of copulas may be said to begin with Fréchet (1951). Fréchet's problem: given the distribution functions F_j $(j = 1, 2, \dots, d)$ of d r.v.'s $X_1, X_2, \dots, X_d$ defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, what can be said about the set $\Gamma(F_1, F_2, \dots, F_d)$ of the d–dimensional d.f.'s whose marginals are the given F_j?

$$H \in \Gamma(F_1, \dots, F_d) \iff H(+\infty, \dots, +\infty, t, +\infty, \dots, +\infty) = F_j(t)$$

The set $\Gamma(F_1, \dots, F_d)$ is called the Fréchet class of the F_j's. Notice $\Gamma(F_1, \dots, F_d) \neq \emptyset$ since, if $X_1, X_2, \dots, X_d$ are independent, then

$$H(x_1, x_2, \dots, x_d) = \prod_{j=1}^{d} F_j(x_j).$$

But, it was not clear which the other elements of $\Gamma(F_1, \dots, F_d)$ were.

Bibliography–1

For Fréchet's work see, e.g.,

M. Fréchet, Sur les tableaux de corrélation dont les marges sont donnés, *Ann. Univ. Lyon, Science,* **4**, 13–84 (1951)

G. Dall'Aglio, Fréchet classes and compatibility of distribution functions, *Symposia Math.,* **9**, 131–150 (1972)

In this latter paper Dall'Aglio studies under which conditions there is just one d.f. belonging to $\Gamma(F_1, F_2)$.

Enters Sklar

In 1959, Sklar obtained the most important result in this respect, by introducing the notion, and the name, of a *copula*, and proving the theorem that now bears his name.

Correspondence with Fréchet

He and Bert Schweizer had been making progress in their work on statistical metric spaces, to the extent that Menger suggested it would be worthwhile to communicate their results to Fréchet. Fréchet was interested, and asked to write an announcement for the *Comptes Rendus*. This lead to an exchange of letters between Sklar and Fréchet, in the course of which Fréchet sent Sklar several packets of reprints, mainly dealing with the work he and his colleagues were doing on distributions with given marginals. These reprints were important for much of the subsequent work. At the time, though, the most significant reprint for Sklar was that of Féron (1956).

Sklar–2

Féron, in studying three-dimensional distributions had introduced auxiliary functions, defined on the unit cube, that connected such distributions with their one-dimensional margins. Sklar saw that similar functions could be defined on the unit d–cube for all $d \geq 2$ and would similarly serve to link d–dimensional distributions to their one–dimensional margins. Having worked out the basic properties of these functions, he wrote about them to Fréchet, in English.

Sklar–3

Fréchet asked Sklar to write a note about them in French. While writing this, Sklar decided he needed a name for these functions. Knowing the word "copula" as a grammatical term for a word or expression that links a subject and predicate, he felt that this would make an appropriate name for a function that links a multidimensional distribution to its one-dimensional margins, and used it as such. Fréchet received Sklar's note, corrected one mathematical statement, made some minor corrections to Sklar's French, and had the note published by the Statistical Institute of the University of Paris (Sklar, 1959).

A curiosity

Curiously, it should be noted that in that paper, the author "Abe Sklar" is named as "M. Sklar" (should it be intended as "Monsieur"?)

Lack of a proof

The proof of Sklar's theorem was not given in (Sklar, 1959), but a sketch of it was provided in (Sklar, 1973). (see also (Schweizer & Sklar, 1974)), so that for a few years practitioners in the field had to reconstruct it relying on the hand–written notes by Sklar himself; this was the case, for instance, of the present speaker. It should be also mentioned that some "indirect" proofs of Sklar's theorem (without mentioning copula) were later discovered by Moore & Spruill and Deheuvels.

For about 15 years, all the results concerning copulas were obtained in the framework of the theory of Probabilistic Metric spaces (Schweizer & Sklar, 1974). The event that arose the interest of the statistical community in copulas occurred in the mid seventies, when Bert Schweizer, in his own words (Schweizer, 2007),

> *quite by accident, reread a paper by A. Rényi, entitled On measures of dependence and realized that [he] could easily construct such measures by using copulas.*

The first building blocks were the announcement by Schweizer & Wolff in the *Comptes Rendus de l'Académie des Sciences* (1976) and Wolff's Ph.D. Dissertation at the University of Massachusetts at Amherst (1977). These results were presented to the statistical community in (Schweizer & Wolff, 1981) (see also (Wolff, 1980)).

However, for several other years, Chapter 6 of the 1983 book by Schweizer & Sklar, devoted to the theory of Probabilistic metric spaces, was the main source of basic information on copulas. Again in Schweizer's words from (Schweizer, 2007),

> *After the publication of these articles and of the book ... the pace quickened as more ... students and colleagues became involved. Moreover, since interest in questions of statistical dependence was increasing, others came to the subject from different directions. In 1986 the enticingly entitled article "The joy of copulas" by C. Genest and R.C MacKay (1986), attracted more attention.*

Finance

At end of the nineties, the notion of copulas became increasingly popular. Two books about copulas appeared and were to become the standard references for the following decade. In 1997 Joe published his book on multivariate models, with a great part devoted to copulas and families of copulas. In 1999 Nelsen published the first edition of his introduction to copulas (reprinted with some new results in 2006).
But, the main reason of this increased interest has to be found in the discovery of the notion of copulas by researchers in several applied field, like finance. Here we should like briefly to describe this explosion by quoting Embrechts's comments (Embrechts, 2009).

Embrechts

...the notion of copula is both natural as well as easy for looking at multivariate d.f.'s. But why do we witness such an incredible growth in papers published starting the end of the nineties (recall, the concept goes back to the fifties and even earlier, but not under that name)? Here I can give three reasons: finance, finance, finance. In the eighties and nineties we experienced an explosive development of quantitative risk management methodology within finance and insurance, a lot of which was driven by either new regulatory guidelines or the development of new products .... Two papers more than any others "put the fire to the fuse": the ...1998 RiskLab report (Embrechts et al., 2002) and at around the same time, the Li credit portfolio model (Li, 2001).

Today

The advent of copulas in finance originated a wealth of investigations about copulas and, especially, applications of copulas. At the same time, different fields like hydrology discovered the importance of this concept for constructing more flexible multivariate models. Nowadays, it is near to impossible to give a complete account of all the applications of copulas to the many fields where they have be used.
Since the field is still *in fieri*, it is important from time to time to survey the progresses that have been achieved, and the new questions that they pose. The aim of this talk is to survey the recent literature.

Today–2

To quote Schweizer again:

> *The "era of i.i.d." is over: and when dependence is taken seriously, copulas naturally come into play. It remains for the statistical community at large to recognize this fact. And when every statistics text contains a section or a chapter on copulas, the subject will have come of age.*

Random variables and vectors

When a r.v. $\mathbf{X} = (X_1, X_2, \dots, X_d)$ is given, two problems are interesting:

- to study the probabilistic behaviour of each one of its components;
- to investigate the relationship among them.

It will be seen how copulas allow to answer the second one of these problems in an admirable and thorough way.
It is a general fact that in probability theory, theorems are proved in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, while computations are usually carried out in the measurable space $(\overline{\mathbb{R}}^d, \mathcal{B}(\overline{\mathbb{R}}^d))$ endowed with the law of the random vector $\mathbf{X}$.

Distribution functions

The study of the law $\mathbb{P}_{\mathbf{X}}$ is made easier by the knowledge of the distribution function(=d.f.), as defined here.
Given a random vector $\mathbf{X} = (X_1, X_2, \dots, X_d)$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, its *distribution function* $F_{\mathbf{X}} : \overline{\mathbb{R}}^d \to \mathbb{I}$ is defined by

$$F_{\mathbf{X}}(x_1, x_2, \dots, x_d) = \mathbb{P}\left(\cap_{i=1}^{d} \{X_i \le x_i\}\right) \qquad (1)$$

if all the x_i's are in $\mathbb{R}$, while:

- $F_{\mathbf{X}}(x_1, x_2, \dots, x_d) = 0$, if at least one of the arguments equals $-\infty$
- $F_{\mathbf{X}}(+\infty, +\infty, \dots, +\infty) = 1$.

C–volume

A d–box is a cartesian product

$$[\mathbf{a}, \mathbf{b}] = \prod_{j=1}^{d} [a_j, b_j],$$

where, for every index $j \in \{1, 2, \dots, d\}$, $0 \le a_j \le b_j \le 1$.
For a function $C : \mathbb{I}^d \to \mathbb{I}$, the C–volume V_C of the box $[\mathbf{a}, \mathbf{b}]$ is defined via

$$V_C\left([\mathbf{a}, \mathbf{b}]\right) := \sum_{\mathbf{v}} \operatorname{sign}(\mathbf{v})\, C(\mathbf{v})$$

where the sum is carried over all the 2^d vertices $\mathbf{v}$ of the box $[\mathbf{a}, \mathbf{b}]$; here

$$\operatorname{sign}(\mathbf{v}) = \begin{cases} 1, & \text{if } v_j = a_j \text{ for an even number of indices,} \\ -1, & \text{if } v_j = a_j \text{ for an odd number of indices.} \end{cases}$$

Properties of distribution functions

Theorem

The d.f. $F_{\mathbf{X}}$ *of the r.v.* $\mathbf{X} = (X_1, X_2, \dots, X_d)$ *has the following properties:*

- F *is isotone, i.e.* $F(\mathbf{x}) \le F(\mathbf{y})$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $\mathbf{x} \le \mathbf{y}$;
- *for all* $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \in \mathbb{R}^{d-1}$, *the function*

$$\mathbb{R} \ni t \mapsto F_X(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_d)$$

  is right–continuous;
- *for every* d*–box* $[\mathbf{a}, \mathbf{b}]$, $V_{F_{\mathbf{X}}}([\mathbf{a}, \mathbf{b}]) \ge 0$.

Marginals

Let F be a d–dimensional d.f. ($d \geq 2$). Let $\sigma = (j_1, \dots, j_m)$ a subvector of $(1, 2, \dots, d)$, $1 \leq m \leq d-1$. We call σ–*marginal* of F the d.f. $F_\sigma : \overline{\mathbb{R}}^m \to \mathbb{I}$ defined by setting $d - m$ arguments of F equal to $+\infty$, namely, for all $x_1, \dots, x_m \in \overline{\mathbb{R}}$,

$$F_\sigma(x_1, \dots, x_m) = F(y_1, \dots, y_d),$$

where, for every $j \in \{1, 2, \dots, d\}$, $y_j = x_j$ if $j \in \{j_1, \dots, j_m\}$, and $y_j = +\infty$ otherwise.

In particular, when $\sigma = \{j\}$, $F_{(j)}$ is usually called 1–*dimensional marginal* and it is denoted by F_j.

If F is the d.f. of the r.v. $\vec{X} = (X_1, X_2, \dots, X_d)$, then the σ–marginal of F is the d.f. of the subvector $(X_{j_1}, \dots, X_{j_m})$.

The definition

Definition

For $d \geq 2$, a d–*dimensional copula* (shortly, a d–copula) is a d–variate d.f. on $\mathbb{I}^d$ whose univariate marginals are uniformly distributed on $\mathbb{I}$.

Each d-copula may be associated with a r.v. $\mathbf{U} = (U_1, U_2, \dots, U_d)$ such that $U_i \sim \mathcal{U}(\mathbb{I})$ for every $i \in \{1, 2, \dots, d\}$ and $\mathbf{U} \sim C$.
Conversely, any r.v. whose components are uniformly distributed on $\mathbb{I}$ is distributed according to some copula.
The class of all d–copulas will be denoted by $\mathcal{C}_d$.

A characterization

Theorem

A function $C : \mathbb{I}^d \to \mathbb{I}$ *is a copula if, and only if, the following properties hold:*

- *for every* $j \in \{1, 2, \dots, d\}$, $C(\mathbf{u}) = u_j$ *when all the components of* $\mathbf{u}$ *are equal to* 1 *with the exception of the* j*–th one that is equal to* $u_j \in \mathbb{I}$;
- C *is isotonic, i.e.* $C(\mathbf{u}) \leq C(\mathbf{v})$ *for all* $\mathbf{u}, \mathbf{v} \in \mathbb{I}^d$ *such that* $\mathbf{u} \leq \mathbf{v}$;
- C *is* d*–increasing.*

The special case $d = 2$

Explicitly, a bivariate copula is a function $C : \mathbb{I}^2 \to \mathbb{I}$ such that

- $\forall u \in [0,1] \qquad C(u,0) = C(0,u) = 0$
- $\forall u \in [0,1] \qquad C(u,1) = C(1,u) = u$
- for all u, u', v, v' in $\mathbb{I}$ with $u \le u'$ and $v \le v'$

$$C(u',v') - C(u',v) - C(u,v') + C(u,v) \geq 0$$

This last inequality is referred to as the rectangular inequality; a function that satisfies it is said to be 2–increasing.

Consequences

- $C(\mathbf{u}) = 0$ for every $\mathbf{u} \in \mathbb{I}^d$ having at least one of its components equal to 0
- (The 1–Lipschitz property): for all $\mathbf{u}, \mathbf{v} \in \mathbb{I}^d$,

$$|C(\mathbf{u}) - C(\mathbf{v})| \leq \sum_{i=1}^{d} |u_i - v_i|.$$

- $\mathcal{C}_d$ is a compact set in the set $C(\mathbb{I}^d, \mathbb{I})$ of all continuous functions from $\mathbb{I}^d$ into $\mathbb{I}$ equipped with the topology of pointwise convergence.
- Pointwise and uniform convergence are equivalent in $\mathcal{C}_d$.

Examples–1

- The *independence copula* $\Pi_d(\mathbf{u}) = u_1\, u_2 \cdots u_d$ associated with a random vector $\mathbf{U} = (U_1, U_2, \dots, U_d)$ whose components are independent and uniformly distributed on $\mathbb{I}$.
- The *comonotonicity copula* $\mathrm{Min}_d(\mathbf{u}) = \min\{u_1, u_2, \dots, u_d\}$ associated with a vector $\mathbf{U} = (U_1, U_2, \dots, U_d)$ of r.v.'s uniformly distributed on $\mathbb{I}$ and such that $U_1 = U_2 = \cdots = U_d$ almost surely.
- The *countermonotonicity copula* $W_2(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\}$ associated with a bivariate vector $\mathbf{U} = (U_1, U_2)$ of r.v.'s uniformly distributed on $\mathbb{I}$ and such that $U_1 = 1 - U_2$ almost surely.

Examples–2: Convex combinations

Convex combinations of copulas: Let $\mathbf{U}_1$ and $\mathbf{U}_2$ be two d–dimensional r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$ distributed according to the copulas C_1 and C_2, respectively. Let Z be a Bernoulli r.v. such that $\mathbb{P}(Z = 1) = \alpha$ and $\mathbb{P}(Z = 2) = 1 - \alpha$ for some $\alpha \in \mathbb{I}$. Suppose that $\mathbf{U}_1$, $\mathbf{U}_2$ and Z are independent. Now, consider the d–dimensional r.v. $\mathbf{U}^*$

$$\mathbf{U}^* = \sigma_1(Z)\,\mathbf{U}_1 + \sigma_2(Z)\,\mathbf{U}_2$$

where, for $i \in \{1, 2\}$, $\sigma_i(x) = 1$, if $x = i$, $\sigma_i(x) = 0$, otherwise. Then, $\mathbf{U}^*$ is distributed according to the copula $\alpha\, C_1 + (1 - \alpha)\, C_2$.

Examples–3

- Fréchet–Mardia family of copulas

$$C_d^{\mathbf{FM}}(\mathbf{u}) = \lambda\, \Pi_d(\mathbf{u}) + (1-\lambda)\, M_d(\mathbf{u})$$

  for every $\lambda \in \mathbb{I}$. A convex sum of the copulas Π_d and M_d.

- Cuadras–Augé family; for $\alpha \in \mathbb{I}$,

$$C_d^{\mathbf{CA}}(\mathbf{u}) = (\Pi_d(\mathbf{u}))^{\alpha}\, (M_d(\mathbf{u}))^{1-\alpha},$$

The derivatives

Consider a bivariate copula $C \in \mathcal{C}_2$. For every $v \in \mathbb{I}$, the functions

$$\mathbb{I} \ni t \to C(t, v)$$
$$\mathbb{I} \ni t \to C(v, t)$$

are increasing; therefore, their first derivatives exists almost everywhere with respect to Lebesgue measure and are positive, where they exist. Because of the Lipschitz condition, they are also bounded above by 1

$$0 \leq D_1 \, C(s,t) \leq 1 \qquad 0 \leq D_2 C(s,t) \leq 1 \quad \text{a.e.}$$

where

$$D_1 \, C(s,t) := \frac{\partial C(s,t)}{\partial s} \quad \text{and} \quad D_2 \, C(s,t) := \frac{\partial C(s,t)}{\partial t}$$

A useful formula

The following integration–by–parts formula is sometimes useful in the computation of statistical quantities.

Theorem

Let A and B be 2–copulæ, and let the function $\varphi : \mathbb{I} \to \mathbb{R}$ be continuously differentiable, i.e., $\varphi \in C_1$. Then

$$\begin{aligned}\int_{[0,1]^2} \varphi \circ A \, dB &= \int_0^1 \varphi(t)\, dt - \int_{[0,1]^2} \varphi'(A)\, D_1 A\, D_2 B\, du\, dv \\ &= \int_0^1 \varphi(t)\, dt - \int_{[0,1]^2} \varphi'(A)\, D_2 A\, D_1 B\, du\, dv\end{aligned}$$

Fréchet–Hoeffding bounds

Theorem

For every $C_d \in \mathcal{C}_d$ *and for every* $\mathbf{u} \in \mathbb{I}^d$,

$$W_d(\mathbf{u}) = \max\left\{\sum_{i=1}^{d} u_i - d + 1, 0\right\} \leq C(\mathbf{u}) \leq M_d(\mathbf{u}).$$

These bounds are sharp:

$$\inf_{C \in \mathcal{C}_d} C(\mathbf{u}) = W_d(\mathbf{u}), \qquad \sup_{C \in \mathcal{C}_d} C(\mathbf{u}) = M_d(\mathbf{u}).$$

Notice that, while W_2 is a copula, W_d is not a copula for $d \geq 3$.

The marginals of a copula

A marginal of an d–copula C is obtained by setting some of its argument equal 1. A k–marginal of C, $k < d$, is obtained by setting exatly $d - k$ arguments equal to 1; therefore, there are

$$\binom{d}{k}$$

k–marginals of the d–copula C.
In particular, the d 1–marginals are easily computed:

$$C(1, 1, \dots, 1, u_j, 1, \dots, 1) = u_j \qquad (j = 1, 2, \dots, d)$$

Sklar's Theorem

Theorem

Given an d–dimensional d.f. H there exists an d–copula C such that for all $(x_1, x_2, \ldots, x_d) \in \overline{\mathbb{R}}^n$

$$H(x_1, x_2, \ldots, x_d) = C\left(F_1(x_1), F(x_2), \ldots, F_d(x_d)\right) \tag{2}$$

The copula C is uniquely defined on $\prod_{j=1}^{d} \text{ran } F_j$ and is therefore unique if all the marginals are continuous.
Conversely, if F_1, F_2,..., F_d are d (1–dimensional) d.f.'s, then the function H defined through eq. (2) is an d–dimensional d.f..

How to obtain a copula from a joint d.f.

Given a d–variate d.f. F, one can derive a copula C. Specifically, when the marginals F_i are continuous, C can be obtained by means of the formula

$$C(u_1, u_2, \dots, u_d) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_d^{-1}(u_d)),$$

where F_i^{-1} denoted the pseudo–inverse of F_i,

$$F_i^{-1}(s) = \inf\{t \mid F_i(t) \geq s\}.$$

Thus, copulæ are essentially a way for transforming the r.v. $(X_1, X_2 \dots, X_d)$ into another r.v.

$$(U_1, U_2, \dots, U_d) = (F_1(X_1), F_2(X_2), \dots, F_d(X_d))$$

having the margins uniform on $\mathbb{I}$ and preserving the dependence among the components.

The uniqueness question

Sklar's theorem immediately poses the question of the uniqueness of the copula C:

If the r.v.'s involved, or, equivalently, their d.f.'s, are both continuous, then the copula C is unique.

If at least one of the d.f.'s has a discrete component, then the copula C is uniquely defined only on the product of the ranges $\text{ran } F_1 \times \text{ran } F_2 \times \cdots \times \text{ran } F_d$, and there may well be more than one copula extending C from this cartesian product to the whole unit cube $\mathbb{I}^d$. In this latter case it is costumary to have recourse to a procedure of bilinear interpolation in order to single out a unique copula; this allow to speak of *the* copula of the pair (X, Y). See Lemma 2.3.5 in (Nelsen, 2006) or (Darsow, Nguyen & Olsen, 1992)

Comments

- Notice that in many papers where copulæ are applied there is often hidden the assumption that the r.v.'s involved are continuous; this avoids the uniqueness question.
- If all the d.f.'s involved are continuous then to each joint d.f. in the Fréchet class $\Gamma(F_1, F_2, \dots, F_d)$ there corresponds a unique d–copula $C \in \mathcal{C}_d$; otherwise, to each $H \in \Gamma(F_1, F_2, \dots, F_d)$ there corresponds the set of copulas in $\mathcal{C}_d$ that coincide on

$$\prod_{j=1}^{d} \operatorname{ran} F_j$$

Comments–2

The second part of Sklar's theorem is very easy to prove, but it is extremely important for the applications; it is, in fact, the very foundation of all the models built on copulas. Models are built according to the following scheme:

- the d rv's $X_1, X_2, \dots, X_d$ are individually described by their 1–dimensional d.f.'s $F_1, F_2, \dots, F_d$
- then a copula $C \in \mathcal{C}_d$ is introduced; this contains every piece of information about the dependence relationship among the r.v.'s $X_1, X_2, \dots, X_d$, *independently* of the choice of the marginals $F_1, F_2, \dots, F_d$.

In particular, copulas can serve for modelling situations where a different distribution is needed for each marginal, providing a valid alternative to several classical multivariate d.f.'s such Gaussian, Pareto, Gamma, etc.. This fact represents one of the main advantage of the copula's idea.

Caution–2

Sklar's theorem should be used with some caution when the margins have jumps. In fact, even if there exists a copula representation for non–continuous joint d.f.'s, it is no longer unique. In such cases, modelling and interpreting dependence through copulas needs some caution. The interested readers should refer to the paper (Marshall, 1996) and to the in–depth discussion by Genest and Nešlehová (2007).

Survival copulæ

Sklar's Theorem can be formulated in terms of survival functions instead of d.f.'s. Specifically, given a r.v. $\mathbf{X} = (X_1, X_2, \dots X_d)$ with joint survival function $\overline{F}$ and univariate survival marginals $\overline{F}_i$ $(i = 1, 2, \dots, d)$, for all $(x_1, x_2, \dots, x_n) \in \overline{\mathbb{R}}^d$

$$\overline{F}(x_1, x_2, \dots, x_d) = \widetilde{C}\left(\overline{F}_1(x_1), \overline{F}_2(x_2), \dots, \overline{F}_d(x_d)\right).$$

for some copula $\widetilde{C}$, usually called the *survival copula* of $\mathbf{X}$ (the copula associated with the survival function of $\mathbf{X}$).

Survival copulæ–2

In particular, let C be the copula of $\mathbf{X}$ and let $\mathbf{U} = (U_1, U_2, \dots, U_d)$ be a vector such that $\mathbf{U} \sim C$. Then,

$$\widetilde{C}(\mathbf{u}) = \overline{C}(1 - u_1, 1 - u_2, \dots, 1 - u_d),$$

where $\overline{C}(\mathbf{u}) = \mathbb{P}(U_1 > u_1, U_2 > u_2, \dots, U_d > u_d)$ is the survival function associated with C, explicitly given by

$$\overline{C}(\mathbf{u}) = 1 + \sum_{k=1}^{d} (-1)^k \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} C_{i_1 i_2 \cdots i_k}(u_{i_1}, u_{i_2}, \dots, u_{i_k}),$$

with $C_{i_1 i_2 \cdots, i_k}$ denoting the marginal of C related to $(i_1, i_2, \cdots, i_k)$.

Singular and absolutely continuous components

For simplicity's sake, we consider here only the case $d = 2$.
Every copula $C \in \mathcal{C}_2$ may be expressed in the form

$$C = C_{ac} + C_s$$

where C_{ac} is *absolutely continuous* and C_s is *singular*.

For an absolutely continuous copula C one has a density c such that

$$C(u,v) = \int_{\mathbb{I}^2} c(s,t)\, ds\, dt = \int_0^1 ds \int_0^1 c(s,t)\, dt$$

The density c is found by differentiation

$$c(u,v) = D_1\, D_2\, C(u,v) = \frac{\partial^2 C(u,v)}{\partial u\, \partial v} \qquad a.e.$$

Singular and absolutely continuous components–2

The presence of a singular component in a copula often causes analytical difficulties. Nevertheless, there are specific applications in which this presence is actually a useful feature; for instance, in default models described by two random variables X and Y, the fact that the event $\{X = Y\}$ may have non–zero probability implies, on the one hand, the existence of a singular component in their copula, and, on the other hand, the possibility of joint defaults of X and Y.

A special case

Notice, however, that, as a consequence of the Lipschitz condition, for every copula $C \in \mathcal{C}_2$ and for every $v \in \mathbb{I}$, both functions $t \mapsto C(t, v)$ and $t \mapsto C(v, t)$ are absolutely continuous so that

$$C(t, v) = \int_0^t c_{1,v}(s)\, ds \quad \text{and} \quad C(v, t) = \int_0^t c_{2,v}(s)\, ds$$

This latter representation has so far found no application.
Notice also that it possible to prove that, for a 2–copula C,

$$D_1 D_2 C = D_2 D_1 C \qquad a.e.$$

Examples–1

Both the copulæ W_2 and M_2 are singular:

- M_2 uniformly spreads the probability mass on the main diagonal $v = u$ $(u \in \mathbb{I})$ of the unit square;
- W_2 uniformly spreads the probability mass on the opposite diagonal $v = -u$ $(u \in \mathbb{I})$ of the unit square.

The product copula $\Pi_2(u, v) := u\, v$ is absolutely continuous and its density π is given by

$$\pi(u, v) = 1_{\mathbb{I}^2}(u, v)$$

Rank–invariant property

Theorem

Let $\mathbf{X} = (X_1, \dots, X_d)$ *be a r.v. with continuous d.f.* F*, univariate marginals* F_1*,* $F_2, \dots,$ F_d*, and copula* C*. Let* $T_1, \dots, T_d$ *be strictly increasing functions from* $\mathbb{R}$ *to* $\mathbb{R}$*. Then* C *is also the copula of the r.v.* $(T_1(X_1), \dots, T_d(X_d))$*.*

> *the study of rank statistics – insofar as it is the study of properties invariant under such transformations – may be characterized as the study of copulas and copula-invariant properties.*

(Schweizer & Wolff, 1981)

Independence

Theorem

Let $(X_1, X_2, \dots, X_d)$ *be a r.v. with continuous joint d.f.* F *and univariate marginals* $F_1, \dots, F_d$. *Then the copula of* $(X_1, \dots, X_d)$ *is* Π_d *if, and only if,* $X_1, \dots, X_d$ *are independent.*

Comonotonicity and countermonotonicity

Theorem

Let $(X_1, X_2, \ldots, X_d)$ be a r.v. with continuous joint d.f. F and univariate marginals $F_1, \ldots, F_d$. Then the copula of $(X_1, \ldots, X_d)$ is M_d if, and only if, there exists a r.v. Z and increasing functions $T_1, \ldots, T_d$ such that $\mathbf{X} = (T_1(Z), \ldots, T_d(Z))$ almost surely.

Theorem

Let (X_1, X_2) be a r.v. with continuous d.f. F and univariate marginals F_1, F_2. Then (X_1, X_2) has copula W_2 if, and only if, for some strictly decreasing function T, $X_2 = T(X_1)$ almost surely.

Stochastic measures

Definition

A measure μ on the measurable space $(\mathbb{I}^d, \mathcal{B}(\mathbb{I}^d))$ will said to be stochastic if, for every Borel set A and for every $j \in \{1, 2, \dots, d\}$,

$$\mu(\underbrace{\mathbb{I} \times \cdots \times \mathbb{I}}_{j-1} \times A \times \mathbb{I} \times \cdots \times \mathbb{I}) = \lambda(A),$$

where λ denotes the (restriction to $\mathcal{B}(\mathbb{I})$ of the) Lebesgue measure.

Copulæ and stochastic measures

Theorem

Every copula $C \in \mathcal{C}_d$ *induces a stochastic measure* μ_C *on the measurable space* $(\mathbb{I}^d, \mathcal{B}(\mathbb{I}^d))$ *defined on the rectangles* $R = [\mathbf{a}, \mathbf{b}]$ *contained in* $\mathbb{I}^d$*, by*

$$\mu_C(R) := V_C\left([\mathbf{a}, \mathbf{b}]\right).$$

Conversely, to every stochastic measure μ *on* $(\mathbb{I}^d, \mathcal{B}(\mathbb{I}^d))$ *there corresponds a unique copula* $C_\mu \in \mathcal{C}_d$ *defined by*

$$C_\mu(\mathbf{u}) := \mu\left([\mathbf{0}, \mathbf{u}]\right).$$

Markov operators

Definition

Given two probability spaces $(\Omega_1, \mathcal{F}_2, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$, a linear operator $T : L^\infty(\Omega_1) \to L^\infty(\Omega_2)$ is said to be a Markov operator if

- T is positive, viz. $Tf \geq 0$ whenever $f \geq 0$;
- $T1 = 1$ (here 1 denotes the constant function $f \equiv 1$);
- $\mathbb{E}_2(Tf) = \mathbb{E}_1(f)$ for every function $f \in L^\infty(\Omega_1)$ ($\mathbb{E}_j$ denotes the expectation in the probability space $(\Omega_j, \mathcal{F}_j, \mathbb{P}_j)$ $(j = 1, 2)$)

Theorem

Every Markov operator $T : L^\infty(\Omega_1) \to L^\infty(\Omega_2)$ has an extension to a bounded operator $T : L^p(\Omega_1) \to L^p(\Omega_2)$ for every $p \geq 1$.

Copulæ and Markov operators

Theorem

For every copula $C \in \mathcal{C}_2$ *the operator* T_C *defined on* $L^1(\mathbb{I})$ *via*

$$(T_C f)(x) := \frac{d}{dx} \int_0^1 D_2 C(x,t) \, f(t) \, dt$$

is a Markov operator on $L^\infty(\mathbb{I})$.
Conversely, for every Markov operator T *on* $L^1(\mathbb{I})$ *the function* C_T *defined on* $\mathbb{I}^2$ *via*

$$C_T(x,y) := \int_0^x \left(T 1_{[0,y]}\right)(s) \, ds$$

is a 2*–copula.*

Examples

$$\left(T_{W_2} f\right)(x) = f(1-x)$$
$$\left(T_{M_2} f\right)(x) = f(x)$$
$$\left(T_{\Pi_2} f\right)(x) = \int_0^1 f \, d\lambda$$

Theorem

For the adjoint $(T_C)^\dagger$ of the Markov operator T_C in the space L^p with $p \in \,]1, +\infty[$ one has $(T_C)^\dagger = T_{C^T}$, where the transpose C^T of the copula C is defined by $C^T(x, y) := C(y, x)$.

The extension to the case $d > 2$

For $d > 2$, consider the factorization $\mathbb{I}^d = \mathbb{I}^p \times \mathbb{I}^q$, where $d = p + q$. While for $d = 2$ there is only one possible factorization, $p = 1$ and $q = 1$, this factorization is not unique when $d > 2$.
Let $C \in \mathcal{C}_d$ be given; it induces a probability measure μ_C on $(\mathbb{I}^d, \mathcal{B}(\mathbb{I}^d))$. Denote the marginals of μ_C on $(\mathbb{I}^p, \mathcal{B}(\mathbb{I}^p))$ and on $(\mathbb{I}^q, \mathcal{B}(\mathbb{I}^q))$ by μ_p and μ_q, respectively.
Given a decomposition $d = p + q$, there is a unique Markov operator $T : L^\infty(\mathbb{I}^p) \to L^\infty(\mathbb{I}^q)$ associated with μ_C and, hence, with the copula C. Therefore, to every copula $C \in \mathcal{C}_d$ there correspond as many Markov operators as there are solutions in natural numbers p and q of the Diophantine equation $p + q = d$. Since the number of these solutions is $d - 1$, there are $d - 1$ possible different Markov operators corresponding to a d–copula when $d \geq 3$.

An introduction to Copulas

Carlo Sempi

Dipartimento di Matematica “Ennio De Giorgi”
Università del Salento
Lecce, Italy
carlo.sempi@unisalento.it

The 33rd Finnish Summer School on Probability Theory and Statistics, June 6th–10th, 2011

Outline

Measure–preserving transformations

$(\Omega, \mathcal{F}, \mu)$ and $(\Omega', \mathcal{F}', \nu)$ — two measure spaces.
$f : \Omega \to \Omega'$ is a measure–preserving transformations (=MPT) if

- $\forall B \in \mathcal{F}' \quad f^{-1}(B) \in \mathcal{F}$
- $\forall B \in \mathcal{F}' \quad \mu\left(f^{-1}(B)\right) = \nu(B)$

From now on $(\Omega, \mathcal{F}, \mu) = (\Omega', \mathcal{F}', \nu) = (\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$
$\mathcal{B}(\mathbb{I})$ — the Borel sets $\mathbb{I}$
λ — the (restriction) of Lebesgue measure to $\mathcal{B}(\mathbb{I})$.

Copulæ and MPT's

Theorem

If f_1, $f_2, \ldots$, f_d are MPT's, the function $C_{f_1,f_2,\ldots,f_d} : \mathbb{I}^n \to \mathbb{I}$ defined by

$$C_{f_1,f_2,\ldots,f_d}(x_1, x_2, \ldots, x_d) := \lambda\left(f_1^{-1}[0, x_1] \cap \cdots \cap f_d^{-1}[0, x_d]\right)$$

is a copula. Conversely, for every d–copula $C \in \mathcal{C}_d$, there exist d MPT's f_1, $f_2, \ldots f_d$ such that

$$C = C_{f_1,f_2,\ldots,f_d}.$$

This representation is not unique: if φ is another MPT on $\mathbb{I}$, then

$$C_{f_1,f_2,\ldots,f_d} = C_{f_1\circ\varphi,f_2\circ\varphi,\ldots,f_d\circ\varphi}.$$

Special MPT's

A transformation f is said to be *ergodic* if, for all measurable sets A and B, one has

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu\left(f^{-k} A \cap B\right) = \mu(A)\,\mu(B);$$

f is said to be *strongly mixing* iIf f satisfies the stronger property

$$\lim_{n\to+\infty} \mu\left(f^{-n} A \cap B\right) = \mu(A)\,\mu(B)$$

Two corollaries

Corollary

If f is strongly mixing, then, for all $x, y \in [0, 1]$,

$$\lim_{n\to+\infty} C_{f^n,g}(x, y) = xy = \Pi_2(x, y).$$

Corollary

If f is ergodic, then, for all $x, y \in [0, 1]$,

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} C_{f^j,g}(x, y) = xy = \Pi_2(x, y).$$

Two examples

For the copula M_2 one has

$$\begin{aligned}\lambda\left(f^{-1}[0,x]\cap f^{-1}[0,y]\right) &= \lambda\left(f^{-1}([0,x]\cap[0,y])\right)\\ &= \lambda([0,x]\cap[0,y]) = \min\{x,y\} = M_2(x,y).\end{aligned}$$

for *every* measure–preserving transformation f.
As for the copula W_2, recall that it concentrates all the probability mass uniformly on the the diagonal $\varphi(t) = 1 - t$ of the unit square. In this case $\varphi = \varphi^{-1}$, so that

$$\begin{aligned}\lambda\left(\varphi^{-1}[0,x]\cap[0,y]\right) &= \lambda([1-x,1]\cap[0,y])\\ &= \begin{cases}0, & \text{if } x \le 1-y,\\ x+y-1, & \text{if } x > 1-y;\end{cases}\end{aligned}$$

therefore

$$W_2(x,y) = \lambda\left(\varphi^{-1}[0,x]\cap[0,y]\right).$$

The independence copula

Theorem

Let f and g be measure–preserving transformations. The following conditions are equivalent for $C_{f,g} \in \mathcal{C}_2$:

(a) $C_{f,g} = \Pi_2$

(b) *f and g, when regarded as random variables on the standard probability space $(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$, are independent.*

Patchwork

- An at most countable family $(S_i)_{i \in I}$ of closed and connected subsets of $\mathbb{I}^2$

$$S_i \cap S_j \subset \partial S_i \cap \partial S_j$$

- C – a copula
- a continuous function $F_i : S_i \to \mathbb{I}^2$ that is isotone in each place and agrees with C (called background) on the the boundary ∂S_i of S_i, namely $F_i(u, v) = C(u, v)$ for every $(u, v) \in \partial S_i$

The function $F : \mathbb{I}^2 \to \mathbb{I}$

$$F(u, v) := \begin{cases} F_i(u, v)\,, & (u, v) \in S_i\,, \\ C(u, v)\,, & \text{elsewhere,} \end{cases}$$

is called the patchwork of $(F_i)_{i \in I}$ into C.

Patchwork copulæ

Theorem

Given the family $(R_i)_{i \in I}$ *of rectangles, for the patchwork of the family* $(F_i)_{i \in I}$ *into the copula* C *the following statements are equivalent:*

(a) F *is a copula;*

(b) *for every* $i \in I$, F_i *is* 2*–increasing on* R_i *and coincides with* C *on the boundary* ∂R_i *of* R_i.

Ordinal sums

- J be a finite or countable subset of the natural numbers $\mathbb{N}$
- $(]a_k, b_k[)_{k\in J}$ be a family of sub–intervals of $\mathbb{I}$ indexed by J. It is required that any two of them have at most an endpoint in common.
- $(C_k)_{k\in J}$ a family of copulas also indexed by J

Definition

The ordinal sum C of $(C_k)_{k\in J}$ with respect to family of intervals $(]a_k, b_k[)_{k\in J}$ is defined, for all $\mathbf{u} = (u_1, u_2) \in \mathbb{I}^2$ by

$$C(u,v) := \begin{cases} a_k + (b_k - a_k)\, C_k\left(\frac{u-a_k}{b_k-a_k}, \frac{v-a_k}{b_k-a_k}\right), & (u,v) \in [a_k, b_k]^2, \\ \min\{u,v\}, & \text{elsewhere.} \end{cases}$$

Ordinal sums–2

Theorem

The ordinal sum of the family of copulas $(C_k)_{k\in J}$ with respect to the family of intervals $(]a_k, b_k[)_{k\in J}$ is a copula.

An ordinal sum is a special case of the construction of patchwork copulas; it suffices to choose

- the copula M_2 as the background copula;
- $S_k =]a_k, b_k[\times]a_k, b_k[$ for every $k \in J$;
- for every $k \in J$, F_k is a version of the copula C_k rescaled in such a way as to meet the requirements of a patchwork

W_2–ordinal sums

Theorem

Let $C \in \mathcal{C}_2$ be a copula for which there exists $x_0 \in \,]0,1[$ such that $C(x_0, 1-x_0) = 0$. Then there exist two 2–copulæ $C_1 \in \mathcal{C}_2$ and $C_2 \in \mathcal{C}_2$ such that

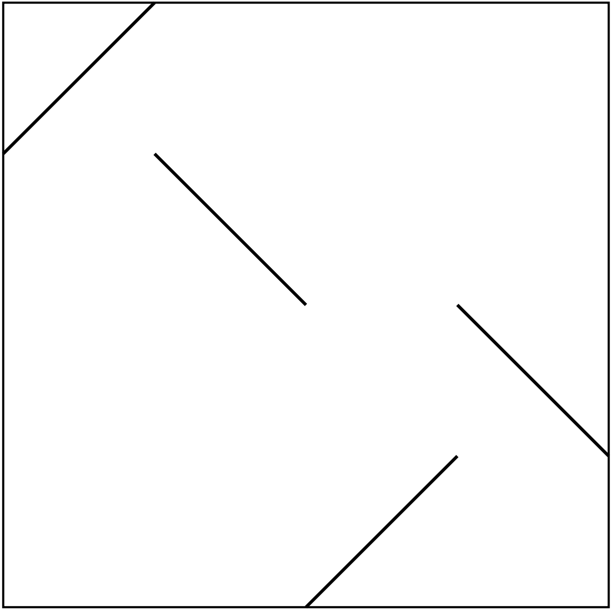
$$C(u,v) = \begin{cases} x_0\, C_1\left(\frac{u}{x_0}, \frac{x_0+v-1}{x_0}\right) & (u,v) \in [0,x_0] \times [1-x_0, 1] \\ (1-x_0)\, C_2\left(\frac{u-x_0}{1-x_0}, \frac{v}{1-x_0}\right), & (u,v) \in [x_0,1] \times [0,1-x_0] \\ W_2(u,v), & \text{elsewhere.} \end{cases}$$

Shuffles of Min

A copula is said to be a shuffle of Min it is obtained through the following procedure:

- the probability mass is placed on the support of the copula M_2, namely on the main diagonal of the unit square;
- then the unit square is cut into a finite number of vertical strips;
- these vertical strips are permuted ("shuffled") and, possibly, some of them are flipped about their vertical axes of symmetry;
- finally the vertical strips are reassembled to form the unit square again;
- to the probability mass thus obtained there corresponds a unique copula C, which is a shuffle of Min.

Shuffles of Min were introduced in (Mikusiński et al. (1992)).



A different presentation

Two continuous random variables X and Y have a shuffle of Min C as their copula is if, and only if, one of them is an invertible piecewise linear function of the other one.

The set of Shuffles of Min is dense in $\mathcal{C}_2$.

Density of the shuffles

Theorem

Let X and Y be continuous random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let F and G be their marginal d.f.'s and H their joint d.f.. Then, for every $\epsilon > 0$ there exist two random variables X_ϵ and Y_ϵ on the same probability space and a piecewise linear function $\varphi : \mathbb{R} \to \mathbb{R}$ such that

(a) $Y_\epsilon = \varphi \circ X_\epsilon$

(b) $F_\epsilon := F_{X_\epsilon} = F$ and $G_\epsilon := F_{Y_\epsilon} = G$

(c) $\|H - H_\epsilon\|_\infty < \epsilon$

where H_ϵ is the joint d.f. of X_ϵ and Y_ϵ, and $\| \cdot \|_\infty$ denotes the L^∞–norm on $\overline{\mathbb{R}}^2$.

A surprising consequence

The last result has a surprising consequence. Let X and Y be independent (and continuous) random variables on the same probability space, let F and G be their marginal d.f.'s and $H = F \otimes G$ their joint d.f.. Then, according to the previous theorem, it is possible to construct two sequences (X_n) and (Y_n) of random variables such that, for every $n \in \mathbb{N}$, their joint d.f. H_n approximates H to within $1/n$ in the L^∞–norm, but Y_n is almost surely a (piecewise linear) function of X_n.

A generalization; preliminaries–1

- $(\Omega, \mathcal{F}, \mu)$ – a measure space
- $(\Omega_1, \mathcal{F}_1)$ – a measurable space
- $\varphi : \Omega \to \Omega_1$ – a measurable function
- $\mathcal{T}$ – the set of all measure–preserving transformations of $(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$
- $\mathcal{T}_p$ – the set of all measure-preserving permutations (automorphisms) of this space
- image measure of μ under φ

$$\mu_\varphi(A) = (\mu \circ \varphi)(A) = \mu\left(\varphi^{-1} A\right) \quad (A \in \mathcal{F}_1)$$

$\mathcal{T}$ equipped with the composition of mappings is a semigroup and $\mathcal{T}_p$ is a subgroup of $\mathcal{T}$.

Interval exchange transformations

- $\{J_{1,i}\}$ $(i = 1, 2, \dots, n)$ – partition of $\mathbb{I}$ into the non–degenerate intervals $J_{1,i} = [a_{1,i}, b_{1,i}[$ and the singleton $J_{1,n} = \{1\}$.
- $\{J_{2,i}\}$ $(i = 1, 2, \dots, n)$ – another such partition such that, $\lambda(J_{1,i}) = \lambda(J_{2,i})$
- the interval exchange transformation

$$T(x) = \begin{cases} x - a_{1,1} + a_{2,1}, & \text{if } x \in J_{1,i}, \\ \lambda\left(\left(\mathbb{I} \setminus \bigcup_{i=1}^{n} J_{1,i}\right) \cap [0, x]\right) + \sum_{i=1}^{n} (b_{2,i} - a_{2,a})\, \mathbf{1}_{[a_2,1]}(x) \\ \qquad \text{otherwise,} \end{cases}$$

A mapping on $\mathbb{I}^2$

Given $T : \mathbb{I} \to \mathbb{I}$ define $S_T : \mathbb{I}^2 \to \mathbb{I}^2$ via

$$S_T(u, v) := (T(u), v) . \qquad ((u, v) \in \mathbb{I}^2)$$

- J – a (possibly degenerate) interval in $\mathbb{I}$
- the (vertical) strip $J \times \mathbb{I}$
- the partition of the unit square $\mathbb{I}^2$ into possibly infinitely many, vertical strips.

Generalized shuffling

A shuffling of a strip partition $\{J_i \times \mathbb{I}\}_{i \in \mathcal{I}}$ (card $\mathcal{I} \leq \aleph_0$) is any permutation S of the unit square such that

(1_{Sh}) admits the representation $S = S_T$ for some $T : \mathbb{I} \to \mathbb{I}$

(2_{Sh}) is measure–preserving on the space $(\mathbb{I}^2, \mathcal{B}(\mathbb{I}^2), \lambda_2)$

(3_{Sh}) the restriction $S|_{J_i \times \mathbb{I}}$ of S to every strip $J_i \times \mathbb{I}$ is continuous with respect to the standard product topology on $\mathbb{I}^2$

Generalized shuffling–2

Intuitively, shuffling is just a reordering of the strips. This feature is captured by the condition (1_{Sh}), which represents the shuffling by a single transformation T of the unit interval. In particular, S_T is a permutation of $\mathbb{I}^2$ if, and only if, T is a permutation of $\mathbb{I}$. Because of (2_{Sh}) the single strips maintain their measure after shuffling. Finally, condition (3_{Sh}) is just a technical tool for ensuring that, during shuffling, the integrity of strips is preserved.

Shuffles: the new characterization

Lemma

Consider the image measure of a doubly stochastic measure μ under S_T. Then the following statements are equivalent:

(a) μ_{S_T} *is doubly stochastic*

(b) T *is in* $\mathcal{T}$.

Theorem

The following statements are equivalent:

(a) *a copula* $C \in \mathcal{C}_2$ *is a shuffle of Min;*

(b) *there exists a piece–wise continuous* $T \in \mathcal{T}$ *such that* $\mu_C = \mu_{M_2} \circ S_T^{-1}$

Shuffles: the new definition

Definition

A copula $C \in \mathcal{C}_2$ is a *generalized shuffle of Min* if $\mu_C = \mu_{M_2} \circ S_T^{-1}$ for some $T \in \mathcal{T}$. Such a shuffle of Min is denoted by M_T.

In this definition, T is allowed to have *countably* many discontinuity points, which is a quite natural generalization of the original notion of shuffle of Min.

Shuffling an arbitrary copula

Definition

Let $C \in \mathcal{C}_2$ be a copula. A copula A is a shuffle of C if there exists $T \in \mathcal{T}$ such that $\mu_A = \mu_C \circ S_T^{-1}$. In this case, A is also called the T–shuffle of C and denoted by C_T.

If a copula C is represented by means of two measure–preserving transformations f and g, $C_{f,g}$, then

$$(C_{f,g})_T = C_{T \circ f, g}$$

Orbits

The mapping which assigns to every $T \in \mathcal{T}$ and to every copula $C \in \mathcal{C}_2$ the corresponding shuffle C_T defines an action of the group $\mathcal{T}$ on the set of all copulas. The orbit of a copula C with respect to this action is the set $\mathcal{T}(C) = \{C_T \mid T \in \mathcal{T}\}$ constituted by all shuffles of C. The general theory of group actions guarantees that the classes of type $\mathcal{T}(C)$ form a partition of the set of all copulas. The orbit of a copula is exactly the collection of all its shuffles.

Theorem

For a copula $C \in \mathcal{C}_2$ the following statements are equivalent:

(a) $C = \Pi_2$;

(b) $\mathcal{T}(C) = \{C\}$.

More on shuffles

Theorem

If $C \in \mathcal{C}_2$ is absolutely continuous then so are all its shuffles.

Theorem

Every copula $C \in \mathcal{C}_2$ other than Π_2 has a non–exchangeable shuffle.

Theorem

For every copula $C \in \mathcal{C}_2$, the independence copula Π_2 can be approximated uniformly by elements of $\mathcal{T}(C)$.

Generators

A function $\varphi : \mathbb{R}_+ \to \mathbb{I}$ is said to be an (outer additive) generator if it is continuous, decreasing and $\varphi(0) = 1$, $\lim_{t\to+\infty} \varphi(t) = 0$ and is strictly decreasing on $[0, t_0]$, where $t_0 := \inf\{t > 0 : \varphi(t) = 0\}$. If the function φ is invertible, or, equivalently, strictly decreasing on $\mathbb{R}_+$, then the generator is said to be strict. If φ is strict, then $\varphi(t) > 0$ for every $t > 0$ (and $\lim_{t\to+\infty} \varphi(t) = 0$).

Archimedan copulæ

A copula $C \in \mathcal{C}_d$ is said to be Achimedean if a generator φ exists such that

$$C(\mathbf{u}) = \varphi\left(\varphi^{(-1)}(u_1) + \varphi^{(-1)}(u_2) + \cdots + \varphi^{(-1)}(u_d)\right) \qquad \mathbf{u} \in \mathbb{I}^d.$$

Such a copula will be denoted by C_φ
When φ is strict the copula C_φ is said to be strict; in this case, C_φ has the representation

$$C_\varphi(\mathbf{u}) = \varphi\left(\varphi^{-1}(u_1) + \cdots + \varphi^{-1}(u_d)\right).$$

d–monotone functions

A function $f :]a, b[\to \mathbb{R}$ is called d–monotone in $]a, b[$, where $-\infty \leq a < b \leq +\infty$ if

- it is differentiable up to order $d - 2$;
- for every $x \in]a, b[$, its derivatives satisfy the inequalities

$$(-1)^k f^{(k)}(x) \geq 0, \qquad (k = 0, 1, \dots, d - 2)$$

- $(-1)^{d-2} f^{(d-2)}$ is decreasing and convex in $]a, b[$

f is 2–monotone function iff it is decreasing and convex. If f has derivatives of every order and if

$$(-1)^k f^{(k)}(x) \geq 0,$$

for every $x \in]a, b[$ and for every $k \in \mathbb{Z}_+$ is said to be completely monotonic.

Characterization of Archimedean copulas

Theorem

(McNeil & Nešlehová) Let $\varphi : \mathbb{R}_+ \to \mathbb{I}$ *be a generator. Then the following statements are equivalent:*

(a) φ *is* d*–monotone on* $]0, +\infty[$;

(b) $C_\varphi(\mathbf{u}) := \varphi\left(\varphi^{(-1)}(u_1) + \cdots + \varphi^{(-1)}(u_d)\right)$ *is a* d*–copula.*

Corollary

Let $\varphi : \mathbb{R}_+ \to \mathbb{I}$ *be a generator. Then the following statements are equivalent:*

(a) φ *is completely monotone on* $]0, +\infty[$

(b) $C_\varphi : \mathbb{I}^d \to \mathbb{I}$ *is a* d*–copula for every* $d \geq 2$

Examples

The copula Π_2 is Archimedean: take $\varphi(t) = e^{-t}$; since $\lim_{t\to+\infty} \varphi(t) = 0$ and $\varphi(t) > 0$ for every $t > 0$, φ is strict; then $\varphi^{-1}(t) = -\ln t$ and

$$\varphi\left(\varphi^{-1}(u) + \varphi^{-1}(v)\right) = \exp\left(-\left(-\ln u - \ln v\right)\right) = uv = \Pi_2(u, v).$$

Also W_2 is Archimedean; take $\varphi(t) := \max\{1 - t, 0\}$. Since $\varphi(1) = 0$, φ is not strict. Its quasi–inverse is $\varphi^{(-1)}(t) = 1 - t$. On the contrary, the upper Fréchet–Hoeffding bound M_2 is not Archimedean.

The Gumbel–Hougaard family

$$C_{\theta}^{\mathbf{GH}}(\mathbf{u}) = \exp\left(-\left(\sum_{i=1}^{d}(-\log(u_i))^{\theta}\right)^{1/\theta}\right)$$

where $\theta \geq 1$. For $\theta = 1$ we obtain the independence copula as a special case, and the limit of $C_{\theta}^{\mathbf{GH}}$ for $\theta \to +\infty$ is the comonotonicity copula. The Archimedean generator of this family is given by $\varphi(t) = \exp\left(-t^{1/\theta}\right)$. Each member of this class is absolutely continuous.

The Mardia–Takahasi–Clayton family

The standard expression for members of this family of d–copulas is

$$C_\theta^{\mathbf{MTC}}(u,v) = \max\left\{\left(\sum_{i=1}^{d} u_i^{-\theta} - (d-1)\right)^{-1/\theta}, 0\right\}$$

where $\theta \geq \frac{-1}{d-1}$, $\theta \neq 0$. The limiting case $\theta = 0$ corresponds to the independence copula.
The Archimedean generator of this family is given by

$$\varphi_\theta(t) = \left(\max\{1+\theta t, 0\}\right)^{-1/\theta}.$$

For every d–dimensional Archimedean copula C and for every $\mathbf{u} \in \mathbb{I}^d$, $C_\theta \theta_L \mathbf{u} \leq C(\mathbf{u})$ for $\theta_L = -\frac{1}{d-1}$.

Frank's family

$$C_\theta^{\mathbf{Fr}}(\mathbf{u}) = -\frac{1}{\theta} \log\left(1 + \frac{\prod_{i=1}^{d}\left(e^{-\theta u_i} - 1\right)}{\left(e^{-\theta} = 1\right)^{d-1}}\right),$$

where $\theta > 0$. The limiting case $\theta = 0$ corresponds to Π_d. For the case $d = 2$, the parameter θ can be extended also to the case $\theta < 0$.

Copulas of this type have been introduced by Frank in relation with a problem about associative functions on $\mathbb{I}$. They are absolutely continuous.

The Archimedean generator is given by

$$\varphi_\theta(t) = -\frac{1}{\theta} \log\left(1 - \left(1 - e^{-\theta}\right) e^{-t}\right)$$

EFGM copulæ–1

For $d \geq 2$ let $\mathcal{S}$ be the class of all subsets of $\{1, 2, \dots, d\}$ having at least 2 elements; $\mathcal{S}$ contains $2^d - d - 1$ elements. To each $S \in \mathcal{S}$, we associate a real number α_S, with the convention that, when $S = \{i_1, i_2, \dots, i_k\}$, $\alpha_S = \alpha_{i_1 i_2 \dots i_k}$.
An EFGM copula can be expressed in the following form:

$$C_d^{\mathbf{EFGM}}(\mathbf{u}) = \prod_{i=1}^{d} u_i \left(1 + \sum_{S \in \mathcal{S}} \alpha_S \prod_{j \in S} (1 - u_j)\right),$$

for suitable values of the α_S's.
For the bivariate case EFGM copulæ have the following expression:

$$C_2^{\mathbf{EFGM}} u_1, u_2 = u_1 u_2 \left(1 + \alpha_{12}(1 - u_1)(1 - u_2)\right),$$

EFGM copulæ–2

EFGM copulæ are absolutely continuous with density

$$c_d^{\mathbf{EFGM}}(\mathbf{u}) = 1 + \sum_{S \in \mathcal{S}} \alpha_S \prod_{j \in S} (1 - 2u_j).$$

As a consequence, the parameters α_S's have to satisfy the following inequality

$$1 + \sum_{S \in \mathcal{S}} \alpha_S \prod_{j \in S} \xi_j \geq 0$$

for every $\xi_j \in \{-1, 1\}$. In particular, $|\alpha_S| \leq 1$.

A necessary detour: associativity

Definition

A binary operation T on $\mathbb{I}$ is said to be associative if, for all s, t and u in $\mathbb{I}$,

$$T\left(T(s,t),u\right) = T\left(s, T(t,u)\right)$$

Definition

The T–powers of an element $t \in \mathbb{I}$ under the associative function T are defined recursively by

$$t^1 := t \qquad \text{and} \qquad \forall n \in \mathbb{N} \quad t^{n+1} := T\left(t^n, t\right),$$

t–norms

Definition

A triangular norm, or, briefly, a t–norm T is a function $T : \mathbb{I}^2 \to \mathbb{I}$ that is associative, commutative, isotone in each place, viz., both the functions

$$\mathbb{I} \ni \; t \mapsto T(t,s) \qquad \text{and} \qquad \mathbb{I} \ni \; t \mapsto T(s,t)$$

are isotone for every $s \in \mathbb{I}$ and such that $T(1,t) = t$ for every $t \in \mathbb{I}$.

Definition

A t–norm T is said to be Archimedean if, for all s and t in $]0,1[$, there is $n \in \mathbb{N}$ such that $s^n < t$.

Copulæ and t–norms

Theorem

For a t–norm T the following statements are equivalent:

(a) T *is a* 2*–copula;*

(b) T *satisfies the Lipschitz condition:*

$$T(x', y) - T(x, y) \leq x' - x \qquad x, x', y \in \mathbb{I} \quad x \leq x'$$

Theorem

For an Archimedean t–norm T, which has φ as an outer additive generator, the following statements are equivalent:

(a) T *is a* 2*–copula;*

(b) *either* φ *or* $\varphi^{(-1)}$ *is convex.*

Two important concepts

Definition

An element $a \in \,]0,1[$ is said to be a nilpotent element of the t–norm T if there exists $n \in \mathbb{N}$ such that $a_T^{(n)} = 0$.

Definition

A t–norm T is said to be strict if it is continuous on $\mathbb{I}^2$ and is strictly increasing on $]0,1[$; it is said to be nilpotent if it is continuous on $\mathbb{I}^2$ and every $a \in \,]0,1[$ is nilpotent.

The t–norm $\Pi_2(u,v) := uv$ is strict, while $W_2(u,v) := \max\{u+v-1, 0\}$ is nilpotent.

$$\forall\, a \in \,]0,1[\qquad a_{W_2}^n = \max\{na - (n-1), 0\},$$

so that $a_{W_2}^n = 0$ for $n \geq 1/(1-a)$.

Representation of t–norms

Under mild conditions the t–norm T has the following representation

$$T(x, y) = \varphi\left(\varphi^{(-1)}(x) + \varphi^{(-1)}(y)\right) \qquad x, y \in \mathbb{I},$$

where $\varphi : \overline{\mathbb{R}}_+ \to \mathbb{I}$ is continuous, decreasing and $\varphi(0) = 1$, while $\varphi^{(-1)} : \mathbb{I} \to \overline{\mathbb{R}}_+$ is a quasi–inverse of φ that is continuous, strictly decreasing on $\mathbb{I}$ and such that $\varphi^{(-1)}(1) = 0$

Isomorphisms of generators

$\varphi : \mathbb{R}_+ \to \mathbb{I}$ — an Archimedean generator
ψ — a stricly increasing bijection on $\mathbb{I}$, in particular, $\psi(0) = 0$ and $\psi(1) = 1$. Then $\psi \circ \varphi$ is also a generator.
If T_φ is the Archimedean t–norm generated by the outer generator φ, then, as is immediately checked, $\psi \circ \varphi$ is the generator of the t–norm

$$T_{\psi\circ\varphi}(u, v) = (\psi \circ \varphi) \left(\varphi^{(-1)} \circ \psi^{-1}(u) + \varphi^{(-1)} \circ \psi^{-1}(v) \right) = \psi \left(T_\varphi \left(\psi^{-1}(u), \psi^{-1}(v) \right) \right) .$$

Isomorphisms of generators–2

Definition

Two generators φ_1 and φ_2 are said to be isomorphic if there exists a strictly increasing bijection $\psi : \mathbb{I} \to \mathbb{I}$ such that $\varphi_2 = \psi \circ \varphi_1$.
Two t–norms T_1 and T_2 are said to be isomorphic if there exists a strictly increasing bijection $\psi : \mathbb{I} \to \mathbb{I}$ such that, for all u and v in $\mathbb{I}$,

$$T_2(u, v) = \psi\left(T_1\left(\psi^{-1}(u), \psi^{-1}(v)\right)\right).$$

Two results on t–norms

Theorem

For a function $T : \mathbb{I}^2 \to \mathbb{I}$, the following statements are equivalent:

(a) T *is a strict t–norm;*

(b) T *is isomorphic to* Π_2.

Theorem

For a function $T : \mathbb{I}^2 \to \mathbb{I}$, the following statements are equivalent:

(a) T *is a nilpotent t–norm;*

(b) T *is isomorphic to* W_2.

Isomorphisms for copulas–1

Theorem

For an Archimedean 2*–copula* $C \in \mathcal{C}_2$*, the following statements are equivalent:*

(a) C *is strict;*

(b) C *is isomorphic to* Π_2*;*

(c) *every additive generator* φ *of* C *is isomorphic to* $\varphi_{\Pi_2}(t) = e^{-t}$ $(t \in \mathbb{R}_+)$

Isomorphisms for copulas–2

Theorem

For an Archimedean 2*–copula* $C \in \mathcal{C}_2$*, the following statements are equivalent:*

(a) C *is nilpotent;*

(b) C *is isomorphic to* W_2*;*

(c) *every outer additive generator* φ *of* C *is isomorphic to* $\varphi_{W_2}(t) = \max\{1 - t, 0\}$ $(t \in \mathbb{R}_+)$

An example

The copula

$$C(u, v) := \frac{uv}{u + v - uv}$$

usually denoted by $\Pi/(\Sigma - \Pi)$ in the literature is strict; its generator is

$$\varphi(t) = \frac{1}{1 + t} \qquad (t \in \mathbb{R}_+).$$

The isomorphism with φ_{Π_2} is realized by the function $\psi : \mathbb{I} \to \mathbb{I}$ defined by

$$\psi(s) = \frac{1}{1 - \ln s} .$$

Brownian motion

In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $\{B_t^{(1)} : t \geq 0\}$ and $\{B_t^{(2)} : t \geq 0\}$ be two Brownian motions (=BM's). We explicitly assume that the BM is continuous and consider, for every $t \geq 0$, the random vector

$$B_t := \left(B_t^{(1)}, B_t^{(2)}\right)$$

Then $\{B_t : t \geq 0\}$ defines a stochastic process with values in $\mathbb{R}^2$. The literature deals mainly with the independent case, viz., $B_t^{(1)}$ and $B_t^{(2)}$ are independent for every $t \geq 0$; this is usually called the *two–dimensional* BM.

Distribution functions

For every $t \geq 0$, let $F_t^{(1)}$ and $F_t^{(2)}$ be the (right–continuous) distribution functions (=d.f.'s) of $B_t^{(1)}$ and $B_t^{(2)}$, respectively; thus, for every $x \in \mathbb{R}$,

$$F_t^{(j)}(x) = \mathbb{P}\left(B_t^{(j)} \leq x\right) \qquad (j = 1, 2).$$

Actually, For every $t \geq 0$, $F_t^{(1)}(x) = F_t^{(2)}(x) = \Phi(x/\sqrt{t})$, where Φ is the d.f. of the standard normal distribution $N(0, 1)$.

Coupled BM–1

For every $t \geq 0$, let C_t, which depends on t, be the bivariate copula of the random pair $(B_t^{(1)}, B_t^{(2)})$. Then the d.f. $H_t : \mathbb{R}^2 \to \mathbb{I}$ of the random pair B_t, is given, for all x and y in $\mathbb{R}$, by

$$H_t(x, y) = C_t\left(F_t^{(1)}(x), F_t^{(2)}(y)\right).$$

Since both $B_t^{(1)}$ and $B_t^{(2)}$ are normally distributed the copula C_t is uniquely determined for every $t \geq 0$.

Coupled BM–2

Through an abuse of notation we shall write

$$B_t := C_t\left(B_t^{(1)}, B_t^{(2)}\right)$$

Notice that, in principle, a different copula is allowed for every $t \geq 0$. The process $\{B_t : t \geq 0\}$ will be called the *2–dimensional coupled Brownian motion*.

The traditional two–dimensional BM is included in the picture; in order to recover it, it suffices to choose the independence copula $\Pi_2(u,v) := u\,v \; ((u,v) \in \mathbb{I}^2)$ and set $C_t = \Pi_2$ for every $t \geq 0$

$$H_t(x,y) = F_t^{(1)}(x)\,F_t^{(2)}(y) \qquad ((x,y) \in \mathbb{R}^2).$$

Properties to be studied

The (one–dimensional) BM is the example of a stochastic process that has three properties

- it a Markov process;
- it is a martingale in continuous time;
- it is a Gaussian process.

These three aspects will be examined for a coupled BM.

The Markov property

Since the Markov property for a d–dimensional process $\{X_t : t \geq 0\}$ disregards the dependence relationship of its components at every $t \geq 0$, but is solely concerned with the dependence structure of the random vector X_t at different times, the traditional proof for the ordinary (independent) BM holds for the coupled BM $\{B_t := C_t(B_t^{(1)}, B_t^{(2)}) : t \geq 0\}$. Therefore,

Theorem

A coupled Brownian motion $\{B_t := C_t(B_t^{(1)}, B_t^{(2)}) : t \geq 0\}$ *is a Markov process.*

The coupled BM is a martingale

Theorem

The coupled Brownian motion $\{B_t := C_t(B_t^{(1)}, B_t^{(2)}) : t \geq 0\}$ *is a martingale.*

Gaussian processes

One has first to state what is meant by the expression *Gaussian process* when a stochastic process with values in $\mathbb{R}^2$ is considered. We shall adopt the following definition.

Definition

A stochastic process $\{X_t : t \geq 0\}$ with values in $\mathbb{R}^d$ is said to be *Gaussian* if, for every $n \in \mathbb{N}$, and for every choice of n times $0 \leq t_1 < t_2 < \cdots < t_n$, the random vector $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ has a $(d \times n)$–dimensional normal distribution.

Is a coupled BM a Gaussian process?

Let the copula C_t coincide, for every $t \geq 0$, with M_2, i.e., $M_2(u,v) = \min\{u,v\}$, u and v in $\mathbb{I}$. Then

$$H_t(x,y) = \frac{1}{\sqrt{2\pi\, t}} \min\left\{ \int_{-\infty}^{x} \exp\{-v^2/(2t)\}\, du, \int_{-\infty}^{y} \exp\{-u^2/(2t)\}\, dv \right\} = \Phi\left(\frac{\min\{x,y\}}{\sqrt{t}}\right).$$

A simple calculation shows that

$$\frac{\partial^2 H_t(x,y)}{\partial x\, \partial y} = 0 \qquad a.e.$$

with respect to the Lebesgue measure λ_2, so that H_t is not even absolutely continuous.

Example–2

If the copula C_t is given, for every $t \geq 0$, by W_2, where

$$W_2(u, v) := \max\{u + v - 1, 0\},$$

then the d.f. H_t of B_t is given by

$$H_t(x, y) = \max\left\{\Phi\left(\frac{x}{\sqrt{t}}\right) + \Phi\left(\frac{y}{\sqrt{t}}\right) - 1, 0\right\},$$

which again leads, after simple calculations, to the conclusion that, again, B_t is not even absolutely continuous.

Singular copulæ

The two previous examples represent extreme cases; in fact, since the d.f.'s involved are continuous, the copula of two random variables is M_2 if, and only if, they are comonotone, namely, each of them is an increasing function of the other, while their copula is W_2 if, and only if, they are countermonotone, namely, each of them is a decreasing function of the other. In this sense both examples are the opposite of the independent case, which is characterized by the copula Π_2.
We recall that a copula can be either absolutely continuous or singular or, again, a mixture of the two types. In general, if the copula C is singular, namely the d.f. of a probability measure concentrated on a subset of zero Lebesgue measure λ_2 in the unit square $\mathbb{I}^2$, then also B_t is singular.

The absolutely continuous case

Now let the copula C_t be absolutely continuous with density c_t; a simple calculation shows that B_t is absolutely continuous and that its density is given a.e. by

$$h_t(x, y) = \frac{1}{2\pi\, t} \exp\left(-\frac{x^2 + y^2}{2t}\right) c_t\left(\Phi\left(\frac{x}{\sqrt{t}}\right), \Phi\left(\frac{y}{\sqrt{t}}\right)\right)$$

As a consequence, B_t has a normal law if, and only if, $c_t(u, v) = 1$ for almost all u and v in $\mathbb{I}$; together with the boundary conditions, this implies $C_t(u, v) = u\, v = \Pi_2(u, v)$.

The special position of independence

Theorem

In a coupled Brownian motion

$$\left\{ B_t = C_t \left(B_t^{(1)}, B_t^{(2)} \right) : t \geq 0 \right\},$$

B_t *has a normal law if, and only if,* $C_t = \Pi_2$*, viz., if, and only if, its components* $B_t^{(1)}$ *and* $B_t^{(2)}$ *are independent.*

An introduction to Copulas

Carlo Sempi

Dipartimento di Matematica "Ennio De Giorgi"
Università del Salento
Lecce, Italy
carlo.sempi@unisalento.it

The 33rd Finnish Summer School on Probability Theory and Statistics, June 6th–10th, 2011

Outline

The $*$–product

Definition

Given two copulas A and B in $\mathcal{C}_2$, define a map via

$$(A * B)(x, y) := \int_0^1 D_2A(x, t)\, D_1B(t, y)\, dt\,.$$

Theorem

For all copulas A and B, $A * B$ is a copula, namely $A * B \in \mathcal{C}_2$, or, equivalently, $* : \mathcal{C}_2 \times \mathcal{C}_2 \to \mathcal{C}_2$.

The $*$–product–2

Lemma

For every pair A and B of 2–copulas, one has

$$T_A \circ T_B = T_{A*B} .$$

Continuity in one variable

Theorem

Consider a sequence $(A_n)_{n\in\mathbb{N}}$ of copulas and a copula B. If the sequence (A_n) converges (uniformly) to $A \in \mathcal{C}$, $A_n \to A$ then both

$$A_n * B \xrightarrow[n\to+\infty]{} A * B \qquad \text{and} \qquad B * A_n \xrightarrow[n\to+\infty]{} B * A,$$

in other words the $*$–product is continuous in each place with respect to the uniform convergence of copulas.

A consequence

Theorem

The binary operation $*$ is associative, viz.
$A * (B * C) = (A * B) * C$, *for all* 2*–copulas* A, B, *and* C.

Corollary

The set of copulas endowed with the $*$*–product,* $(\mathcal{C}_2, *)$ *is a semigroup with identity.*

However...

...the $*$–product is not commutative, so that the semigroup $(\mathcal{C}_2, *)$ is not abelian.

Let $C_{1/2}$ be the copula belonging to the Cuadras–Augé family, defined by

$$C_{1/2}(u, v) = \begin{cases} u\sqrt{v}, & u \leq v, \\ \sqrt{u}\, v, & u \geq v. \end{cases}$$

$$(W_2 * C_{1/2})\left(\frac{1}{4}, \frac{1}{2}\right) = \frac{1}{4} - \frac{\sqrt{2}}{8} \neq \frac{1}{2} - \frac{\sqrt{3}}{4} = (C_{1/2} * W_2)\left(\frac{1}{4}, \frac{1}{2}\right)$$

Special cases

$$\Pi_2 * C = C * \Pi_2 = \Pi_2,$$
$$M_2 * C = C * M_2 = C,$$
$$(W_2 * C)(u, v) = v - C(1 - u, v),$$
$$(C * W_2)(u, v) = u - C(u, 1 - v).$$

In particular, one has $W_2 * W_2 = M_2$.

Theorem

The copulæ Π_2 and M_2 are the (right and left) annihilator and the identity of the $*$–product, respectively.

Copulæ and Conditional Expectations

Theorem

Let C be the copula of the continuous random variables X and Y defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$; then, for almost every $\omega \in \Omega$,

$$\mathbb{E}\left(\mathbf{1}_{\{X \le x\}} \mid Y\right)(\omega) = D_2 C\left(F_X(x), F_Y(Y(\omega))\right)$$

and

$$\mathbb{E}\left(\mathbf{1}_{\{Y \le y\}} \mid X\right)(\omega) = D_1 C\left(F_X(X(\omega)), F_Y(y)\right).$$

An important consequence

Corollary

Let X, Y *and* Z *be continuous random variables on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *If* X *and* Z *are conditionally independent given* Y, *then*

$$C_{XZ} = C_{XY} * C_{YZ}.$$

$*$–product and Markov processes

Theorem

Let $(X_t)_{t\in T}$ be a real stochastic process, let each random variable X_t be continuous for every $t \in T$ and let C_{st} denote the (unique) copula of the random variables X_s and X_t $(s, t \in T)$. Then the following statements are equivalent:

(a) *for all s, t, u in T,*

$$C_{st} = C_{su} * C_{ut};$$

(b) *the transition probabilities $\mathbb{P}(s, x, t, A) := \mathbb{P}\left(X_t \in A \mid X_s = x\right)$ satisfy the Chapman–Kolmogorov equations*

$$\mathbb{P}(s, x, t, A) = \int_{\mathbb{R}} \mathbb{P}(u, \xi, t, A)\, \mathbb{P}(s, x, u, d\xi)$$

The $\star$–product

The Chapman–Kolmogorov equation is a necessary but not a sufficient condition for a Markov process. This motivates the introduction of another operation on copulas.

Definition

Let $A \in \mathcal{C}_m$ and $B \in \mathcal{C}_n$; the $\star$–product of A and B is the mapping $A \star B : \mathbb{I}^{m+n-1} \to \mathbb{I}$ defined by

$$(A \star B)(u_1, \dots, u_{m+n-1}) \\ := \int_0^{x_m} D_m A(u_1, \dots, u_{m-1}, \xi)\, D_1 B(\xi, u_{m+1}, \dots, u_{m+n-1})\, d\xi.$$

Properties of the *star*–product

(a) for all copulas $A \in \mathcal{C}_m$ and $B \in \mathcal{C}_n$ the $\star$–product $A \star B$ is an $(m+n-1)$–copula, viz. $\star : \mathcal{C}_m \times \mathcal{C}_n \to \mathcal{C}_{m+n-1}$

(b) the $\star$–product is continuous in each place: if the sequence $(A_k)_{k\in\mathbb{N}}$ converges uniformly to $A \in \mathcal{C}_m$, then, for every $B \in \mathcal{C}_n$ one has both

$$A_k \star B \xrightarrow[k\to+\infty]{} A \star B \qquad \text{and} \qquad B \star A_k \xrightarrow[k\to+\infty]{} B \star A$$

(c) the $\star$–product is associative:

$$(A \star B) \star C = A \star (B \star C)$$

Characterization of Markov processes

Theorem

For a stochastic process $(X_t)_{t \in T}$ such that each random variable X_t has a continuous distribution the following statements are equivalent:

(a) (X_t) *is a Markov process;*

(b) *for every choice of* $n \geq 2$ *and of* $t_1, t_2, \dots, t_n$ *in* T *such that* $t_1 < t_2 < \cdots < t_n$

$$C_{t_1, t_2, \dots, t_n} = C_{t_1 t_2} \star C_{t_2 t_3} \star \cdots \star C_{t_{n-1} t_n},$$

where $C_{t_1, t_2, \dots, t_n}$ *is the unique copula of the random vector* $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ *and* $C_{t_j t_{j+1}}$ *is the (unique) copula of the random variables* X_{t_j} *and* $X_{t_{j+1}}$.

The role of the Chapman–Kolmogorov equations

It is now possible to see from the standpoint of copulas why the Chapman–Kolmogorov equations alone do not garantee that a process is Markov. One can construct a family of n–copulas with the following two requirements:

- they do not satisfy the conditions of the equations

$$C_{t_1,t_2,\dots,t_n} = C_{t_1 t_2} \star C_{t_2 t_3} \star \cdots \star C_{t_{n-1} t_n}$$

- they do satisfy the conditions of the equations

$$C_{st} = C_{su} * C_{ut}$$

  and are, as a consequence, compatible with the 2–copulas of a Markov process and, hence, with the Chapman–Kolmogorov equations.

Construction of the example

Consider a stochastic process (X_t) in which the random variables are pairwise independent. Thus the copula of every pair of random variables X_s and X_t is given by Π_2. Since, $\Pi_2 * \Pi_2 = \Pi_2$, the Chapman–Kolmogorov equations are satisfied. It is now an easy task to verify that for every $n > 2$, the n–fold $\star$–product of Π_2 yields

$$(\Pi_2 \star \Pi_2 \star \cdots \star \Pi_2)(u_1, u_2, \dots, u_n) = \Pi_n(u_1, u_2, \dots, u_n) \,,$$

so that it follows that the only Markov process with pairwise indedependent (continuous) random variables is one where all finite subsets of random variables in the process are independent.

Construction of the example–2

On the other hand, there are many 3–copulæ whose 2–marginals coincide with Π_2; such an instance is represented by the family of copulas

$$C_\alpha(u_1, u_2, u_3) := \Pi_3(u_1, u_2, u_3) + \alpha\, u_1\,(1 - u_1)\, u_2\,(1 - u_2)\, u_3\,(1 - u_3)\,,$$

for $\alpha \in\,]-1, 1[$. Now consider a process (X_t) such that

- three of its random variables, call them X_1, X_2 and X_3, have C_α as their copula;
- every finite set not containing all three of X_1, X_2 and X_3 is made of independent random variables;
- the n–copula $(n > 3)$ of a finite set containing all three of them is given by

$$C_{t_1,\dots,t_n}(u_1, \dots, u_n) = C_\alpha(u_1, u_2, u_3)\, \Pi_{n-3}(u_4, \dots, u_n)\,,$$

where we set $\Pi_1(t) := t$.

Construction of the example–3

Such a process exists since it is easily verified that the resulting joint distribution satisfy the compatibility of Kolmogorov's consistency theorem; this ensures the existence of a stochastic process with the specified joint distributions. Since any two random variables in this process are independent, the Chapman–Kolmogorov equations are satisfied. However, the copula of X_1, X_2 and X_3 is inconsistent with the set of equations with the $\star$–product, so that the process is not a Markov process.

A comparison

It is instructive to compare the traditional way of specifying a Markov process with the one due to Darsow, Olsen and Nguyen. In the traditional approach a Markov process is singled out by specifying the initial distribution F_0 a family of transition probabilities $\mathbb{P}(s, x, t, A)$ that satisfy the Chapman–Kolmogorov equations. Notice that in the classical approach, the transition probabilities are fixed, so that changing the initial distribution simultaneously varies all the marginal distributions. In the present approach, a Markov process is specified by giving all the marginal distributions and a family of 2–copulas that satisfies

$$C_{st} = C_{su} * C_{ut}$$

As a consequence, holding the copulas of the process fixed and varying the initial distribution does not affect the other marginals.

Copulæ and Conditional expectations–2

Definition

A copula C will be said to be idempotent (with respect to the $*$–product) if

$$C * C = C,$$

or, equivalently if, for all $(u, v) \in \mathbb{I}^2$, it satisfies the integro–differential equation

$$C(u, v) = \int_0^1 D_2 C(u, t)\, D_1 C(t, v)\, dt.$$

Both the copulæ Π_2 and M_2 are idempotent.

Pfanzagl's characterization

Theorem

Let $\mathcal{H}$ be a subset of $L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that $\alpha f \in \mathcal{H}$ $(f \in \mathcal{H},\ \alpha \in \mathbb{R})$, $1 + \mathcal{H} \in \mathcal{H}$ $(f \in \mathcal{H})$, $f \wedge g \in \mathcal{H}$ $(f,\ g \in \mathcal{H})$ and such that if $(f_n)_{n\in\mathbb{N}}$ is a decreasing sequence of elements of $\mathcal{H}$ that tends to a function $f \in L^1$, then $f \in \mathcal{H}$. Then an operator $T : \mathcal{H} \to \mathcal{H}$ is the restriction to $\mathcal{H}$ of a conditional expectation if, and only if, (a) $Tf \leq Tg$ whenever $f \leq g$ $(f, g \in \mathcal{H})$; (b) $T(\alpha f) = \alpha\, Tf$ $(\alpha \in \mathbb{R},\ f \in \mathcal{H}$; (c) $T(1 + f) = 1 + Tf$ $(f \in \mathcal{H})$, (d) $\mathbb{E}(Tf) = \mathbb{E}(f)$ $(f \in \mathcal{H})$, (e) $T^2 := T \circ T = T$. when these conditions are satisfied, then $T = \mathbb{E}_{\mathcal{G}}$, where

$$\mathcal{G} = \{A \in \mathcal{F} : T\,\mathbf{1}_A = \mathbf{1}_A\}.$$

Idempotent copulæ and Markov operators

Theorem

A Markov operator $T : L^\infty(\mathbb{I}) \to L^\infty(\mathbb{I})$ is the restriction to $L^\infty(\mathbb{I})$ of a CE *if, and only if, it is idempotent, viz.* $T^2 = T$*; when this latter condition is satisfied, then* $T = \mathbb{E}_{\mathcal{G}}$*, where* $\mathcal{G} := \{A \in \mathcal{B}(\mathbb{I}) : T\,\mathbf{1}_A = \mathbf{1}_A\}$.

Theorem

A Markov operator T is idempotent with respect to composition $T^2 = T$*, if, and only if, the copula* $C_T \in \mathcal{C}_2$ *that corresponds to it is idempotent,* $C_T = C_T * C_T$.

Copulæ and Conditional expectations–3

Theorem

For a copula C, the following statements are equivalent:

(a) *the corresponding Markov operator T_C is a CE restricted to $L^\infty(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$*

(b) *the corresponding Markov operator T_C is idempotent*

(c) *C is idempotent*

Copulæ and Conditional expectations–4

Theorem

To every sub–σ–field $\mathcal{G}$ of $\mathcal{B}$, the Borel σ–field of $\mathbb{I}$, there corresponds a unique idempotent copula $C(\mathcal{G})$ such that $\mathbb{E}_{\mathcal{G}} = T_{C(\mathcal{G})}$. Conversely, to every idempotent copula C there corresponds a unique sub–σ–field $\mathcal{G}(C)$ of $\mathcal{B}$ such that $T_C = \mathbb{E}_{\mathcal{G}(C)}$.

$$T_{\Pi_2} f = \mathbb{E}(f) = \int_0^1 f(t)\,dt \qquad \text{and} \qquad T_{M_2} f = f$$

for every f in $L^1(\mathbb{I})$. Therefore $T_{\Pi_2} = \mathbb{E}_{\mathcal{N}}$, where $\mathcal{N}$ is the trivial σ–field $\{\emptyset, \mathbb{I}\}$, and $T_{M_2} = \mathbb{E}_{\mathcal{B}}$; thus Π_2 and M_2 represent the extreme cases of copulas corresponding to CE's.

Extreme copulæ

Definition

Given a copula $C \in \mathcal{C}_2$, a copula $A \in \mathcal{C}_2$ will be said to be a left inverse of C if $A * C = M_2$, while a copula $B \in \mathcal{C}_2$ will be said to be a right inverse of C if $C * B = M_2$.

Definition

A copula $C \in \mathcal{C}_2$ is said to be extreme if the equality $C = \alpha A + (1 - \alpha) B$ with $\alpha \in]0, 1[$ implies $C = A = B$.

Theorem

If a copula $C \in \mathcal{C}_2$ possesses either a left or right inverse, then it is extreme.

Inverses of copulas

Theorem

When they exist, left and right inverses of copulas in $(\mathcal{C}_2, *)$ are unique.

Theorem

For a copula C the following statements are equivalent:

(a) *for every $v \in \mathbb{I}$ there exists $a = a(v) \in]0, 1[$ such that $D_1 C(u, v) = \mathbf{1}_{[a(v),1]}(u)$, for almost every $u \in \mathbb{I}$;*

(b) *C has a left inverse;*

(c) *there exists a Borel–measurable function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $Y = \varphi \circ X$ a.e..*

In either case the transpose C^T of C is a left inverse of C.

Kendall distribution function

If X is a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and if its d.f. F is continuous, then the random variable $F \circ X = F(X)$ is uniformly distributed on $\mathbb{I}$. This is called the probability integral transform (PIT for short)

Definition

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and on this let X and Y be random variables with joinf d.f. given by H and with marginals F and G, respectively. Then the Kendall distribution function of X and Y is the d.f. of the random variable $H(X, Y)$,

$$K_H(t) := \mathbb{P}\left(H(X, Y) \leq t\right) = \mu_H\left(\left\{(x, y) \in \overline{\mathbb{R}}^2 : H(x, y) \leq t\right\}\right).$$

Kendall distribution function–2

K_H depends only on the copula C of X and Y:

$$K_C(t) := \mathbb{P}\left(C(U,V) \leq t\right) = \mu_C\left(\{(u,v) \in \mathbb{I}^2 : C(u,v) \leq t\}\right).$$

Consider an Archimedean copula with inner generator f,

$$C_f(u,v) = g\left(f(u) + f(v)\right)$$

then

$$K_{C_f}(t) = t - \frac{f(t)}{f'(t)}$$

A characterization of Kendall d.f.

Theorem

For every copula $C \in \mathcal{C}_2$, K_C *is a d.f. in* $\mathbb{I}$ *such that, for every* $t \in \mathbb{I}$,

(a) $t \leq K_C(t) \leq 1$

(b) $\ell^- K_C(0) = 0$

Moreover the bounds of (a) *are attained, since* $K_{M_2}(t) = t$ *and* $K_{W_2}(t) = 1$ *for every* $t \in \mathbb{I}$.
For every d.f. F *that satisfies properties* (a) *and* (b) *there exists a copula* $C \in \mathcal{C}_2$ *for which* $F = K_C$.

Kendall's tau

Let (X_1, Y_1) and (X_2, Y_2) be a pair of independent random vectors defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with joint d.f. H; then the population version of Kendall's tau is defined as the difference of the probabilities of concordance and discordance, respectively, namely

$$\tau_{X,Y} := \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right].$$

The concordance function

Theorem

Let X_1, Y_1, X_2, Y_2 be continuous random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let the random vectors (X_1, Y_1) and (X_2, Y_2) be independent, let H_1 and H_2 be their respective joint d.f.'s and let the marginals d.f.'s satisfy $F_{X_1} = F_{X_2} = F$ and $F_{Y_1} = F_{Y_2} = G$, so that H_1 and H_2 both belong to the Fréchet class $\Gamma(F, G)$ and $H_1(x, y) = C_1(F(x), G(y))$ and $H_2(x, y) = C_2(F(x), G(y))$, where C_1 and C_2 are the (unique) copulæ of (X_1, Y_1) and (X_2, Y_2), respectively. Define

$$Q := \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right].$$

Then Q depends only on C_1 and C_2 and is given by

$$Q(C_1, C_2) = 4 \int_{\mathbb{I}^2} C_2(s, t)\, dC_1(s, t) - 1$$

Kendall's tau and copulæ

Corollary

The Kendall's tau of two continuous random variables X and Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ depends only on the (unique) copula C of X and Y and is given by

$$\tau_{X,Y} = 4 \int_{\mathbb{I}^2} C(s,t)\, dC(s,t) - 1\,.$$

In terms of the Kendall d.f.

$$\tau(C) = 3 - \int_0^1 K_C(t)\, dt$$

Examples

$$\tau(M_2) = 1 \qquad \tau(W_2) = -1 \qquad \tau(\Pi_2) = 0$$

For the Farlie–Gumbel–Morgenstern copula C_θ

$$\tau_\theta = \frac{2}{9}\,\theta \in \tau_\theta \in \left[-\frac{2}{9}, \frac{2}{9}\right]$$

For the Fréchet family of 2–copulas

$$C_{\alpha,\beta} = \alpha\, M_2 + (1 - \alpha - \beta)\,\Pi_2 + \beta\, W_2,$$

where $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \beta \leq 1$

$$\tau(C_{\alpha,\beta}) = \frac{1}{3}\,(\alpha - \beta)\,(\alpha + \beta + 2)$$

The case of Archimedean copulas

Theorem

The population version of Kendall's tau $\tau(C_f)$ *for an Archimedean copula* C_f *with inner additive generator* f *is given by*

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

Spearman's rho

Let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) there independent continuous random vectors having a common joint distribution function H, with marginals F and G and copula C. Then Spearman's rho ρ_{XY} is defined to be proportional to the difference between the probability of concordance and the probability of discordance for the two vectors (X_1, Y_1) and (X_2, Y_3); notice that the distribution function of the second vector is $F \otimes G$, since X_2 and Y_3 are independent. Then

$$\rho_{X,Y} := 3\left(\mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_3) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_3) < 0\right]\right)$$

Spearman's rho and copulæ

Theorem

If C is the copula of two continuous random variables X and Y, then the population version of Spearman's rho of X and Y depends only on C, will be denoted indifferently by $\rho_{X,Y}$ or by ρ_C or by $\rho(C)$, and is given by

$$\rho_{X,Y} = \rho_C = 12 \int_{\mathbb{I}^2} u\, v\, dC(u,v) - 3 = 12 \int_{\mathbb{I}^2} C(u,v)\, du\, dv - 3$$
$$= 12 \int_{\mathbb{I}^2} \{C(u,v) - u\, v\}\, du\, dv$$

The Schweizer–Wolff measure of dependence

Let X and Y be continuous random variables and let F and G be their d.f.'s, H their joint d.f., and C their (unique) connecting copula. The graph of C is a surface over the unit square, which is bounded above by the surface $z = M_2(u, v)$, and is bounded below by the surface $z = W_2(u, v)$. If X and Y happen to be independent, then the surface $z = C(u, v)$ is the hyperbolic paraboloid $z = u\,v$. The volume between the surfaces $z = C(u, v)$ and $z = u\,v$ can be used as a measure of dependence. The Schweizer–Wolff measure of dependence

$$\sigma(X, Y) := 12 \int_{\mathbb{I}^2} |C(u, v) - u\,v|\; du\,du = 12 \int_{\mathbb{I}^2} |C - \Pi_2|\; d\lambda_2$$

$$= 12 \int_{\mathbb{I}^2} |H(u, v) - F(u)\,G(v)|\; dF(u)\,dG(v)$$

Properties of then SW measure

(SW1) σ is defined for every pair of continuous random variables X and Y defined on the same probability space $(\Omega, F, \mathbb{P})$

(SW2) $\sigma(X, Y) = \sigma(Y, X)$

(SW3) $\sigma(X, Y) \in [0, 1]$

(SW4) $\sigma(X, Y) = 0$ if, and only if, X and Y are independent;

(SW5) $\sigma(X, Y) = 1$ if either $X = \varphi \circ Y$ or $Y = \psi \circ X$ for some strictly monotone functions $\varphi, \psi : \mathbb{R} \to \mathbb{R}$

(SW6) $\sigma(\varphi \circ X, \psi \circ Y) = \sigma(X, Y)$ for strictly monotone if $\varphi, \psi : \mathbb{R} \to \mathbb{R}$

(SW7) $\sigma(X, Y) = 6/\pi \ \arcsin(|\rho|/2)$ for the bivariate normal distribution with correlation coefficient ρ

(SW8) if (X_n, Y_n) has joint continuous d.f. H_n and converges in law to the random vector (X, Y) with continuous joint d.f. H_0, then $\sigma\,(X_n, Y_n) \longrightarrow \sigma(X, Y)$

Rényi's axioms

(R1) R is defined for any pair of random variables X and Y that are not a.e. constant

(R2) R is symmetric, $R(X, Y) = R(Y, X)$

(R3) for every pair of non–constant random variables X and Y, $R(X, Y)$ belongs to $[0, 1]$

(R4) $R(X, Y) = 0$ if, and only if, X and Y are independent

(R5) $R(X, Y) = 1$ if either $x = f \circ Y$ or $Y = g \circ X$ for some Borel measurable functions f and g

(R6) if $f, g : \mathbb{R} \to \mathbb{R}$ are Borel–measurable and one–to–one, then $R(f \circ X, g \circ Y) = R(X, Y)$

(R7) if the joint distribution of X and Y is a bivariate normal distribution with correlation coefficient ρ, then $R(X, Y) = |\rho|$

Other measures of dependence

- the L^∞ norm:

$$\sigma_\infty(X,Y) := k_\infty \, \| C - \Pi_2 \|_\infty = k_\infty \sup_{(u,v)\in\mathbb{I}^2} |C(u,v) - \Pi_2(u,v)| \, ;$$

- the L^p norm:

$$\sigma_p(X,Y) := k_p \left(\int_{\mathbb{I}^2} |C(u,v) - \Pi_2(u,v)|^p \, d\lambda_2 \right)^{1/p}$$

Measures of non–exchangeability

Let $\mathcal{H}(F)$ be the class of all random pairs (X, Y) such that X and Y are identically distributed with continuous joint d.f. F.

Definition

A function $\hat{\mu} : \mathcal{H}(F) \to \mathbb{R}_+$ is called a measure of non–exchangeability if

(A1) $\hat{\mu}$ is bounded, $\hat{\mu}(X, Y) \leq K$

(A2) $\hat{\mu}(X, Y) = 0$ if, and only if, (X, Y) is exchangeable

(A3) $\hat{\mu}$ is symmetric:$\hat{\mu}(X, Y) = \hat{\mu}(Y, X)$

(A4) $\hat{\mu}(X, Y) = \hat{\mu}(f(X), f(Y))$ for every strictly monotone function f

(A5) if (X_n, Y_n) and (X, Y) are pairs of random variables with joint d.f.'s H_n and H, respectively, and if H_n converges weakly to H, then $\hat{\mu}(X_n, Y_n)$ converges to $\hat{\mu}(X, Y)$

In the language of copulas

Definition

A function $\mu : \mathcal{C} \to \mathbb{R}_+$ is called a measure of non–exchangeability for $\mathcal{H}(F)$ if it satisfies the following properties:

(B1) $\mu(C) \leq K$

(B2) $\mu(C) = 0$ if, and only if, C is symmetric;

(B3) $\mu(C) = \mu(C^t)$

(B4) $\mu(C) = \mu(\hat{C})$

(B5) if $C_n \xrightarrow[to+\infty]{} C$ uniformly, then $\mu(C_n) \xrightarrow[to+\infty]{} \mu(C)$

An explicit measure

Theorem

The mapping $\mu_p : \mathcal{C} \to \mathbb{R}_+$ *defined by*

$$\mu_p(C) := d_p(C, C^t)$$

is a measure of non–exchangeability for every $p \in [1, +\infty]$.

Theorem

For every $p \in [1, +\infty[$ *and for every* $C \in \mathcal{C}_2$, *one has*

$$\mu_p(C) \leq \left(\frac{2 \cdot 3^{-p}}{(p+1)\,(p+2)} \right)^{1/p} \leq \frac{1}{3}\,.$$

Quasi–copulæ

Definition

A track B in $\mathbb{I}^d$ is a subset of unit cube $\mathbb{I}^d$ that can be written in the form

$$B := \{(F_1(t), F_2(t), \dots, F_d(t)) : t \in \mathbb{I}\}$$

where F_1, F_2, ..., F_d are continuous d.f.'s such that $F_j(0) = 0$ and $F_j(1) = 1$ for $j = 1, 2, \dots, d$

Definition

A d–quasi–copula is a function $Q : \mathbb{I}^d \to \mathbb{I}$ such that for every track B in $\mathbb{I}^d$ there exists a d–copula C_B that coincides with Q on B, namely such that, for every point $\mathbf{u} \in B$,

$$Q(\mathbf{u}) = C_B(\mathbf{u}).$$

An equivalent definition

Theorem

A d–quasi–copula Q satisfies the following properties:

(a) *for every* $j \in \{1, 2, \dots, d\}$, $Q(1, \dots, 1, u_j, 1, \dots, 1) = u_j$

(b) Q *is increasing in each place*

(c) Q *satisfies Lipschitz condition, if* $\mathbf{u}$ *and* $\mathbf{v}$ *are in* $\mathbb{I}^d$, *then*

$$|Q(\mathbf{v}) - Q(\mathbf{u})| \leq \sum_{j=1}^{d} |v_j - u_j|$$

Conversely if $Q : \mathbb{I}^d \to \mathbb{I}$ *satisfies properties* (a), (b) *and* (b), *then it is a quasi–copula.*

An immediate consequence

For $d > 2$ the function $W_d(\mathbf{u}) := \max\{u_1 + \cdots + u_d - d + 1, 0\}$ is a d–quasi–copula, but not a copula. For $d > 2$ consider the d–box

$$[\mathbf{1/2}, \mathbf{1}] = [1/2, 1] \times [1/2, 1] \times \cdots \times [1/2, 1] .$$

Then W_d–volume of this d–box is, for $d > 2$,

$$V_{W_d}\left([\mathbf{1/2}, \mathbf{1}]\right) = 1 - \frac{d}{2} < 0,$$

so that W_d cannot be a copula for $d > 2$, but is a proper quasi–copula.

A surprising result

Let μ_Q the real measure induced by the quasi–copula Q on $(\mathbb{I}^2, \mathcal{B}(\mathbb{I}^2))$.

Theorem

For all given $\epsilon > 0$ and $M > 0$, there exist a quasi–copula Q and a Borel subset S of $\mathbb{I}^2$ such that

(a) $\mu_Q(S) < -M$

(b) *for all u and v in $\mathbb{I}$, $|Q(u, v) - \Pi_2(u, v)| < \epsilon$*

Quasii–copulæ form a lattice

Given a set $\mathbf{S}$ of functions from $\mathbb{I}^d$ into $\mathbb{I}$ one defines

$$\bigwedge \mathbf{S}(\mathbf{u}) := \inf\{S(\mathbf{u}) : S \in \mathbf{S}\}.$$

Theorem

Both the upper and the lower bounds, $\bigvee \mathbf{Q}$ *and* $\bigwedge \mathbf{Q}$ *of every set* $\mathbf{Q}$ *of* d*–quasi–copulas are quasi–copulæ,* $\bigvee \mathbf{Q} \in \mathcal{Q}_d$ *and* $\bigwedge \mathbf{Q} \in \mathcal{Q}_d$.

Corollary

Both the upper and the lower bounds, $\bigvee \mathbf{C}$ *and* $\bigwedge \mathbf{C}$ *of every set* $\mathbf{C}$ *of* d*–copulas are* d*–quasi–copulæ,* $\bigvee \mathbf{C} \in \mathcal{Q}_d$ *and* $\bigwedge \mathbf{C} \in \mathcal{Q}_d$.

An example

For $\theta \in \mathbb{I}$ consider the copula

$$C_\theta(s,t) = \begin{cases} \min\{s, t-\theta\}, & (s,t) \in [0, 1-\theta] \times [\theta, 1], \\ \min\{s+\theta-1, t\}, & (s,t) \in [1-\theta, 1] \times [0, \theta], \\ W_2(s,t), & \text{elsewhere}, \end{cases}$$

If U and V are uniform rv's with $V = U + \theta \pmod 1$; then C_θ is their copula. Set $\mathbf{C} = \{C_{1/3}, C_{2/3}\}$, then $\bigvee \mathbf{C}$ is given by

$$\bigvee \mathbf{C}(s,t) = \begin{cases} \max\{0, s-1/3, t-1/3, s+t-1\}, & -1/3 \leq t-s \leq 2/3 \\ W_2(s,t), & \text{elsewhere}. \end{cases}$$

Notice

$$V_{\bigvee \mathbf{C}}\left([1/3, 2/3]^2\right) = -1/3 < 0$$

$\mathcal{Q}_d$ as a lattice

A partially ordered set $P \neq \emptyset$ is said to be a lattice if both the join $x \vee y$ and $x \wedge y$ of every pair x and y of elements of P are in P. A lattice P is said to be complete if both $\vee S$ and $\wedge S$ belong to P for every subset S of P.

Theorem

The set $\mathcal{Q}_d$ of d–quasi–copulas is a complete lattice under pointwise suprema and infima.

Theorem

Neither the family $\mathcal{C}_d$ of copulas nor the family $\mathcal{Q}_d \setminus \mathcal{C}_d$ of proper quasi–copulas is a lattice.

An introduction to Copulas

Carlo Sempi

Dipartimento di Matematica "Ennio De Giorgi"
Università del Salento
Lecce, Italy
carlo.sempi@unisalento.it

The 33rd Finnish Summer School on Probability Theory and Statistics, June 6th–10th, 2011

Outline

An example: the tent map

Choose θ in $]0,1[$ and consider the probability mass θ spread on the segment joining the points $(0,0)$ and $(\theta,1)$ and the probability mass $1-\theta$ spread on the segment joining the points $(\theta,1)$ and $(1,1)$. It is now easy to find the expression for the copula C_θ of the resulting probability distribution on the unit square:

$$C_\theta(u,v) = \begin{cases} u, & u \in [0, \theta\, v]\,, \\ \theta\, v, & u \in]\theta\, v, 1-(1-\theta)\, v[\,, \\ u+v-1, & u \in [1-(1-\theta)\, v, 1]\,. \end{cases}$$

The diagonal of a copula

The diagonal section δ_C of a copula $C \in \mathcal{C}_d$ is the function $\delta_C : \mathbb{I} \to \mathbb{I}$, defined by $\delta_C(t) := C(t, t, \dots, t)$.
The diagonal section has a probabilistic meaning. If U_1, U_2,..., U_d are random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, having uniform distribution on $(0, 1)$ and C as their (unique) copula, then

$$\begin{aligned}\delta_C(t) = C(t, t, \dots, t) &= \mathbb{P}\left(\bigcap_{j=1}^{d}\{U_j \leq t\}\right) \\ &= \mathbb{P}\left(\max\{U_1, U_2, \dots, U_d\} \leq t\right) = \mathbb{P}\left(\bigvee_{j=1}^{d} U_j \leq t\right),\end{aligned}$$

Then δ_C is the d.f. of the random variable $\max\{U_1, U_2, \dots, U_d\}$

Properties of the diagonal section

Theorem

The diagonal section δ_C *of a copula* $C \in \mathcal{C}_d$, *or of a quasi–copula* $Q \in \mathcal{Q}_d$, *satisfies the following properties:*

(D1) $\delta_C(0) = 0$ *and* $\delta_C(1) = 1$

(D2) $\forall t \in \mathbb{I} \quad \delta_C(t) \leq t$

(D3) *the function* $\mathbb{I} \ni t \to \delta_C(t)$ *is isotone;*

(D4) $|\delta_C(t') - \delta_C(t)| \leq d\,|t' - t|$ *for all* t *and* t' *in* $\mathbb{I}$

The set of diagonals will be denoted by $\mathcal{D}$

Questions

(Q.1) whether, given a diagonal $\delta \in \mathcal{D}$, there exists a copula C whose diagonal section δ_C coincides with δ, namely whether the class $\mathcal{C}_\delta$ is non–empty;

(Q.2) whether there exist bounds for the family $\mathcal{C}_\delta$; these, if they exist, are necessarily sharper than the Fréchet–Hoeffding ones;

(Q.3) whether these bounds, when they exist, are the best possible.

Answer to (Q.1)

Theorem

For every $\delta \in \mathcal{D}$, *the function* $K_\delta : \mathbb{I}^2 \to \mathbb{I}$ *defined by*

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

is a copula with diagonal δ, *so that* K_δ *belongs to* $\mathcal{C}_\delta$; *it will be called the diagonal copula associated with* δ.

The probabilistic meaning

Theorem

Let X and Y be continuous random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with a common d.f. F and copula C. Then the following statements are equivalent:

(a) *The joint d.f. of the random variables $\min\{X, Y\}$ and $\max\{X, Y\}$ is the Fréchet–Hoeffding upper bound*

(b) *C is a diagonal copula.*

More probability

Lemma

For every diagonal δ *and for every symmetric copula* $C \in \mathcal{C}_\delta$ *one has* $C \leq K_\delta$.

Theorem

For a diagonal δ *the following statements are equivalent:*

(a) δ *is the diagonal section of an absolutely continuous copula* $C \in \mathcal{C}_d$

(b) *the set* $\{t \in \mathbb{I} : \delta(t) = t\}$ *has Lebesgue measure* 0, $\lambda(\{t \in \mathbb{I} : \delta(t) = t\}) = 0$

The Bertino copula

For a given diagonal δ defined $\hat{\delta}(t) := t - \delta(t)$

Theorem

For every diagonal $\delta \in \mathcal{D}$, the function $B_\delta : \mathbb{I}^2 \to \mathbb{I}$ defined by

$$B_\delta(u, v) := \min\{u, v\} - \min\{\hat{\delta}(t) : t \in [u \wedge v, u \vee v]\}$$
$$= \begin{cases} u - \min_{t\in[u,v]}\{t - \delta(t)\}, & u \leq v, \\ v - \min_{t\in[v,u]}\{t - \delta(t)\}, & v \leq u \end{cases}$$

is a symmetric 2–copula having diagonal equal to δ, i.e., $B_\delta \in \mathcal{C}_\delta$. B_δ is called the Bertino copula of δ.

Bounds for copulas with given diagonal–1

Theorem

For every diagonal $\delta \in \mathcal{D}$*, the function* $A_\delta : \mathbb{I}^2 \to \mathbb{I}$ *defined by*

$$A_\delta(u,v) := \min\left\{u, v, \max\{u,v\} - \max\{\hat{\delta}(t) : t \in [u \wedge v, u \vee v]\}\right\}$$

$$= \begin{cases} \min\left\{u, v - \max_{t\in[u,v]}\{t - \delta(t)\}\right\}, & u \le v, \\ \min\left\{v, u - \max_{t\in[v,u]}\{t - \delta(t)\}\right\}, & v \le u \end{cases}$$

is a symmetric 2*–quasi–copula having diagonal equal to* δ*, i.e.,* $A_\delta \in \mathcal{Q}_\delta$.

Bounds for copulas with given diagonal–2

Theorem

For every diagonal δ and for every copula $C \in \mathcal{C}_\delta$ one has $B_\delta \leq C \leq A_\delta$.

Theorem

The quasi–copula A_δ is a copula if, and only if, $A_\delta = K_\delta$.

Theorem

For the quasi–copula A_δ the following statements are equivalent:

(a) $A_\delta = K_\delta$

(b) *the graph of the function $t \mapsto \delta(t)$ is piecewise linear; each segment has slope equal to 0, 1 or 2 and has at least one of its endpoints on the diagonal $v = u$.*

Statement of the problem

In its most general form, the problem runs as follows. If k and d with $1 < k \leq d$ are natural numbers, the d–copula C has $\binom{d}{k}$ k–marginals, which are obtained by setting $d - k$ of its arguments equal to 1. In the other direction, if at most $\binom{d}{k}$ k–copulæ are given, there may not exist a d–copula of which the given k–copulæ are the k–marginals. This may easily be seen in the case $d = 3$ and $k = 2$; if, for instance, the three two copulæ are all equal to W_2, then, in view of the probabilistic meaning of the copula W_2, there is no 3–copula C of which they are the marginals. On the other hand, if an d–copula exists of which the given copulæ are the k–marginals, then these are said to be compatible.

The special case $d = 3$ and $k = 2$

Let A and B be 2–copulæ, $A, B \in \mathcal{C}_2$, and denote by $\mathcal{D}(A, B)$ the set of all 2–copulas that are compatible with A and B, in the sense that, if C is in $\mathcal{D}(A, B)$, then there exists a 3–copula $\widetilde{C}$ such that, for all $(u, v, w) \in [0, 1]^3$,

$$\widetilde{C}(u, v, 1) = A(u, v), \quad \widetilde{C}(1, v, w) = B(v, w), \quad \widetilde{C}(u, 1, w) = C(u, w).$$

Theorem

Given any two 2–copulas A and B, there always exists a 2–copula C that is compatible with A and B, namely $\mathcal{D}(A, B) \neq \emptyset$, for instance $A * B$.

Examples

$$C_{W_2,W_2}(u,v,w) = \max\{0, v + (u \wedge w) - 1\},$$
$$C_{M_2,M_2}(u,v,w) = u \wedge v \wedge w = M_3(u,v,w),$$
$$C_{W_2,M_2}(u,v,w) = \max\{0, u + (v \wedge w) - 1\},$$
$$C_{M_2,W_2}(u,v,w) = \max\{0, (u \wedge v) - 1 + w\},$$
$$C_{\Pi_2,\Pi_2}(u,v,w) = uvw = \Pi_3(u,v,w),$$
$$C_{\Pi_2,M_2}(u,v,w) = u\, M_2(v,w),$$
$$C_{M_2,\Pi_2}(u,v,w) = w\, M_2(u,v).$$

Properties of $\mathcal{D}(A, B)$

Theorem

The set $\mathcal{D}(A, B)$ of copulas that are compatible with two given bivariate copulas A and B is convex and compact with respect to the topology of uniform convergence in $\mathbb{I}^2$.

Minimality of $\mathcal{D}(A, B)$

The class $\mathcal{D}(A, B)$ is said to be minimal when $\mathcal{D}(A, B) = \{A * B\}$. It is worth asking: when is this the case? The following theorem provides a sufficient condition for this to happen.

Theorem

Let A and B be two bivariate copulas with $A = C_{f,g}$ and $B = C_{p,r}$, where f, g, p and r are measure–preserving transformations from $\mathbb{I}$ into $\mathbb{I}$, and either pair (f, g) or (p, r) is made of one–to–one transformations. Then $\mathcal{D}(A, B)$ is minimal.

Corollary

If either A or B (or both) is a shuffle of Min, then

$$\mathcal{D}(A, B) = \{A * B\}.$$

Gluing of two copulas–1

Let A and B be d–copulæ, $A, B \in \mathcal{C}_d$, let $i \in \{1, 2, \dots, n\}$, and choose θ in $]0, 1[$. Define the $(u_i = \theta)$–*gluing* of A and B via

$$\left(A \bigotimes_{u_i=\theta} B \right) (u_1, \dots, u_{i-1}, u_i, u_{i+1}, \dots, u_d)$$

$$:= \theta\, A \left(u_1, \dots, u_{i-1}, \frac{u_i}{\theta}, u_{i+1}, \dots, u_d \right)$$

for $u_i \in [0, \theta]$

Gluing of two copulas–2

$$\left(A \bigotimes_{u_i=\theta} B \right) (u_1, \dots, u_{i-1}, u_i, u_{i+1}, \dots, u_d)$$
$$:= \theta\, A\, (u_1, \dots, u_{i-1}, 1, u_{i+1}, \dots, u_d)$$
$$+ (1-\theta)\, B \left(u_1, \dots, u_{i-1}, \frac{u_i - \theta}{1-\theta}, u_{i+1}, \dots, u_d \right)$$

for $u_i \in [\theta, 1]$.

Theorem

For every pair A and B of d–copulas, for every index $i \in \{1, 2, \dots, d\}$, and for every $\theta \in]0, 1[$, $A \bigotimes_{u_i=\theta} B$ is a d–copula.