Je dédie cette thèse

à ma chère maman et à mon cher papa.

Vous m'avez toujours aidés et soutenus pour réussir mes études et devenir ce que je suis aujourd'hui, car toute réussite dans ma vie est grâce à vos sacrifices et tout ce que je pourrai vous dire ne suffira pour vous exprimer ma gratitude et ma reconnaissance pour votre dévouement et votre grande patience.

Avec l'expression de mon profond amour.

"Ce qui compte ne peut pas toujours être compté, et ce qui peut être compté ne compte pas forcément."

Albert Einstein

Remerciements

J'ai pue réaliser cette thèse avec l'aide et les conseils des différentes personnes.

Je souhaite d'abord exprimer ma profonde gratitude à mon Directeur de thèse, M. RABHI Abbes, pour son temps précieux qu'il m'a consacré, pour son appui scientifique, qui m'a été indispensable, pour son savoir qu'il m'a communiqué, pour son aimable attention, pour sa disponibilité, pour ses encouragements, et sa grande patience.

Je suis très honoré que M. OUAKAS Seddik a accepté de présider le jury et d'avoir ainsi accepté de juger mon travail.

Merci à M. KANDOUCI Abdeldjebbar, Directeur du Laboratoire Modèles Stochastiques,Statistiques et Applications, de m'avoir accepté comme membre au laboratoire, pour ses encouragements, ses conseils et de me faire aussi l'honneur d'avoir accepté de participer à ce jury et de juger mon travail.

Je suis particulièrement sensible à la présence de Mlle BOUCHENTOUF Amina Angelika dans le jury. Je la remercie chaleureusement pour son soutien constant tout au long de la préparation de ce travail, pour ses encouragements, ses orientations et ses conseils.

Mes remerciements et mon profond respect iront de même à M. BENAISSA Samir et à M. GUENDOUZI Toufik, pour avoir bien voulu examiner ce travail et pour le travail que cela leur demande.

Mes pensées vont à mes chers parents, pour tous les efforts qu'ils m'ont consentis pour m'aider à réaliser ce travail et l'affection qu'ils m'ont témoignée dans les moments difficiles, pour leurs précieux conseils et leur grande patience.

À mes chères petites nièces, Farah,"ma princesse", merci pour ton soutien et ton affection. Kenza, "ma belle", tes bisous et câlins du matin m'ont rempli de joie et apporter des encouragements. Alaa, "bébé", ton grand sourire m'a apporter beaucoup de bonheur. Sans oublier mon neveu, Salah, "mon choux" tu m'as apporté que du meilleur. Vous avez rempli mes moments de joie et de bonheur. Je vous souhaite tout ce qu'il y a de meilleur.

Je remercie, mes sœurs Malika, Soumeya et Aziza, pour leur soutien moral tout au long de la préparation de cette thèse. À mes beau-frères Ali, Sofiane et Mohamed, pour leurs encouragements.

Je tien à remercier Sid Ahmed, mon époux, pour son soutien et ses encouragements qu'il m'a apporté dans les moments difficiles. Merci pour ta patience.

Mes remerciements iront de même à ma belle famille, pour les encouragements et le soutien qu'ils m'ont apporté.

Je remercie M. ABDELHAK Yacine, mon enseignant du secondaire de Mathématiques, de m'avoir donné le goût des mathématiques. Merci pour tout.

A special tanks to Miss. BAGHLI Asmaa, "my friend" for her great moral support in the development of this work.

Sans oublier de remercier, toutes les personnes qui de loin ou de prés, m'ont apporté des encouragements pour élaborer ce travail, notamment mes collègues Doctorants: Zouaouia BOULENOIR, Soumia SAKKAL, Asmaa BENCHAKOR, Imene MAAMAR, Imene MAACHOU, Soumia IDRISSI, Sara SOLTANI, Tayeb DJABBOURI.

Merci à tous.

Contents

Chapter 1

General introduction

Natural and Physical Sciences are traditionally attached to the understanding of the underlying phenomenon observations by creating a model, whose validity and relevance can be questioned during an experiment. It is through this dialectic between theory and experimental work, that scientific knowledge progresses. To quote H. Poincaré *Science and Hypothesis*, note that if science is built on facts learned, *"an accumulation of facts is no more a science than a heap of stones is a house"*. Also, in developing his theory, the scientist must constantly submit its assumptions verification by experiment and to accomplish this validation work, *" above all, the scientist must provide"*.

That is to say, besides its eminently practical interest, the importance of the problem of prediction, includes basis for the use of statistics for scientific purposes.

Obviously, the purpose of this thesis is much smaller compared to these epistemological issues. This work deals with the prediction Statistics. It tackles this problem from a point of view nonparametric.

Indeed, most of the physical phenomena in nature have a random element in their structure, whereby the magnitudes are variable and can not be predicted with certainty. It is then natural to adopt an approach Statistics. A probabilistic model is then supposed to describe the behavior of the phenomenon, which evolves according to a probability law.

The approch discussed in this thesis is somewhat different. The probabilistic model is nonparametric, and a sample of n pairs of random variables are observed $(X_i, y_i); i = 1, \ldots, n$ independent, identically distributed. It then tries to predict the sense of explaining the variable Y by the predictor variable X. The interest then focuses on the parameter estimation of conditional position, constructed from a regression estimator. For this purpose there is provided a new and studying the regression estimator (Chapters 2 and 3).

Let $(X_i, y_i); i = 1, \ldots, n$ independent samples identically distributed random variables with real value (X, Y) sitting on a given probability space. To predict the response Y of the input variable X at a given location x, it is of great interest to estimate the conditional mean or the regression function $E(Y/X = x) = r(x)$. A natural approach to estimate the regression would be to exploit the identity

$$r(x) = E(Y/X = x) = \int yf(y/x)dy = \int y\frac{f(x, y)}{f(x)}dy$$

Where f_{XY} and f_X denote the joint density of (X, Y) and X , respectively. By introducing the kernel estimator of Nadaraya-Watson regression , namely ,

$$\hat{r}(x) = \frac{\sum_{i=1}^{n} Y_i K(\frac{x-X_i}{h_n})}{\sum_{i=1}^{n} K(\frac{x-X_i}{h_n})}$$

However, its form as a quotient of two estimators, the probabilistic behavior of the Nadaraya-Watson estimator is difficult to study. It is usually treated with a centering waiting for the numerator and denominator of the inverse linearization, see e.g. Fan, J. and Yao, Q. (2005) or Bosq, D. (1998) for details. As a result, the practical applications of this estimator can lead to numerical instability when the denominator is close to zero.

So, What criteria can we choose to measure the performance of our prediction? Several methods exist in the literature, however there is no method universally better than the other.

To overcome these problems, we propose an estimator which is based on the idea of using synthetic data, i,e. a data representation more adapted to the problem as the original. By transforming the data by quantile transformations and using the copula function, the estimator turns out to have a remarkable product form

$$\hat{r}(x) = E(Y/X = x) = Y\hat{c}(F(x), G(y))$$

His study then is particularly simple: it reduces to those already made on the estimation of nonparametric regression.

Copula theory, following the works of Sklar in 1959, allows a flexible modeling of dependence between two or more random variables. In recent years, the growing interest for this theory is phenomenal. Thomas Mikosch stated that in September 2005, a Google search on the term "copula" produced 650,000 results. Then, in January 2007, this same query generates more than 1.13 million. Given the number of publications in scientific journals and the number of papers available on Internet, it is undeniable that passion to the copula theory is still booming.

The progress of applications of this theory is wide in the field of finance, risk management, performance evaluation of assets, the valuation of derivatives, the extreme value theory, contagion require flexible and practical models of addiction. The construction and properties of copulas have been studied rather extensively during the last 15 years. Hutchinson and Lai (1990) were among the early authors who popularized the study of copulas. Nelsen (1999) presented a comprehensive treatment of bivariate copulas, while Joe (1997) devoted a chapter of his book to multivariate copulas. Further authoritative updates on copulas are given in Nelsen (2006). Copula methods have many important applications in insurance and finance Cherubini et al. (2004) and Embrechts et al. (2003).

Briefly speaking, copulas are functions that join multivariate distributions to their one-dimensional marginal distribution functions. Equivalently, copulas are multivariate distributions whose marginals are uniform on the interval (0, 1). In this thesis, our attention id restricted to bivariate copulas. Fisher (1997) gave two major reasons as to why copulas are of interest to statisticians: *firstly, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions."*

Specifically, copulas are an important part of the study of dependence between two variables since they allow us to separate the effect of dependence from the effects of the marginal distributions. This feature is analogous to the bivariate normal distribution where the mean vectors are unlinked to the covariance matrix and jointly determine the distribution. Many authors have studied constructions of bivariate distributions with given marginals: This may be viewed as constructing a copula. Nonparametric estimators of copula densities have been suggested by Gijbels and Mielnicsuk (1990) and Fermanian and Scaillet (2005), who used kernel methods, Sancetta (2003) and Sancetta and Satchell (2004), who used techniques based on Bernstein polynomials. Biau and Wegkamp (2006) proposed estimating the copula density through a minimum distance criterion. Faugeras (2008) in his thesis studied the quantile copula approach to conditional density estimation.

There is a fast-growing industry for copulas. They have useful applications in econometrics, risk management, finance, insurance, etc. The commercial statistics software SPLUS provides a module in FinMetrics that include copula fitting written by Carmona

(2004). One can also get copula modules in other major software packages such as R, Mathematica, Matlab, etc. The International Actuarial Association (2004) in a paper on Solvency II,1 recommends using copulas for modeling dependence in insurance portfolios. Moodyïs uses a Gaussian copula for modeling credit risk and provides software for it that is used by many financial institutions. Basle II2, *copulas are now standard tools in credit risk management*. There are many other applications of copulas, especially the Gaussian copula, the extreme-value copulas, and the Archimedean copula. Now, we classify these applications into several categories.

1.1 Some generalities on Copulas

The study of copulas and their applications in statistics is a rather modern phenomenon. Until quite recently, it was difficult to even locate the word "copula" in the statistical literature. There is no entry for "copula" in the nine volume Encyclopedia of Statistical Sciences, nor in the supplement volume. However, the first update volume, published in 1997, does have such an entry (Fisher 1997). The first reference in the Current Index to Statistics to a paper using "copula" in the title or as a keyword is in Volume 7 (1981) [the paper is (Schweizer and Wolff 1981)]-indeed, in the first eighteen volumes (1975-1992) of the Current Index to Statistics there are only eleven references to papers mentioning copulas. There are, however, 71 references in the next ten volumes (1993-2002).

Further evidence of the growing interest in copulas and their applications in statistics and probability in the past fifteen years is afforded by five international conferences devoted to these ideas: the "Symposium on Distributions with Given Marginals (Fréchet Classes)" in Rome in 1990; the conference on "Distributions with Fixed Marginals, Doubly Stochastic Measures, and Markov Operators" in Seattle in 1993; the conference on "Distributions with Given Marginals and Moment Problems" in Prague in 1996; the conference on "Distributions with Given Marginals and Statistical Modelling" in Barcelona in 2000; and the conference on "Dependence Modelling: Statistical Theory and Applications in Finance and Insurance" in Québec in 2004. As the titles of these conferences indicate, copulas are intimately related to study of distributions with "fixed" or "given" marginal distributions. The published proceedings of the first four conferences (Dall'Aglio et al. 1991; Rüschendorf et al. 1996; Beneš and *Štěpn* 1997; Cuadras et al.2002) are among the most accessible resources for the study of copulas and their applications.

What are copulas? From one point a view, copulas are functions that join or "couple" multivariate distribution functions to their onedimensional marginal distribution functions. Alternatively, copulas are multivariate distribution functions whose one-dimensional margins are uniform on the interval $(0,1)$. this chapter will be devoted to presenting a complete answer to this question.

Why are copulas of interest to students of probability and statistics? As Fisher (1997) answers in his article in the first update volume of the Encyclopedia of Statistical Sciences, "Copulas [are] of interest to statisticians for two main reasons: Firstly, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions, sometimes with a view to simulation."

Proceed briefly to the history of the developement and the study of copulas. For more details on those who participated in the evolution of the topic, see documents by Dall'Aglio (1991) and Schweizer (1991) the work of the Conference of Rome and Article Sklar (1996) acts conference in Seattle.

The word copula is a Latin noun that means "a link, tie, bond" (Cassell's Latin Dictionary) and is used in grammar and logic to describe "that part of a proposition which connects the subject and predicate" (Oxford English Dictionary). The word copula was first employed in a mathematical or statistical sense by Abe Sklar (1959) in the theorem (which now bears his name) describing the functions that "join together" one-dimensional distribution functions to form multivariate distribution functions .In (Sklar 1996) we have the following account of the events leading to this use of the term copula:

Feron (1956), in studying three-dimensional distributions had introduced auxiliary functions, defined on the unit cube, that connected such distributions with their one-dimensional margins. I saw that similar functions could be defined on the unit n-cube for all $n \geq 2$ and would similarly serve to link n-dimensional distributions to their one-dimensional margins. Having worked out the basic properties of these functions, I wrote about them to Frechet, in English. He asked me to write a note about them in French. While writing this, I decided I 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, I 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. Frechet received my note, corrected one mathematical statement, made some minor corrections to my French, and had the note published by the Statistical Institute of the University of Paris as Sklar (1959).

But as Sklar notes, the functions themselves predate the use of the term copula. They appear in the work of Fréchet, Dall'Aglio, Féron, and many others in the study of multivariate distributions with fixed univariate marginal distributions. Indeed, many of the basic results about copulas can be traced to the early work of Wassily Hoeffding. In (Hoeffding 1940, 1941) one finds bivariate "standardized distributions" whose support is contained in the square $[-\frac{1}{2}, \frac{1}{2}]^2$ and whose margins are uniform on the interval $[-\frac{1}{2}, \frac{1}{2}]$. (As Schweizer (1991) opines, "had Hoeffding chosen the unit square $[0,1]^2$ instead of $[-\frac{1}{2}, \frac{1}{2}]^2$ for his normalization, he would have discovered copulas.")

Hoeffding also obtained the basic best-possible bounds inequality for these functions, characterized the distributions ("functional dependence") corresponding to those bounds, and studied measures of dependence that are "scale-invariant," i.e., invariant under strictly increasing transformations. Unfortunately, until recently this work did not receive the attention it deserved, due primarily to the fact the papers were published in relatively obscure German journals at the outbreak of the Second World War. However, they have recently been translated into English and are among Hoeffdin's collected papers, recently published by Fisher and Sen (1994). Unaware of Hoeffding's work, Fréchet (1951) independently obtained many of the same results, which has led to the terms such as "Fréchet bounds" and "Fréchet classes." In recognition of the shared responsibility for these important ideas, we will refer to "Fréchet-Hoeffding bounds" and "Fréchet-Hoeffding classes." After Hoeffding, Fréchet, and Sklar, the functions now known as copulas were rediscovered by several other authors. Kimeldorf and Sampson (1975b) referred to them as uniform representations, and Galambos (1978) and Deheuvels (1978) called them dependence functions.

At the time that Sklar wrote his 1959 paper with the term "copula", he was collaborating with Berthold Schweizer in the development of the theory of probabilistic metric spaces, or PM spaces. During the period from 1958 through 1976, most of the important results concerning copulas were obtained in the course of the study of PM spaces. Recall that (informally) a metric space consists of a set S and a metric d that measures "distances" between points, say p and q, in S. In a probabilistic metric space, we replace the distance $d(p,q)$ by a distribution function F_{pq} , whose value $F_{pq}(x)$ for any real x is the probability that the distance between p and q is less than x. The first difficulty in the construction of probabilistic metric spaces comes when one tries to find a "probabilistic" analog of the triangle inequality $d(p,r) \leq d(p,q) + d(q,r)$. What is the corresponding relationship among the distribution functions F_{pr}, F_{pq} , and F_{qr} for all p, q, and r in S? Karl Menger (1942) proposed $F_{pr}(x + y) \geq T(F_{pq}(x), F_{qr}(y))$; where T is a triangle norm or t-norm. Like a copula, a t-norm maps $[0, 1]^2$ to $[0, 1]$, and joins distribution functions. Some t-norms are copulas, and conversely, some copulas are t-norms. So, in a sense, it was inevitable that copulas would arise in the study of PM spaces. For a thorough treatment of the theory of PM spaces and the history of its development, see (Schweizer and Sklar 1983; Schweizer 1991).

Among the most important results in PM spaces-for the statistician is the class of Archimedean t-norms, those t-norms T that satisfy $T(u, u) < u$ for all u in $(0, 1)$. Archimedean t-norms that are also copulas are called Archimedean copulas. Because of their simple forms, the ease with which they can be constructed, and their many nice properties,Archimedean copulas frequently appear in discussions of multivariate distributions - see, for example, (Genest and MacKay 1986a,b; Marshall and Olkin 1988; Joe 1993, 1997).

We now turn our attention to copulas and dependence. The earliest paper explicitly relating copulas to the study of dependence among random variables appears to be (Schweizer and Wolff 1981). In that paper, Schweizer and Wolff discussed and modified Rényi's (1959) criteria for measures of dependence between pairs of random variables, presented the basic invariance properties of copulas under strictly monotone transformations of random variables, and introduced the measure of dependence now known as Schweizer and Wolff's σ. In their words, since

> *... under almost surely increasing transformations of (the random variables), the copula is invariant while the margins may be changed at will, it follows that it is precisely the copula which captures those properties of the joint distribution which are invariant under almost surely strictly increasing transformations. Hence 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.*

Of course, copulas appear implicitly in earlier work on dependence by many other authors, too many to list here, so we will mention only two. Foremost is Hoeffding. In addition to studying the basic properties of "standardized distributions" (i.e., copulas), Hoeffding (1940, 1941) used them to study nonparametric measures of association such as Spearman's rho and his "dependence index" ϕ^2. Deheuvels (1979, 1981a,b,c) used "emperical dependence functions" (i.e., empirical copulas, the sample analogs of copulas) to estimate the population copula and to construct various nonparametric tests of independence. Section 1.1.2 is devoted to an introduction to the role played by copulas in the study of dependence.

The study of copulas and the role they play in probability, statistics, and stochastic processes is a subject still in its infancy. There are many open problems and much work to be done.

1.1.1 Definitions and Basic Properties

As we have just refer that the copulas are "functions or join multivariate distribution functions to their one-dimensional marginal distribution functions" and also "distribution functions whose one-dimensional margins are uniform". But neither of these statements is a definition-hence we will devote this section to giving a precise definition of copulas and to examining some of their elementary properties.

But first we present a glimpse of where we are headed. Consider for a moment a pair of random variables X and Y, with distribution functions $F(x) = P[X \leq x]$ and $G(y) = P[Y \leq y]$, respectively, and a joint distribution function $H(x,y) = P[X \leq x, Y \leq y]$ (we will review definitions of random variables, distribution functions, and other important topics as needed in the course of this chapter). To each pair of real numbers (x,y) we can associate three numbers: $F(x), G(y)$, and $H(x,y)$. Note that each of these numbers lies in the interval $[0,1]$. In other words, each pair (x,y) of real numbers leads to a point $(F(x), G(y))$ in the unit square $[0,1] \times [0,1]$, and this ordered pair in turn corresponds to a number $H(x,y)$ in $[0,1]$. We will show that this correspondence, which assigns the value of the joint distribution function to each ordered pair of values of the individual distribution functions, is indeed a function. Such functions are copulas. To accomplish what we have outlined above, we need to generalize the notion of "nondecreasing" for univariate functions to a concept applicable to multivariate functions. We begin with some notation and definitions. Throughout this work we limit ourselves to two-dimensional case.

1.1.1.1 Preliminaries

The focus of this section is the notion of a "2-increasing" function-a two-dimensional analog of a nondecreasing function of one variable. But first we need to introduce some notation. We will let $\mathbb{R}$ denote the ordinary real line $(-\infty, \infty)$, $\overline{\mathbb{R}}$ denote the extended real line $[-\infty, \infty]$, and $\overline{\mathbb{R}}^2$ denote the extended real plane $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$. A rectangle in $\overline{\mathbb{R}}^2$ is the Cartesian product B of two closed intervals: $B = [x_1, x_2] \times [y_1, y_2]$. The vertices of a rectangle B are the points $(x_1, y_1), (x_1, y_2), (x_2, y_1)$, and (x_2, y_2). The unit square I^2 is the product $I \times I$ where $I = [0, 1]$. A 2-place real function H is a function whose domain, $DomH$, is a subset of $\overline{\mathbb{R}}^2$ and whose range, $RanH$, is a subset of $\mathbb{R}$.

Definition 1.1.1. *Let S_1 and S_2 be nonempty subsets of $\overline{\mathbb{R}}$, and let H be a two-place real function such that $DomH = S_1 \times S_2$. Let $B = [x_1, x_2] \times [y_1, y_2]$ be a rectangle all of whose vertices are in $DomH$. Then the H-volume of B is given by*

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1). \tag{1.1}$$

Note that if we define the first order differences of H on the rectangle B as

$$\Delta_{x_1}^{x_2} = H(x_2, y) - H(x_1, y) \ \ \text{and} \ \ \Delta_{y_1}^{y_2} = H(x, y_2) - H(x, y_1),$$

then the H-volume of a rectangle B is the second order difference of H on B,

$$V_H(B) = \Delta_{y_1}^{y_2} \Delta_{x_1}^{x_2} H(x, y).$$

Definition 1.1.2. *A 2-place real function H is 2-increasing if $V_H(B) \geq 0$ for all rectangles B whose vertices lie in $DomH$. When H is 2-increasing, we will occasionally refer to the H-volume of a rectangle B as the H-measure of B. Some authors refer to 2-increasing functions as quasi-monotone. We note here that the statement "H is 2-increasing" neither implies nor is implied by the statement "H is nondecreasing in each argument".*

The following lemmas will be very useful in the next section in establishing the continuity of subcopulas and copulas. The first is a direct consequence of Definitions 1.1.1 and 1.1.2.

Lemma 1.1.1. *[Nelsen (2006)] Let S_1 and S_2 be nonempty subsets of $\overline{\mathbb{R}}$, and let H be a 2-increasing function with domain $S_1 \times S_2$. Let x_1, x_2 be in S_1 with $x_1 \leq x_2$, and let y_1, y_2 be in S_2 with $y_1 \leq y_2$. Then the function $t \mapsto H(t, y_2) - H(t, y_1)$ is nondecreasing on S_1, and the function $t \mapsto H(x_2, t) - H(x_1, t)$ is nondecreasing on S_2.*

As an immediate application of this lemma, we can show that with an additional hypothesis, a 2-increasing function H is nondecreasing each argument. Suppose S_1 has a least element a_1 and that S_2 has a least element a_2 . We say that a function H from $S_1 \times S_2$ into $\mathbb{R}$ is grounded if $H(x, a_2) = 0 = H(a_1, y)$ for all (x, y) in $S_1 \times S_2$. Hence we have

Lemma 1.1.2. *[Nelsen (2006)] Let S_1 and S_2 be nonempty subsets of $\overline{\mathbb{R}}$, and let H be a grounded 2-increasing function with domain $S_1 \times S_2$. Then H is nondecreasing in each argument.*

$$\begin{aligned} &DomF = S_1, \text{ and } F(x) = H(x, b_2) \text{ for all } x \text{ in } S_1; \\ &DomG = S_2, \text{ and } G(y) = H(b_1, y) \text{ for all } y \text{ in } S_2. \end{aligned}$$

Lemma 1.1.3. *Let S_1 and S_2 be nonempty subsets of $\overline{\mathbb{R}}$, and let H be a grounded 2-increasing function, with margins, whose domain is $S_1 \times S_2$. Let (x_1, y_1) and (x_2, y_2) be any points in $S_1 \times S_2$. Then*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|.$$

1.1.1.2 Copulas

We are now in a position to define the functions-copulas that are the subject of this chapter. To do so, we first define subcopulas as a certain class of grounded 2-increasing functions with margins; then we define copulas as subcopulas with domain I^2.

Definition 1.1.3. *A two-dimensional subcopula (or 2-subcopula, or briefly, a subcopula) is a function C' with the following properties:*

1. *$DomC' = S_1 \times S_2$, where S_1 and S_2 are subsets of I containing 0 and 1;*

2. *C' is grounded and 2-increasing;*

3. *For every u in S_1 and every v in S_2,*

$$C'(u, 1) = u \text{ and } C'(1, v) = v. \tag{1.2}$$

Note that for every (u, v) in $DomC'$, $0 \leq C'(u, v) \leq 1$, so that $RanC'$ is also a subset of I.

Definition 1.1.4. *A two-dimensional copula (or 2-copula, or briefly, a copula) is a 2-subcopula C whose domain is I^2.*

Equivalently, a copula is a function C from I^2 to I with the following properties:

1. *For every u, v in I,*
$$C(u,0) = 0 = C(0,v) \tag{1.3}$$
and
$$C(u,1) = u \ \ and \ \ C(1,v) = v; \tag{1.4}$$

2. *For every u_1, u_2, v_1, v_2 in I such that $u_1 \leq u_2$ and $v_1 \leq v_2$,*
$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0. \tag{1.5}$$

Because $C(u,v) = V_C([0,u] \times [0,v])$, one can think of $C(u,v)$ as an assignment of a number in I to the rectangle $[0,u] \times [0,v]$. Thus (1.5) gives an "inclusion-exclusion" type formula for the number assigned by C to each rectangle $[u_1,u_2] \times [v_1,v_2]$ in I^2 and states that the number so assigned must be nonnegative.

The distinction between a subcopula and a copula (the domain) may appear to be a minor one, but it will be rather important in the next section when we discuss Sklar's theorem. In addition, many of the important properties of copulas are actually properties of subcopulas.

Theorem 1.1.1. *Let C' be a subcopula. Then for every (u,v) in $DomC'$,*
$$\max(u+v-1,0) \leq C'(u,v) \leq \min(u,v). \tag{1.6}$$

Because every copula is a subcopula, the inequality in the above theorem holds for copulas. Indeed, the bounds in (1.6) are themselves copulas and are commonly denoted by $M(u,v) = \min(u,v)$ and $W(u,v) = \max(u+v-1,0)$. Thus for every copula C and every (u,v) in I^2,
$$W(u,v) \leq C(u,v) \leq M(u,v). \tag{1.7}$$

Inequality (1.7) is the copula version of the Fréchet-Hoeffding bounds inequality, which we shall encounter later in terms of distribution functions. We refer to M as the Fréchet-Hoeffding upper bound and W as the Fréchet-Hoeffding lower bound. A third important copula that we will frequently encounter is the product copula $\Pi(u,v) = uv$. The following theorem, which follows directly from Lemma 1.1.3, establishes the continuity of subcopulas and hence of copulas via a Lipschitz condition on I^2.

Theorem 1.1.2. *[Nelsen (2006)] Let $C^{'}$ be a subcopula. Then for every $(u_1, u_2), (v_1, v_2)$ in $DomC^{'}$,*

$$|C^{'}(u_2, v_2) - C^{'}(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|. \tag{1.8}$$

Hence $C^{'}$ is uniformly continuous on its domain.

The sections of a copula will be employed in section 1.1.2 to provide interpretations of certain dependence properties.

Definition 1.1.5. *Let C be a copula, and let a be any number in I. The horizontal section of C at a is the function from I to I given by $t \mapsto C(t, a)$; the vertical section of C at a is the function from I to I given by $t \mapsto C(a, t)$; and the diagonal section of C is the function δ_C from I to I defined by $\delta_C(t) = C(t, t)$.*

The following corollary is an immediate consequence of Lemma 1.1.2 and Theorem 1.1.2.

Corollary 1.1.1. *[Nelsen (2006)] The horizontal, vertical, and diagonal sections of a copula C are all nondecreasing and uniformly continuous on I.*

Various applications of copulas that we will encounter in later sections of this chapter involve the shape of the graph of a copula, i.e., the surface $z = C(u, v)$. It follows from Definition 1.1.4 and Theorem 1.1.2 that the graph of any copula is a continuous surface within the unit cube I^3 whose boundary is the skew quadrilateral with vertices $(0, 0, 0), (1, 0, 0), (1, 1, 1)$, and $(0, 1, 0)$; and from Theorem 1.1.1 that this graph lies between the graphs of the Fréchet-Hoeffding bounds, i.e., the surfaces $z = M(u, v)$ and $z = W(u, v)$. In Figure 1.1 we present the graphs of the copulas M and W, as well as the graph of Π, a portion of the hyperbolic paraboloid $z = uv$.

A simple but useful way to present the graph of a copula is with a contour diagram (Conway 1979), that is, with graphs of its level sets-the sets in I^2 given by $C(u, v) = a$ constant, for selected constants in I. In Figure 1.2 we present the contour diagrams of the copulas M, Π,and W. Note that the points $(t, 1)$ and $(1, t)$ are each members of the level set corresponding to the constant t. Hence we do not need to label the level sets in the diagram, as the boundary conditions $C(1, t) = t = C(t, 1)$ readily provide the constant

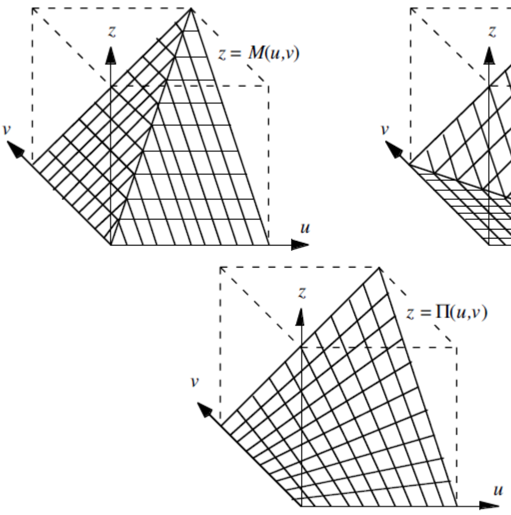

Figure 1.1: Graphs of the copulas M, Π, and W.

for each level set.

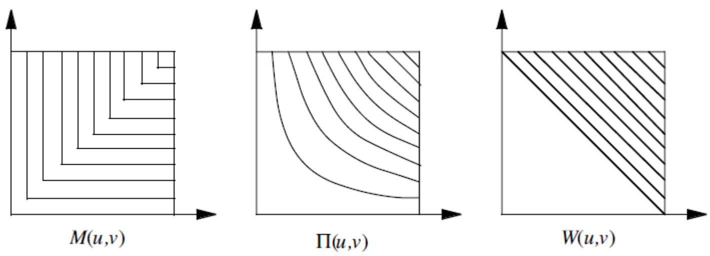

Figure 1.2: Contour diagrams of the copulas M, Π, and W.

Also note that, given any copula C, it follows from (1.7) that for a given t in I the graph of the level set $\{(u,v) \in I^2 / C(u,v) = t\}$ must lie in the shaded triangle in Figure 1.3, whose boundaries are the level sets determined by $M(u,v) = t$ and $W(u,v) = t$.

We conclude this subsection with the two theorems concerning the partial derivatives of copulas. The word "almost" is used in the sense Lebesgue measure.

Theorem 1.1.3. *Let C be a copula. For any v in I, the partial derivative $\frac{\partial C(u,v)}{\partial u}$ exists for almost all u, and for such v and u,*

$$0 \leq \frac{\partial C(u,v)}{\partial u} \leq 1 \tag{1.9}$$

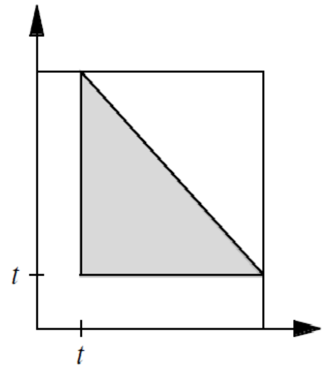

Figure 1.3: The region that contains the level set $\{(u,v) \in I^2 / C(u,v) = t\}$.

Similarly, for any u in I, the partial derivative $\frac{\partial C(u,v)}{\partial v}$ exists for almost all v, and for such u and v,

$$0 \leq \frac{\partial C(u,v)}{\partial v} \leq 1 \tag{1.10}$$

Furthermore, the functions $u \mapsto \frac{\partial C(u,v)}{\partial v}$ and $v \mapsto \frac{\partial C(u,v)}{\partial u}$ are defined and nondecreasing almost everywhere on I.

Theorem 1.1.4. *[Nelsen (2006)] Let C be a copula. If $\frac{\partial C(u,v)}{\partial v}$ and $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ are continous on I^2 and $\frac{\partial C(u,v)}{\partial u}$ exists for all $u \in (0,1)$ when $v = 0$, then*

$$\frac{\partial C(u,v)}{\partial u} \text{ and } \frac{\partial^2 C(u,v)}{\partial v \partial u} \text{ exist in } (0,1)^2 \text{ and } \frac{\partial^2 C(u,v)}{\partial u \partial v} = \frac{\partial^2 C(u,v)}{\partial v \partial u}.$$

1.1.1.3 Sklar's Theorem

The theorem in the title of this subsection is central to the theory of copulas and is the foundation of many, if not most, of the applications of that theory to statistics. Sklar's theorem elucidates the role that copulas play in the relationship between multivariate distribution functions and their univariate margins. Thus we begin this section with a short discussion of distribution functions.

Definition 1.1.6. *A distribution function is a function F with domain $\overline{\mathbb{R}}$ such that*

1. *F is nondecreasing,*

2. *$F(-\infty) = 0$ and $F(\infty) = 1$.*

Definition 1.1.7. *A joint distribution function is a function H with domain $\overline{\mathbb{R}}^2$ such that*

1. *H is 2-increasing,*

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

Thus H is grounded, and because $DomH = \overline{\mathbb{R}}^2$, H has margins F and G given by $F(x) = H(x, \infty)$ and $G(y) = H(\infty, y)$. By virtue of Corollary 1.1.1, F and G are distribution functions.

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

Theorem 1.1.5 (Sklar's Theorem)**.** *[Nelsen (2006)] Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all x, y in $\overline{\mathbb{R}}$,*

$$H(x, y) = C(F(x), G(y)). \tag{1.11}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $RanF \times RanG$. Conversely, if C is a copula and F and G are distribution functions, then the function H defined by (1.11) is a joint distribution function with margins F and G.

This theorem first appeared in (Sklar 1959). The name "copula" was chosen to emphasize the manner in which a copula "couples" a joint distribution function to its univariate margins. The argument that we give below is essentially the same as in (Schweizer and Sklar 1974). Also it is very important because it allows to associate with each a two-dimensional distribution copula. The equation (1.11) gives a canonical representation of the distribution function H, by bringing one side, the distributions F and G of unidimentionnelles distributions and on the other side, the copula which allows "be cemented" these margins; this copula expresses the dependence between the one-dimensional functions.

Thus, it requires two lemmas.

Lemma 1.1.4. *Let H be a joint distribution function with margins F and G. Then there exists a unique subcopula $C^{'}$ such that*

1. *$DomC^{'} = RanF \times RanG$,*

2. *For all x, y in $\overline{\mathbb{R}}$, $H(x, y) = C^{'}(F(x), G(y))$.*

Lemma 1.1.5. *[Nelsen (2006)] Let $C^{'}$ be a subcopula. Then there exists a copula C such that $C(u, v) = C^{'}(u, v)$ for all (u, v) in $DomC^{'}$; i.e., any subcopula can be extended to a copula. The extension is generally non-unique.*

Now, Using the theorem of Sklar, we can express the density of a random vector (X, Y) depending on the density of the copula and its margins F and G by the following definition :

Definition 1.1.8. *Let H be a joint distribution function absolutely continuous with margins F and G, and C the copula such that for all x, y in $\overline{\mathbb{R}}$:*

$$H(x, y) = C(F(x), G(y)).$$

Then we assume that the copula function $C(u, v)$ has a density $c(u, v)$ with respect to the Lebesgue measure on $[0, 1]^2$ in such a way that $c(u, v) := \frac{\partial^2 C(u,v)}{\partial u \partial v}$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u, v)$ and $c(u, v)$ are then the cumulative distribution function (c.d.f.) and density respectively of the transformed variables $(U, V) = (F(x), G(y))$. By differentiating formula of $H(x, y)$, we get for the joint density,

$$h(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial F(x)}{\partial x} \frac{\partial G(y)}{\partial y} \frac{\partial^2 C(F(x), G(y))}{\partial F(x) \partial G(y)} = f(x)g(y)c(F(x), G(y)),$$

Equation (1.11) gives an expression for joint distribution functions in terms of a copula and two univariate distribution functions. But (1.11) can be inverted to express copulas in terms of a joint distribution function and the "inverses" of the two margins. However, if a margin is not strictly increasing, then it does not possess an inverse in the usual sense. Thus we first need to define "quasi-inverses" of distribution functions (recall Definition 1.1.6).

Definition 1.1.9. *Let F be a distribution function. Then a quasi-inverse of F is any function $F^{(-1)}$ with domain I such that*

1. *if t is in $RanF$, then $F^{(-1)}(t)$ is any number x in $\overline{\mathbb{R}}$ such that $F(x) = t$, i.e., for all t in $RanF$,*
$$F(F^{(-1)}(t)) = t;$$
2. *if t is not in $RanF$, then*
$$F^{(-1)}(t) = \inf\left\{x/F(x) \geq t\right\} = \sup\left\{x/F(x) \leq t\right\}.$$

If F is strictly increasing, then it has but a single quasi-inverse, which is of course the ordinary inverse, for which we use the customary notation F^{-1}.

Using quasi-inverses of distribution functions, we now have the following corollary to Lemma 1.1.4.

Corollary 1.1.2. *[Nelsen (2006)] Let H, F, G, and $C^{'}$ be as in Lemma 1.1.4, and let $F^{(-1)}$ and $G^{(-1)}$ be quasi-inverses of F and G, respectively. Then for any (u, v) in $DomC^{'}$,*

$$C^{'}(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)). \tag{1.12}$$

When F and G are continuous, the above result holds for copulas as well and provides a method of constructing copulas from joint distribution functions.

1.1.1.4 The Fréchet-Hoeffding Bounds for Joint Distribution Functions

In Subsect. 1.1.1.2 we encountered the Fréchet-Hoeffding bounds as universal bounds for copulas, i.e., for any copula C and for all u, v in I,

$$W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v).$$

As a consequence of Sklar's theorem, if X and Y are random variables with a joint distribution function H and margins F and G, respectively, then for all x, y in $\overline{\mathbb{R}}$,

$$\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y)) \tag{1.13}$$

Because M and W are copulas, the above bounds are joint distribution functions and are called the Fréchet-Hoeffding bounds for joint distribution functions H with margins

F and G. Of interest in this section is the following question: What can we say about the random variables X and Y when their joint distribution function H is equal to one of its Fréchet-Hoeffding bounds?

To answer this question, we first need to introduce the notions of nondecreasing and nonincreasing sets in $\overline{\mathbb{R}}^2$.

Definition 1.1.10. *A subset S of $\overline{\mathbb{R}}^2$ is nondecreasing if for any (x, y) and (u, v) in $S, x < u$ implies $y \leq v$. Similarly, a subset S of $\overline{\mathbb{R}}^2$ is nonincreasing if for any (x, y) and (u, v) in $S, x < u$ implies $y \geq v$.*

The following figure illustrates a simple nondecreasing set.

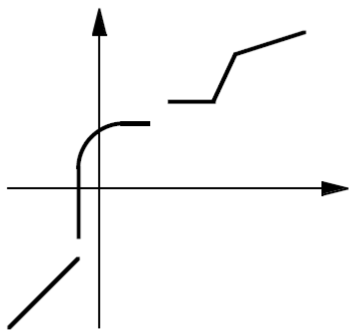

Figure 1.4: The graph of a nondecreasing set.

We will now prove that the joint distribution function H for a pair (X, Y) of random variables is the Fréchet-Hoeffding upper bound (i.e., the copula is M) if and only if the support of H lies in a nondecreasing set. The following proof is based on the one that appears in (Mikusiński, Sherwood and Taylor 1991-1992). But first, we need two lemmas:

Lemma 1.1.6. *Let S be a subset of $\overline{\mathbb{R}}^2$. Then S is nondecreasing if and only if for each (x, y) in $\overline{\mathbb{R}}^2$, either*

1.
$$for \;\; all \;\; (u, v) \;\; in \;\; S, u \leq x \;\; implies \;\; v \leq y; \;\; or \tag{1.14}$$

2.
$$for \;\; all \;\; (u, v) \;\; in \;\; S, v \leq y \;\; implies \;\; u \leq x. \tag{1.15}$$

Lemma 1.1.7. *Let X and Y be random variables with joint distribution function H. Then H is equal to its Fréchet-Hoeffding upper bound if and only if for every (x, y) in $\overline{\mathbb{R}}^2$, either*

$$P[X > x, Y \leq y] = 0 \text{ or } P[X \leq x, Y > y] = 0.$$

We are now ready to prove

Theorem 1.1.6. *Let X and Y be random variables with joint distribution function H. Then H is identically equal to its Fréchet-Hoeffding upper bound if and only if the support of H is a nondecreasing subset of $\overline{\mathbb{R}}^2$.*

Theorem 1.1.7 (Nelsen (2006))**.** *Let X and Y be random variables with joint distribution function H. Then H is identically equal to its Fréchet-Hoeffding lower bound if and only if the support of H is a nonincreasing subset of $\overline{\mathbb{R}}^2$.*

When X and Y are continuous, the support of H can have no horizontal or vertical line segments, and in this case it is common to say that "Y is almost surely an increasing function of X" if and only if the copula of X and Y is M; and "Y is almost surely a decreasing function of X" if and only if the copula of X and Y is W. If U and V are uniform $(0, 1)$ random variables whose joint distribution function is the copula M, then $P[U = V] = 1$; and if the copula is W, then $P[U + V = 1] = 1$.

Random variables with copula M are often called comonotonic, and random variables with copula W are often called countermonotonic.

1.1.2 Dependence

In this section, we explore ways in which copulas can be used in the study of dependence or association between random variables. As Jogdeo (1982) notes,

Dependence relations between random variables is one of the most widely studied subjects in probability and statistics. The nature of the dependence can take a variety of forms and unless some specific assumptions are made about the dependence, no meaningful statistical model can be contemplated.

There are a variety of ways to discuss and to measure dependence. As we shall see, many of these properties and measures are, in the words of Hoeffding (1940, 1941), "scale-invariant", that is, they remain unchanged under strictly increasing transformations of the random variables. As we noted in the Introduction, "*...it is precisely the copula which* *captures those properties of the joint distribution which are invariant under almost surely strictly increasing transformations*" (Schweizer and Wolff 1981). The focus of this section is an exploration of the role that copulas play in the study of dependence.

Dependence properties and measures of association are interrelated, and so there are many places where we could begin this study. Because the most widely known scale-invariant measures of association are the population versions of Kendall's tau and Spearman's rho, both of which "measure" a form of dependence known as concordance, we will begin there.

A note on terminology: we shall reserve the term "correlation coefficient" for a measure of the linear dependence between random variables (e.g., Pearson's product-moment correlation coefficient) and use the more modern term "measure of association" for measures such as Kendall's tau and Spearman's rho.

1.1.2.1 Concordance

Informally, a pair of random variables are concordant if "large" values of one tend to be associated with "large" values of the other and "small" values of one with "small" values of the other. To be more precise, let (x_i, y_i) and (x_j, y_j) denote two observations from a vector (X, Y) of continuous random variables. We say that (x_i, y_i) and (x_j, y_j) are concordant if $x_i < x_j$ and $y_i < y_j$, or if $x_i > x_j$ and $y_i > y_j$. Similarly, we say that (x_i, y_i) and (x_j, y_j) are discordant if $x_i < x_j$ and $y_i > y_j$ or if $x_i > x_j$ and $y_i < y_j$. Note the alternate formulation: (x_i, y_i) and (x_j, y_j) are concordant if $(x_i - x_j)(y_i - y_j) > 0$ and discordant if $(x_i - x_j)(y_i - y_j) < 0$.

1.1.2.2 Kendall's tau

The sample version of the measure of association known as Kendall's tau is defined in terms of concordance as follows (Kruskal 1958; Hollander and Wolfe 1973; Lehmann 1975): Let $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ denote a random sample of n observations from a vector (X, Y) of continuous random variables. There are (n,2) distinct pairs (x_i, y_i) and (x_j, y_j) of observations in the sample, and each pair is either concordant or discordant-let c denote the number of concordant pairs and d the number of discordant pairs. Then Kendall's tau for the sample is defined as

$$\tau = \frac{c - d}{c + d} = (c - d)/(n, 2) \tag{1.16}$$

Equivalently, τ is the probability of concordance minus the probability of discordance for a pair of observations (x_i, y_i) and (x_j, y_j) that is chosen randomly from the sample. The population version of Kendall's tau for a vector (X, Y) of continuous random variables with joint distribution function H is defined similarly. Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random vectors, each with joint distribution function H. Then the population version of Kendall's tau is defined as the probability of concordance minus the probability of discordance:

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{1.17}$$

In order to demonstrate the role that copulas play in concordance and measures of association such as Kendall's tau, we first define a "concordance function" Q, which is the difference of the probabilities of concordance and discordance between two vectors (X_1, Y_1) and (X_2, Y_2) of continuous random variables with (possibly) different joint distributions H_1 and H_2 , but with common margins F and G. We then show that this function depends on the distributions of (X_1, Y_1) and (X_2, Y_2) only through their copulas.

Theorem 1.1.8. *Let (X_1, Y_1) and (X_2, Y_2) be independent vectors of continuous random variables with joint distribution functions H_1 and H_2, respectively, with common margins F (of X_1 and X_2) and G (of Y_1 and Y_2). Let C_1 and C_2 denote the copulas of (X_1, Y_1) and (X_2, Y_2), respectively, so that $H_1(x, y) = C_1(F(x), G(y))$ and $H_2(x, y) = C_2(F(x), G(y))$. Let Q denote the difference between the probabilities of concordance and discordance of (X_1, Y_1) and (X_2, Y_2), i.e., let*

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{1.18}$$

Then

$$Q = Q(C_1, C_2) = 4 \int\int_{I^2} C_2(u, v) dC_1(u, v) - 1. \tag{1.19}$$

Because the concordance function Q in Theorem 1.1.8 plays an important role throughout this subsection, we summarize some of its useful properties in the following corollary.

Corollary 1.1.3. *[Nelsen (2006)] Let C_1, C_2, and Q be as given in Theorem 1.1.8. Then*

1. *Q is symmetric in its arguments: $Q(C_1, C_2) = Q(C_2, C_1)$.*

2. Q is nondecreasing in each argument: if $C_1 \prec C_1'$ and $C_2 \prec C_2'$ for all (u,v) in I^2, then $Q(C_1, C_2) \leq Q(C_1', C_2')$.

3. Copulas can be replaced by survival copulas in Q, i.e., $Q(C_1, C_2) = Q(\hat{C}_1, \hat{C}_2)$.

Theorem 1.1.9. *[Nelsen (2006)] Let X and Y be continuous random variables whose copula is C. Then the population version of Kendall's tau for X and Y (which we will denote by either τ_{XY}, or τ_C) is given by*

$$\tau_{X,Y} = \tau_C = Q(C,C) = 4\int\int_{I^2} C(u,v)dC(u,v) - 1. \tag{1.20}$$

Thus Kendall's tau is the first "concordance axis" in the figure below. Note that the integral that appears in (1.20) can be interpreted as the expected value of the function $C(U,V)$ of uniform $(0,1)$ random variables U and V whose joint distribution function is C, i.e.,

$$\tau_C = 4E(C(U,V)) - 1. \tag{1.21}$$

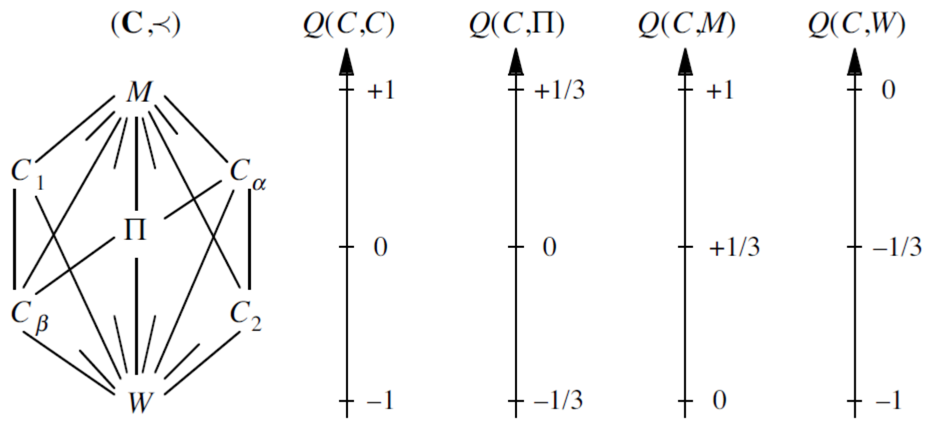

Figure 1.5: The partially ordered set $(C,\prec)$ and several "concordance axes".

When the copula C is a member of a parametric family of copulas (e.g., if C is denoted C_θ or $C_{\alpha,\beta}$,), we will write τ_θ and $\tau_{\alpha,\beta}$, rather than τ_{C_θ} and $\tau_{C_{\alpha,\beta}}$, respectively.

In general, evaluating the population version of Kendall's tau requires the evaluation of the double integral in (1.20). For an Archimedean copula, the situation is simpler, in that Kendall's tau can be evaluated directly from the generator of the copula, as shown in the following corollary (Genest and MacKay 1986a,b). Indeed, one of the reasons that Archimedean copulas are easy to work with is that often expressions with a one-place function (the generator) can be employed rather than expressions with a two-place function (the copula).

Corollary 1.1.4. *[Nelsen (2006)] Let X and Y be random variables with an Archimedean copula C generated by φ in Ω. The population version τ_C of Kendall's tau for X and Y is given by*

$$\tau_C = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt. \tag{1.22}$$

1.1.2.3 Spearman's rho

As with Kendall's tau, the population version of the measure of association known as Spearman's rho is based on concordance and discordance. To obtain the population version of this measure (Kruskal 1958; Lehmann 1966), we now let $(X_1, Y_1), (X_2, Y_2)$, and (X_3, Y_3) be three independent random vectors with a common joint distribution function H (whose margins are again F and G) and copula C. The population version rXY, of Spearman's rho is defined to be proportional to the probability of concordance minus the probability of discordance for the two vectors (X_1, Y_1) and (X_2, Y_3) i.e., a pair of vectors with the same margins, but one vector has distribution function H, while the components of the other are independent:

$$r_{X,Y} = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]) \tag{1.23}$$

(the pair (X_3, Y_2) could be used equally as well). Note that while the joint distribution function of (X_1, Y_1) is $H(x, y)$, the joint distribution function of (X_2, Y_3) is $F(x)G(y)$ (because X_2 and Y_3 are independent).

Theorem 1.1.10. *Let X and Y be continuous random variables whose copula is C. Then the population version of Spearman's rho for X and Y (which we will denote by either ρ_{XY}, or ρ_C) is given by*

$$\rho_{XY} = \rho_C = 3Q(C, \Pi), \tag{1.24}$$

$$= 12\int\int_{I^2} uvdC(u, v) - 3 \tag{1.25}$$

$$= 12\int\int_{I^2} C(u, v)dudv - 3 \tag{1.26}$$

Thus Spearman's rho is essentially the second "concordance axis" in the figure above. The coefficient "3" that appears in (1.23) and (1.24) is a "normalization" constant, because $Q(C, \Pi) \in [-1/3, 1/3]$. As was the case with Kendall's tau, we will write ρ_θ and $\rho_{\alpha,\beta}$, rather than ρ_{C_θ} and $\rho_{C_{\alpha,\beta}}$, respectively, when the copula C is given by C_θ or $C_{\alpha,\beta}$,

Any set of desirable properties for a "measure of concordance" would include those in the following definition (Scarsini 1984).

Definition 1.1.11. *A numeric measure κ of association between two continuous random variables X and Y whose copula is C is a measure of concordance if it satisfies the following properties (again we write κ_{XY}, or κ_C when convenient):*

1. *κ is defined for every pair X, Y of continuous random variables;*

2. *$-1 \leq \kappa_{X,Y} \leq 1, \kappa_{X,X} = 1$, and $\kappa_{X,-X} = -1$;*

3. *$\kappa_{X,Y} = \kappa_{Y,X}$;*

4. *if X and Y are independent, then $\kappa_{X,Y} = \kappa_\Pi = 0$;*

5. *$\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;*

6. *if C_1 and C_2 are copulas such that $C_1 \prec C_2$, then $\kappa_{C_1} \leq \kappa_{C_2}$;*

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

As a consequence of Definition 1.1.11, we have the following theorem.

Theorem 1.1.11. *[Nelsen (2006)] Let κ be a measure of concordance for continuous random variables X and Y:*

1. *if Y is almost surely an increasing function of X, then*

$$\kappa_{X,Y} = \kappa_M = 1;$$

2. if Y is almost surely a decreasing function of X, then

$$\kappa_{X,Y} = \kappa_W = -1;$$

3. if a and b are almost surely strictly monotone functions on $RanX$ and $RanY$, respectively, then

$$\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}.$$

In the next theorem, we see that both Kendall's tau and Spearman's rho are measures of concordance according to the above definition.

Theorem 1.1.12. *If X and Y are continuous random variables whose copula is C, then the population versions of Kendall's tau (1.20) and Spearman's rho (1.24, 1.25 and 1.26) satisfy the properties in Definition 1.1.11 and Theorem 1.1.11 for a measure of concordance.*

The fact that measures of concordance, such as r and t, satisfy the sixth criterion in Definition 1.1.11 is one reason that "$\prec$" is called the *concordance ordering.*

Spearman's rho is often called the "grade" correlation coefficient. Grades are the population analogs of ranks that is, if x and y are observations from two random variables X and Y with distribution functions F and G, respectively, then the grades of x and y are given by $u = F(x)$ and $v = G(y)$. Note that the grades (u and v) are observations from the uniform $(0, 1)$ random variables $U = F(X)$ and $V = G(Y)$ whose joint distribution function is C. Because U and V each have mean 1/2 and variance 1/12, the expression for ρ_C in (1.25) can be re-written in the following form:

$$\begin{aligned} \rho_{X,Y} &= \rho_C = 12 \int\int_{I^2} uvdC(u,v) - 3 = 12E(UV) \\ &= \frac{E(UV)-1/4}{1/2} = \frac{E(UV)-E(U)E(V)}{\sqrt{VarU}\sqrt{VarV}}. \end{aligned}$$

As a consequence, Spearman's rho for a pair of continuous random variables X and Y is identical to Pearson's product-moment correlation coefficient for the grades of X and Y , i.e., the random variables $U = F(X)$ and $V = G(Y)$.

1.1.2.4 The Relationship between Kendall's tau and Spearman's rho

Although both Kendall's tau and Spearman's rho measure the probability of concordance between random variables with a given copula, the values of ρ and τ are often quite different. In this subsection, we will determine just how different ρ and τ can be.

The next theorem1.1.13, due to Daniels (1950), gives universal inequalities for these measures. For the proof see Kruskal (1958).

Theorem 1.1.13. *[Nelsen (2006)] Let X and Y be continuous random variables, and let τ and ρ denote Kendall's tau and Spearman's rho, defined by (1.17) and (1.23), respectively. Then*

$$-1 \leq 3\tau - 2\rho \leq 1. \tag{1.27}$$

The next theorem gives a second set of universal inequalities relating ρ and τ. It is due to Durbin and Stuart (1951); and again the proof is adapted from Kruskal (1958):

Theorem 1.1.14. *[Nelsen (2006)] Let X, Y, τ, and ρ be as in Theorem 1.1.12. Then*

$$\frac{1+\rho}{2} \geq \left(\frac{1+\tau}{2}\right)^2 \tag{1.28}$$

and

$$\frac{1-\rho}{2} \geq \left(\frac{1-\tau}{2}\right)^2 \tag{1.29}$$

The inequalities in the preceding two theorems combine to yield

Corollary 1.1.5. *Let X, Y, τ, and ρ be as in Theorem 1.1.12. Then*

$$\frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2}, \tau \geq 0, \tag{1.30}$$

and

$$\frac{\tau^2 - 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2}, \tau \leq 0,$$

1.1.2.5 Why the copula and not the correlation coefficient?

Undoubtedly, the dependence between random variables plays a very important role in many areas of mathematics. It is a widely studied topics in probability and statistics. A huge variety of this concept has been studied by many authors to propose definitions and useful properties with applications. A measure of dependence is regularly used the linear correlation by Bravais Pearson (1896) ; This correlation measures the linear relationship between two random variables X and Y, and can take any value from the interval

$[-1, 1]$. The linear correlation coefficient is a measure of dependence easily calculated. This indicator is efficient when the dependence is linear and the universe considered Gaussian. It is very useful for families of elliptic distributions (because these distributions for non correlation implies independence). However, this dependence measure often used by practitioners has several limitations; we cite some problems related to this concept:

- The correlation coefficient is undefined if the moments of order two random variables are not finished. This is not an appropriate measure of dependence for heavy-tailed distributions where the variances can be finished.
- It is easy to construct examples where the linear correlation coefficient of Pearson is not invariant under strictly increasing transformations; for example the correlation between two random variables X and Y is not the same between $\log(X)$ and $\log(Y)$; Indeed, changes affect the feedback data correlations .
- The correlation is simply a scalar measured dependence; it can not tell us everything we want to know about the structure of dependence.
- The absolute positive dependence is not necessarily a correlation of $+1$; same negative dependency perfect not necessarily a correlation of -1.

In finance, the Gaussian case is rarely used; to remedy this we use other indicators of dependence based on concordance and discordance observed in a sample. We then use the coefficients of non-linear and non-parametric correlation, such as Kendall's tau or Spearman's rho. These are good indicators of the overall dependence between random variables. In addition, they are between -1 and $+1$, as the linear correlation coefficient, a value $+1$ means a perfect match.

Measure the length using statistical indicators is one thing, the model by a function of dependence is another. Copula meets this objective. Indicators of dependence (linear correlation, Kendall's tau and Spearman's rho) can be defined in this framework from the parameters of the copula (when it is parametric).

The tool is relatively copula innovative modeling the dependency structure of several random variables. Knowledge of this statistical tool is essential to understanding many application areas of quantitative finance: measurement multiple credit risk assessment of structured credit products, replicating the performance of hedge funds, measure multiple risk market, portfolio management using Monte Carlo simulations,... . Thus, whenever it is necessary to model the dependence structure of several random variables, we can use the copula. The relationship between the joint distribution and the copula allows us to study the dependence structure of (X, Y) separately from their marginal.

- Instead of summarizing the structure of dependence by a single number as the linear correlation coefficient, we can use a model that reflects a more detailed knowledge on risk management issues that we deal with.

- The copula is a multivariate distribution function that helps us to understand the multivariate risk factor data, and then find the marginal models for the various individual risk factors and copula models for their dependence structure.

- As an appropriate model we have a wide range of families of copulas which you can select; This allows us to choose a particular family of copulas as random variables of multivariate data that we are trying to model.

- If the marginal distributions are known, the copula can be used to suggest an appropriate form for the joint distribution; this means that we can create functions and marginal distributions we can extract the copula functions from well-known multivariate distributions.

- Sklar's theorem is "powerful" multivariate copula analysis tool, as it allow to build models of multivariate distributions compatible with the one-dimensional marginal models, this compatibility is often very important in financial modeling (for models estimate of the value at risk VaR).

- The copula solves another problem: the development of non-Gaussian models. The family of non-Gaussian distributions is not only small but powerful; the disadvantage is that the margins are identical. But with the copula, one can construct such a distribution with a Gaussian marginal and marginal uniform or Gaussian inverse...

- The function of multivariate distribution carries more information than the different marginal distributions and this generally helps us to avoid the disadvantages of correlation as a measure of dependence.

- Copulas are a way to test to extract the dependence structure from the joint distribution function and separate dependence and marginal behavior.

- It allows us to make possible natural extensions of some results obtained in the univariate case to the multivariate case. Multidimensional distributions are obtained benefit in line with reality, especially in the use of financial statstiques.

1.1.3 Exemples Copulas

There are several methods for the construction of the copula which can be found in the literature (Joe (1997), Nelsen (2006)); in this subsection we will present some of the most used copula. These copulas are of great interest for risk management because they help build parametric or semi-parametric models.

1.1.3.1 Gaussian copula

Definition 1.1.12. *Bivariate Gaussian copula is defined as follows:*

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

Where ρ is the correlation coefficient and Φ_ρ is the standard normal distribution bivariate correlation ρ.

Figures 1.6 and 1.7 to visualize the density of the Gaussian copula for different values of the linear correlation coefficient and the density of a bivariate Gaussian distribution, which is, remember, composed of Gaussian margins and a copula Gaussian.

We have the following expression of Kendall's tau:

$$\tau = \frac{2}{\pi} \arcsin \rho$$

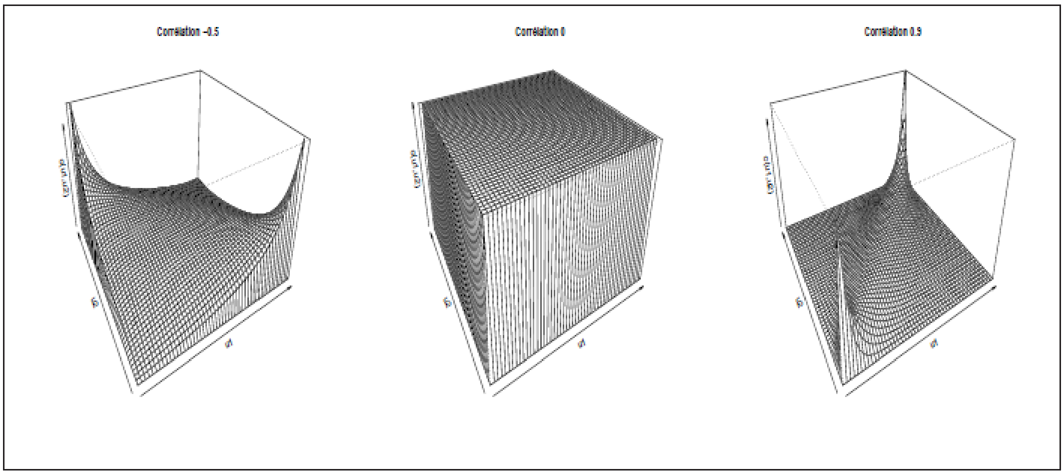

Figure 1.6: Density bivariate of three gaussian copulas for different values of ρ.

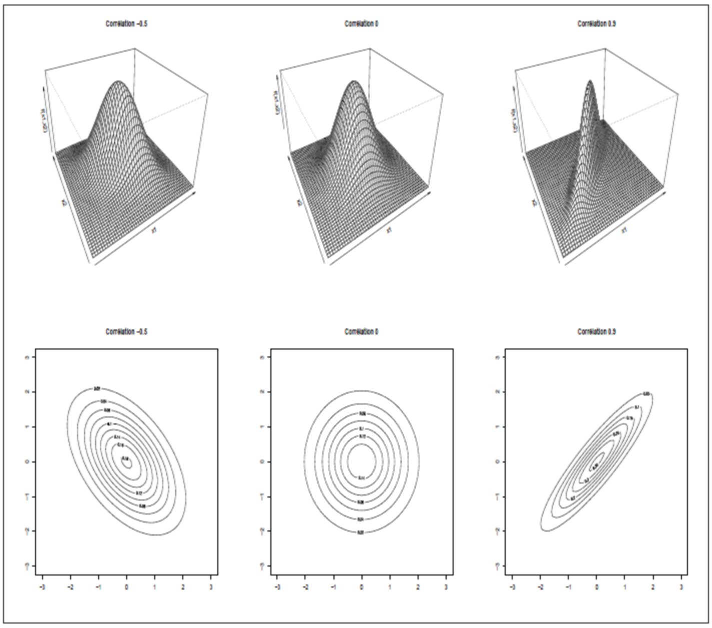

Figure 1.7: Density of three bivariate laws built from Gaussian centered reduced margins and a Gaussian copula for different values of ρ with their curves corresponding levels.

1.1.3.2 Archimedean copulas

Definition 1.1.13. *The Archimedean copulas are defined as follows:*

$$C(u_1u_2) = \begin{cases} \varphi^{-1}(\varphi(u_1), \varphi(u_2)), & if \;\; \varphi(u_1) + \varphi(u_2) \leq \varphi(0); \\ 0, & otherwise. \end{cases}$$

With φ the generating function of the copula, checking

$$\varphi(1) = 0, \varphi^{'}(u) < 0 \;\; and \;\; \varphi^{''}(u) > 0 \;\; for \;\; all \;\; 0 \leq u \leq 1$$

Kendall's tau τ is equal to the Archimedean copulas to:

$$1 + 4\int_0^1 \frac{\varphi(u)}{\varphi'(u)} du$$

Density of three bivariate laws built from Gaussian centered reduced margins and a Gaussian copula for different values of rho with their curves corresponding levels

Some examples of bivariate Archimedean copulas are noting $\tilde{u} = -\ln u$

Nom	Générateur	Copule bivariée
Clayton ($\theta > 0$)	$u^{-\theta} - 1$	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$
Gumbel ($\theta \geq 1$)	$(-\ln u)^{\theta}$	$\exp(-(\tilde{u}_1^{\theta} + \tilde{u}_2^{\theta})^{1/\theta})$
Frank ($\theta \neq 0$)	$-\ln \dfrac{e^{-\theta u} - 1}{e^{-\theta} - 1}$	$-\dfrac{1}{\theta} \ln \left(1 + \dfrac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$

Table 1.1: Examples bivariate Archimedean copula.

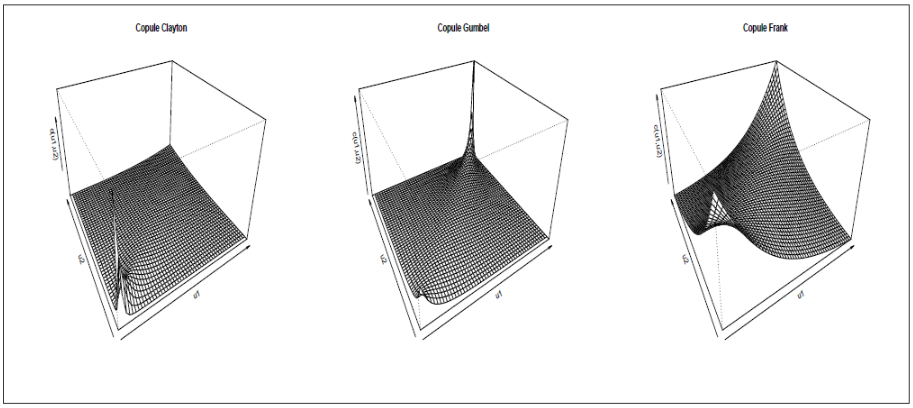

Figure 1.8: Density of three Archimedean copulas with parameter $\theta = 3$.

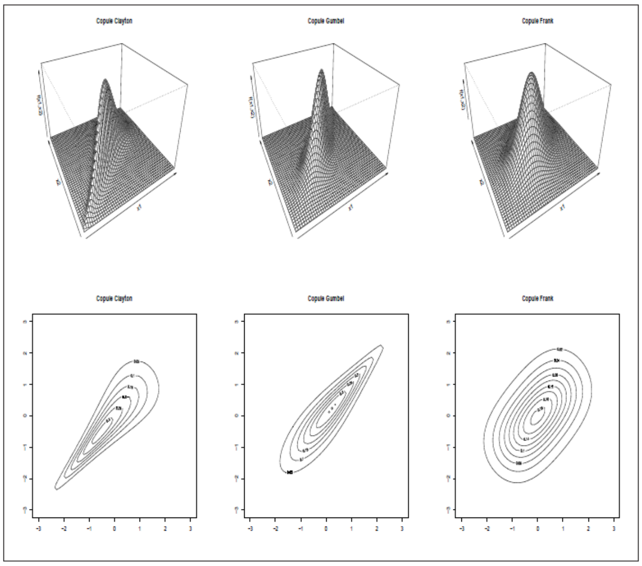

Figure 1.9: Three bivariate densities constructed from Gaussian centered reduced margins laws and Archimedean copula with parameter $\theta = 3$, with their curves corresponding levels.

1.2 Description of the thesis

The objective of this thesis is devoted to the estimation of a regression model by a function of copulae and our work is divided into three chapters and is organized as follows:

In Chapter 1, we first give a general introduction as well as definitions and tools; We start with the basic properties of copulas and then proceeds to present the copula construction methods and examine the role played by copulas in the modeling and the study of addiction. The emphasis is on bivariate copulas, we indicate Sklar's theorem that illuminates the role of copulas in the relationship between the distribution functions of two variables and their univariate marginal.(This chapter was taken from Nelsen (2006) with some slight modifications.).

In chapter 2, we estimate the regression using the copula function , we present our model, and then we study its asymptotic properties and obtain, from classical results of convergence of kernel estimators of regression and under the usual regularity conditions on the densities and kernels, the following results:

- Point consistency in probability,

  Theorem: Let the regularity assumptions (i)-(vii) on the density and the kernel be satisfied, if h_n tends to zero as $n \to \infty$ in such a way that

$$na_n^4 \to \infty, \frac{\sqrt{\ln \ln n}}{na_n^3} \to 0,$$

  then,

$$\hat{r}_n(x) = r(x) + O_P\left(h_n^2 + \frac{1}{\sqrt{nh_n^2}} + \frac{1}{nh_n^4} + \frac{\sqrt{\ln \ln n}}{nh_n^3}\right)$$

  Corrolary: We get the rate of convergence, by choosing the bandwidth which balance the bias and variance trade-off: for an optimal choice of $h_n \simeq n^{-1/6}$, we get

$$\hat{r}_n(x) = r(x) + O_P(n^{-1/3})$$

- Point almost sure consistency,

  Theorem: Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied.If the bandwidth h_n tends to zero as $n \to \infty$ in such a way that

$$\frac{\sqrt{\ln n \ln\ln n}}{nh_n^3} \to 0, \frac{\ln\ln n}{nh_n^4} \to 0,$$

then,

$$\hat{r}_n(x) = r(x) + O_{a.s}\left(h_n^2 + \sqrt{\frac{\ln\ln n}{nh_n^2}} + \frac{\ln\ln n}{nh_n^4} + \frac{\sqrt{\ln n \ln\ln n}}{nh_n^3}\right)$$

Corollary: For $h_n \simeq (\ln\ln n/n)^{1/6}$ which is the optimal trade-off between the bias and the stochastic term, one gets the optimal rate

$$\hat{r}_n(x) = r(x) + O_{a.s}\left(\frac{\ln\ln n}{n}\right)^{1/3}.$$

In Chapter 3, we present a note on the asymptotic normality of the regression model by a copula function for this; we introduce the model, and then we make some notations and assumptions of regularity for our main result contained in the last part of this note.

Thus, one has

Corollary: Consider the model (3.6). If the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied,then,

1.
$$\sqrt{nh}\,(\hat{r}(x) - r(x)) \rightsquigarrow \mathcal{N}\left(0, \frac{(\phi(x) - r^2(x))}{f(x)}\|K\|_2^2\right)$$

2.
$$B_o = E(\hat{r}(x)) - r(x) = B(x)h^2 + o(h^2),$$

and

$$V_0 = Var(\hat{r}(x)) = V(x)\frac{1}{nh} + o\left(\frac{1}{nh}\right)$$

with

$$B(x) = \frac{\int t^2K(t)dt(g^{(2)}(x) - r(x)f^{(2)})}{2f(x)},$$

$$V(x) = \int K^2(t)dt \frac{(\phi(x) - r^2(x))}{f(x)}$$

and $g(x) = \frac{1}{n}\sum Y_i K\left(\frac{x-X_i}{h}\right), f(x) = \frac{1}{n}\sum K\left(\frac{x-X_i}{h}\right), \phi(x) = E(Y^2/X = x)$

These results are then extended to compact $\mathbb{R}$, in the following theorem:

Theorem: Let the regularity assumptions (i)-(vii) on the density and the kernel be satisfied, if $h_n \simeq (\ln n/n)^{1/6}$ then,

$$\sup_{x\in\mathbb{R}} |\hat{r}_n(x) - r(x)| = O_P\left(\left(\frac{\ln n}{n}\right)^{1/3}\right)$$

and

$$\sup_{x\in\mathbb{R}} |\hat{r}_n(x) - r(x)| = O_{a.s}\left(\left(\frac{\ln n}{n}\right)^{1/3}\right)$$

Finally we conclude this thesis by giving some perspectives for future research and appendix consisting of classical results and tools used in this thesis.

Bibliography

Beneš V, Štěpn J., eds *Distributions with Given Marginals and Moment Problems.* Kluwer, Dordrecht(1997)

Conway. DA, *Multivariate Distributions with Specified Marginals.* (Technical Report 145, Stanford University)(1979)

Conway DA "Plackett family of distributions". *In: Kotz S, Johnson NL (eds) Encyclopedia of Statistical Sciences*, Vol 7. Wiley, New York, (1986) pp 1-5

Cuadras CM, Fortiana J, Rodríguez-Lallena JA eds *Distributions with Given Marginals and Statistical Modelling.* Kluwer, Dordrecht(2002)

Dall'Aglio, G., Fréchet classes: the beginnings". *In: Dall'Aglio G*, Kotz S, Salinetti, G.,(eds) "Advances in Probability Distributions with Given Marginals.", Kluwer, Dordrecht,(1991) pp 1-12

Daniels, HE., "Rank correlation and population models.", (1950), *J Roy Statist Soc Ser B*, 12:171-181

Darsow, WF., Nguyen B, Olsen, ET., "Copulas and Markov processes.", *Illinois J Math*, (1992),36:600-642

Deheuvels, P., "Caractérisation complète des lois extrêmes multivariées et de la convergence des types extremes". *Publ Inst Statist Univ Paris* (1978) 23:1-37

Deheuvels, P., "La fonction de dépendence empirique et ses propriétés. Un test non paramétrique d'indépendence.", *Acad Roy Belg Bul Cl Sci* (5) (1979), 65:274-292

Deheuvels, P., "A Kolmogorov-Smirnov type test for independence and multivariate samples.", *Rev Roumaine Math Pures Appl*,(1981a), 26:213-226

Deheuvels, P., "A non parametric test for independence." *Publ Inst Statist Univ Paris*, (1981b),26:29-50

Deheuvels, P., "Multivariate tests of independence." In: Analytical Methods in Probability (Oberwolfach, 1980) *Lecture Notes in Mathematics 861. Springer*, Berlin,(1981c), pp 42-50

Durbin J, "Stuart A Inversions and rank correlations"., *J Roy Statist Soc Ser B*, (1951), 13:303-309

Fan, J., and Yao, Q. *Nonlinear time series*, second ed. Springer Series in Statistics. Springer-Verlag, New York, (2005). Nonparametric and parametric methods.

Féron, R., " Sur les tableaux de corrélation dont les marges sont données, cas de l'espace à trois dimensions". *Publ Inst Statist Univ Paris.*, (1956), 5:3-12

Fisher NI, Sen PK eds *The Collected Works of Wassily Hoeffding.* Springer, New York, (1994)

Fisher, N.I., *Copulas*, Encyclopedia of Statistical Sciences, Updated Volume 1, S. Kotz, C. B. Read, D. L. Banks (eds.), pp. 159-164. John Wiley and Sons, New York (1997).

Fisher NI, Sen PK eds (1994) The Collected Works of Wassily Hoeffding. Springer, New York

Galambos, J., *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York (1978)

Genest, C., MacKay, J., "Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données." *Canad J Statist* (1986a) 14:145-159

Genest, C., MacKay, J., "The joy of copulas: Bivariate distributions with uniform marginals." *Amer Statist*(1986b) 40:280-285

Hoeffding, W., "Masstabinvariante Korrelationstheorie.", *Schriften des Matematischen Instituts und des Instituts für Angewandte Mathematik* der Universität Berlin 5 Heft 3:179-233 [Reprinted as Scale-invariant correlation theory. In: Fisher NI, Sen PK (eds) The Collected Works of Wassily Hoeffding. Springer, New York, (1940), pp 57-107]

Hoeffding, W., "Masstabinvariante Korrelationsmasse für diskontinuierliche Verteilungen."; *Arkiv für matematischen Wirtschaften und Sozialforschung* 7:49-70 [Reprinted as Scale-invariant correlation measures for discontinuous distributions. In: Fisher NI, Sen PK (eds) The Collected Works of Wassily Hoeffding. Springer, New York, (1941), pp 109-133]

Hollander,. M., Wolfe, DA., *Nonparametric Statistical Methods.* Wiley, New York (1973)

Joe, H., "Parametric families of multivariate distributions with given margins.", *J Multivariate Anal*, (1993), 46:262-282

Joe, H., *Multivariate Models and Dependence Concepts*, Chapman and Hall, London (1997).

Jogdeo, K., "Concepts of dependence". *In: Kotz S, Johnson NL (eds) Encyclopedia of Statistical Sciences*, Vol 1. Wiley, New York,(1982), pp 324-334

Kamiński, A., Sherwood, H., Taylor, MD., "Doubly stochastic measures with mass on the graphs of two functions.", *Real Anal Exchange*, (1987-88), 13:253-257

Kimeldorf, G., Sampson, A., "Uniform representations of bivariate distributions"., *Comm Statist A-Theory Methods*, (1975b),4:617-627

Kruskal, WH., "Ordinal measures of association.", *J Amer Statist Assoc*, (1958), 53:814-861

Lehmann, EL., "Some concepts of dependence.", *Ann Math Statist*, (1966), 37:1137-1153.

Lehmann, EL., *Nonparametrics: Statistical Methods Based on Ranks.*, Holden-Day, San Francisco, (1975)

Lounas,Fadhila., *Modèlisation de la dépendance par les copules et applications*, Mémoire de Magister Université Mouloud Mammeri, Tizi-Ouzou.(2011).

Mardia, KV., *Families of Bivariate Distributions.*, Griffin, London, (1967) Marshall, AW., Olkin, I., "Families of multivariate distributions.", *J Amer Statist Assoc*,(1988), 83:834-841

Menger, K., "Statistical metrics". *Proc Natl Acad Sci USA*, (1942),28:535-537

Mikusiński, P., Sherwood, H., Taylor, MD., "Probabilistic interpretations of copulas and their convex sums". *In: Dall'Aglio G, Kotz S, Salinetti G* (eds) "Advances in Probability Distributions with Given Marginals". Kluwer, Dordrecht,(1991), pp 95-112

Mikusiński, P., Sherwood, H., Taylor, MD., "The Fréchet bounds revisited.", *Real Anal Exchange*, (1991-92), 17:759-764

Mikusiński, P., Sherwood, H., Taylor, MD., "Shuffles of Min". *Stochastica*, (1992), 13:61-74

Moore, DS., Spruill, MC., "Unified large-sample theory of general chi-squared statistics for tests of fit". *Ann Statist*,(1975), 3:599-616

Nelsen, R.B., *An Introduction to Copulas*, 2nd edition. Springer-Verlag, New York (2006).

Rüschendorf, L., Schweizer, B., Taylor, MD., eds *Distributions with Fixed Marginals and Related Topics*. Institute of Mathematical Statistics, Hayward, CA, (1996)

Scarsini, M., "On measures of concordance"., *Stochastica*, (1984), 8:201-218

Schweizer, B., Sklar, A., "Operations on distribution functions not derivable from operations on random variables". *Studia Math* (1974) 52:43-52

Schweizer, B., Sklar, A., *Probabilistic Metric Spaces*. North-Holland, New York (1983)

Schweizer, B., Wolff, EF., "On nonparametric measures of dependence for random variables"., *Ann Statist*, (1981),9:879-885

Schweizer, B., "Thirty years of copulas". *In: Dall'Aglio G, Kotz S, Salinetti G (eds) Advances in Probability Distributions with Given Marginals*. Kluwer, Dordrecht,(1991) pp 13-50

Seeley, RT., "Fubini implies Leibniz implies $F_{yx} = F_{xy}$". Am Math Monthly, (1961), 68:56-57

Sherwood, H., Taylor, MD., "Doubly stochastic measures with hairpin support". *Probab Theory Related Fields*,(1988), 78:617-626

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

Sklar, A., "Random variables, joint distributions, and copulas". *Kybernetica*, (1973), 9:449-460.

Chapter 2

Estimation using copula function in regression model

This chapter is the subject of an article published in

Mathematical Sciences and Applications E-Notes Volume 2 No. 1 pp. 105 115 (2014) MSAEN

Date: Received: October 6, 2013; Accepted: April 17, 2014.

Estimation using copula function in regression model

Djamila BENNAFLA, Abbes RABHI, Amina Angelika BOUCHENTOUF

Stochastic Models, Statistics and Applications Laboratory,Moulay Tahar University, Saida
E-mail: dbennafla@yahoo.fr
Mathematics Laboratory, Djillali Liabes University, Sidi Bel Abbes
E-mail: rabhi_abbes@yahoo.fr
Mathematics Laboratory, Djillali Liabes University, Sidi Bel Abbes
E-mail: bouchentouf_amina@yahoo.fr

Abstract

Copula models are becoming an increasingly powerful tool for modeling the dependencies between random variables, they have useful applications in many fields such as biostatistics, actuarial science, and finance. In this paper, we investigate the estimating of a regression model, by use of the copula representation. We study its asymptotic properties; the convergence almost surely as the convergence rate.

Keywords: *62N02, 62H12, 62J05.*
2000 MSC No: Copulas, nonparametric estimation, regression model.

2.1 Introduction

Copula theory, following the works of Sklar in 1959, allows a flexible modeling of dependence between two or more random variables. In recent years, the growing interest for this theory is phenomenal. In [18] Thomas Mikosch stated that in September 2005, a Google search on the term "copula" produced 650,000 results. Then, in January 2007, this same query generates more than 1.13 million. Given the number of publications in scientific journals and the number of papers available on Internet, it is undeniable that passion to the copula theory is still booming.

The progress of applications of this theory is wide in the field of finance, risk management, performance evaluation of assets, the valuation of derivatives, the extreme value theory, contagion require flexible and practical models of addiction.

The construction and properties of copulas have been studied rather extensively during the last 15 years or so. Hutchinson and Lai (1990) [15] were among the early authors who popularized the study of copulas. Nelsen (1999) [20] presented a comprehensive treatment of bivariate copulas, while Joe (1997) [16] devoted a chapter of his book to multivariate copulas. Further authoritative updates on copulas are given in Nelsen (2006) [19]. Copula methods have many important applications in insurance and finance Cherubini *et al.* (2004) [3] and Embrechts *et al.* (2003) [6].

Briefly speaking, copulas are functions that join multivariate distributions to their one-dimensional marginal distribution functions. Equivalently, copulas are multivariate distributions whose marginals are uniform on the interval $(0, 1)$. In this paper, we restrict our attention to bivariate copulas. Fisher (1997) [13] gave two major reasons as to why copulas are of interest to statisticians: firstly, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions." Specifically, copulas are an important part of the study of dependence between two variables since they allow us to separate the effect of dependence from the effects of the marginal distributions. This feature is analogous to the bivariate normal distribution where the mean vectors are unlinked to the covariance matrix and jointly determine the distribution. Many authors have studied constructions of bivariate distributions with given marginals: This may be viewed as constructing a copula.

Nonparametric estimators of copula densities have been suggested by Gijbels and Mielnicsuk [14] and Fermanian and Scaillet [10], who used kernel methods, Sancetta [24] and Sancetta and Satchell [25], who used techniques based on Bernstein polynomials. Biau and Wegkamp[1] proposed estimating the copula density through a minimum distance criterion. Faugeras [7] in his thesis studied the quantile copula approach to conditional density estimation.

The aim of this paper is devoted to the estimation of a regression model via a copulae function, the rest of the paper is organized as follows; at first in section 2 we state Sklar's theorem which elucidates the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals and at the end of the section we introduce our model, then in section 3 we make some regularity assumptions on the kernels and the densities which, although far from being minimal, are somehow customary in kernel density estimation, the main result and its proof is given in the fourth part of this paper. Then we finish this work by a small conclusion.

2.2 The model

Let $(X_i; Y_i); i = 1, 2, \dots, n$ be an independent identically distributed sample from real-valued random variables (X, Y) sitting on a given probability space. For predicting the response Y of the input variable X at a given location x, it is of great interest to estimate not only the conditional mean or regression function $\mathbb{E}(Y/X = x)$, but the full conditional density $f(y/x)$. Indeed, estimating the conditional density is much more informative, since it allows not only to recalculate from the density the conditional expected value $\mathbb{E}(Y/X)$, but also many other characteristics of the distribution such as the conditional variance. In particular, having knowledge of the general shape of the conditional density, is especially important for multi-modal or skewed densities, which often arise from nonlinear or non-Gaussian phenomena, where the expected value might be nowhere near a mode, i.e. the most likely value to appear.

A natural approach to estimate the conditional density $f(y/x)$ of Y given $X = x$ would be to exploit the identity

$$f(y/x) = \frac{f_{XY}(x,y)}{f_X(x)}, \quad f_X(x) \neq 0, \tag{2.1}$$

where f_{XY} and f_X denote the joint density of (X, Y) and X, respectively.

By introducing Parzen-Rosenblatt [21, 22] kernel estimators of these densities, namely,

$$\widehat{f}_{n,XY}(x,y) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x)K_h(Y_i - y),$$

$$\widehat{f}_{n,X}(x) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x),$$

where $K_h(.) = \frac{1}{h}K(./h)$ and $K'_{h'}(.) = \frac{1}{h'}K'(./h')$ are (rescaled) kernels with their associated sequence of bandwidth $h = h_n$ and $h' = h'_n$ going to zero as $n \to 1$, one can construct the quotient

$$\widehat{f}_n(y/x) = \frac{\widehat{f}_{n,XY}(x,y)}{\widehat{f}_{n,X}(x)},$$

and obtain an estimator of the conditional density.

Formally, Sklar's theorem below elucidates the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals see Sklar[28].

Theorem 2.2.1. ***(Sklar 1959)*** *For any bivariate cumulative distribution function* $F_{X,Y}$ *on* $\mathbb{R}^2$*, with marginal cumulative distribution functions* F *of* X *and* G *of* Y*, there exists some function* $C : [0,1]^2 \to [0,1]$*, called the dependence or copula function, such as*

$$F_{X,Y}(x,y) = C(F(x), G(y)), \quad -\infty \leq x, y \leq +\infty. \tag{2.2}$$

If F *and* G *are continuous, this representation is unique with respect to* (F,G)*. The copula function* C *is itself a cumulative distribution function on* $[0,1]^2$ *with uniform marginals.*

This theorem gives a representation of the bivariate c.d.f. as a function of each univariate c.d.f. In other words, the copula function captures the dependence structure among the components X and Y of the vector (X,Y), irrespectively of the marginal distribution F and G. Simply put, it allows to deal with the randomness of the dependence structure and the randomness of the marginals separately.

Copulas appear to be naturally linked with the quantile transform: in the case F and G are continuous, formula (3.2) is simply obtained by defining the copula function as $C(u,v) = F_{X,Y}(F^{-1}(u), G^{-1}(v)), \;\; 0 \leq u \leq 1, \;\; 0 \leq v \leq 1$. For more details regarding copulas and their properties, one can consult for example the book of Joe [17]. Copulas have witnessed a renewed interest in statistics, especially in finance, since the pioneering work of Räuschendorf [23] and Deheuvels [4], who introduced the empirical copula process. Weak convergence of the empirical copula process was investigated by Deheuvels [5], Van der Vaart and Wellner [29], Fermanian, Radulovic and Wegkamp [9]. For the estimation of the copula density, refer to Gijbels and Mielniczuk [14], Fermanian [8] and Fermanian and Scaillet [11].

From now on, we assume that the copula function $C(u,v)$ has a density $c(u,v)$ with respect to the Lebesgue measure on $[0,1]^2$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u,v)$ and $c(u,v)$ are then the cumulative distribution function (c.d.f.) and density respectively of the transformed variables $(U,V) = (F(x), G(y))$. By differentiating formula (3.2), we get for the joint density,

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = f(x)g(y)c(F(x), G(y)),$$

where $c(u,v) := \frac{\partial^2 C(u,v)}{\partial u \partial v}$ is the above mentioned copula density. Eventually, we can obtain the following explicit formula of the conditional density

$$f(y/x) = \frac{f_{XY}(x,y)}{f(x)} = g(y)c(F(x), G(y)), \quad f(x) \neq 0. \tag{2.3}$$

So, let

$$f_n(y/x) = \widehat{g}_n(y)\widehat{c}_n(F_n(x), G_n(y)),$$

be an estimator which builds on the idea of using synthetic data. where $\widehat{g}_n(y)$, $\widehat{c}_n$, $F_n(x)$, $G_n(y)$ are estimators of the density g of Y, the copula density c, the c.d.f. F of X and G of Y respectively. Its study then reveals to be particularly simple: it reduces to the ones already done on nonparametric density estimation.

From now on, we assume that the copula function $C(u,v)$ has a density $c(u,v)$ with respect to the Lebesgue measure on $[0,1]^2$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u,v)$ and $c(u,v)$ are then the cumulative distribution function (c.d.f.) and density respectively of the transformed variables $(U,V) = (F(X), G(Y))$.

Now, To build an estimator of the conditional density we have to use a Parzen-Rosenblatt kernel type non parametric estimator of the marginal density g of Y.

$$\widehat{g}_n(y) := \frac{1}{nh_n} \sum_{i=1}^{n} K_0\left(\frac{y - Y_i}{h_n}\right),$$

the empirical distribution functions $F_n(x)$ and $G_n(y)$ for $F(x)$ and $G(y)$ respectively,

$$F_n(x) = \sum_{j=1}^{n} 1_{X_j \leq x} \text{ and } G_n(y) = \sum_{j=1}^{n} 1_{Y_j \leq y}.$$

Concerning the copula density $c(u,v)$, we noted that $c(u,v)$ is the joint density of the transformed variables $(U,V) := (F(x), G(y))$. Therefore, $c(u,v)$ can be estimated by the bivariate Parzen-Rosenblatt kernel type non parametric density (pseudo) estimator,

$$c_n(u,v) := \frac{1}{nh_nb_n} \sum_{i=1}^{n} K\left(\frac{u - U_i}{h_n}, \frac{v - V_i}{b_n}\right), \tag{2.4}$$

where K is a bivariate kernel and h_n, b_n its associated bandwidth. For simplicity, we restrict ourselves to product kernels, i.e. $K(u,v) = K_1(u)K_2(v)$ with the same bandwidths $h_n = b_n$.

Nonetheless, since F and G are unknown, the random variables (U_i, V_i) are not observable, i.e. c_n is not a true statistic. Therefore, we approximate the pseudo-sample (U_i, V_i), $i = 1, 2, \ldots, n$ by its empirical counterpart $(F_n(X_i), G_n(Y_i))$, $i = 1, 2, \ldots, n$. We therefore obtain a genuine estimator of $c(u,v)$.

$$\widehat{c}_n(u,v) := \frac{1}{nh_n^2} \sum_{i=1}^{n} K_1\left(\frac{u - F_n(X_i)}{h_n}\right) K_2\left(\frac{v - G_n(Y_i)}{h_n}\right). \tag{2.5}$$

Now, let us present Our estimated model, the regression function $r(x)$, is given as follows:

$$r(x) = Yc_n(F(x), G(y)), \;\; |Y| \leq M, \;\; Y, m \in \mathbb{R}.$$

This regression function $r(x)$ is estimated by a function $\widehat{r} = Y\widehat{c}_n(F(x), G(y))$.

To state our main result, we will have to make some regularity assumptions on the kernels and the densities which, although far from being minimal, are somehow customary in kernel density estimation.

2.3 Notations and Assumptions

Set x and y two fixed points in the interior of $supp(f)$ and $supp(g)$ respectively.

The support of the densities function f and g are noted by

$$supp(f) = \overline{\{x \in \mathbb{R}; f(x) > 0\}} \quad \text{and} \quad supp(g) = \overline{\{y \in \mathbb{R}; g(y) > 0\}},$$

where A stands for the closure of a set A.

N.B. $o_P(.)$ and $O_p(.)$ (respectively $o_{a.s}(.)$ and $O_{a.s}(.)$)will stands for convergence and boundedness in probability (respectively almost surely).

Assumptions

- (i) the c.d.f F of X and G of Y are strictly increasing and differentiable.

- (ii) the densities g and c are twice continuously differentiable with bounded second derivatives on their support.

- (iii) the densities g and c are uniformly continuous and non-vanishing almost everywhere on a compact set $J := [a, b]$ and $D \subset (0, 1) \times (0, 1)$ included in the interior of $supp(g)$ and $supp(c)$, respectively.

- (iv) K and K_0 are of bounded support and of bounded variation.

- (v) $0 \leq K \leq C$ and $0 \leq K_0 \leq C$ for some constant C.

- (vi) K and K_0 are second order kernels.

- (vii) K it is twice differentiable with bounded second partial derivatives.

Recall that $c_n(u, v)$ is the kernel copula (pseudo) density estimator from the unobservable, but fixed with respect to n, pseudo data $(F(X_i), G(Y_i))$, and that $\widehat{c}_n(u, v)$ is its analogue made from the approximate data $(F_n(X_i), G_n(Y_i)$. The heuristic of the reason why our estimator works is that the $n^{-1/2}$ in probability rate of convergence in uniform norm of F_n and G_n to F and G is faster than the $1/\sqrt{nh_n^2}$ rate of the non parametric kernel estimator c_n of the copula density c. Therefore, the approximation step of the unknown transformations F and G by their empirical counterparts F_n and G_n does not have any impact asymptotically on the estimation step of c by c_n. Put in another way, one can approximate $\widehat{c}_n(F_n(x), G_n(y))$ by $c_n(F(x), G(y))$ at a faster rate than the convergence rate of $c_n(F(x), G(y))$ to $c(F(x), G(y))$.

2.4 Main Result

This part of the paper is devoted to the asymptotic study the convergence in probability and almost surely (with rate) of our estimators introduced above.

Theorem 2.4.1. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied, if h_n tends to zero as $n \to \infty$ in such a way that*

$$nh_n^4 \to \infty, \frac{\sqrt{\ln \ln n}}{nh_n^3} \to 0,$$

then,

$$\widehat{r}_n(x) = r(x) + O_P\left(h_n^2 + \frac{1}{\sqrt{nh_n^2}} + \frac{1}{nh_n^4} + \frac{\sqrt{\ln \ln n}}{nh_n^3}\right)$$

Proof theorem 3.4.1: Let $\widehat{r}(x) = Y c_n(F(x), G(x))$, to demonstrate that $\widehat{r}(x)$ converge to $r(x)$ it is sufficient to prove that $\widehat{c}_n(U, V) \to c_n(U, V)$, with $U = F(x)$, $V = G(x)$.

For $(X_i, i = 1, 2, \ldots, n)$ an i.i.d. sample of a real random variable X with common c.d.f. F, the Kolmogorov-Smirnov statistic is defined as $D_n := \|F_n - F\|$. Glivenko-Cantelli, Kolmogorov and Smirnov, Chung, Donsker among others have studied its convergence properties in increasing generality (See e.g. [27] and [28] for recent accounts). For our purpose, we only need to formulate these results in the following rough form:

Lemma 2.4.1. *For an i.i.d. sample from a continuous c.d.f. F,*

$$\|F_n - F\|_\infty = O_P\left(\frac{1}{\sqrt{n}}\right), \quad i = 1, 2, \ldots, n, \tag{2.6}$$

$$\|F_n - F\|_\infty = O_{a.s}\left(\frac{\ln \ln n}{n}\right) \quad i = 1, 2, \ldots, n. \tag{2.7}$$

Since F is unknown, the random variables $U_i = F(X_i)$ are not observed. As a consequence of the preceding lemma, one can naturally approximate these variables by the statistics $F_n(X_i)$. Indeed,

$$\|F(X_i) - F_n(X_i)\| \leq \sup_{x \in \mathbb{R}} \|F(x) - F_n(x)\| = \|F_n - F\|_\infty a.s.$$

Let

$$c_n(U, V) = \frac{1}{nh_n^2} \sum_{i=1}^{n} K_1\left(\frac{U - F_n(X_i)}{h_n}\right) K_2\left(\frac{V - G_n(Y_i)}{h_n}\right),$$

$$\widehat{c}_n(U,V) = \frac{1}{nh_n^2}\sum_{i=1}^{n} K_1\left(\frac{U-F_n(X_i)}{h_n}\right) K_2\left(\frac{V-G_n(Y_i)}{h_n}\right).$$

So, we must show that $F_n(X_i)$ converge to $F(X_i)$ and $G_n(Y_i)$ converge to $G(Y_i)$.

$$\widehat{c}_n(U,V) - c_n(U,V) = \frac{1}{nh_n^2}\left(\sum_{i=1}^{n} K_1\left(\frac{U-F_n(X_i)}{h_n}\right) K_2\left(\frac{V-G_n(Y_i)}{h_n}\right)\right.$$

$$\left. - \sum_{i=1}^{n} K_1\left(\frac{U-F(X_i)}{h_n}\right) K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right),$$

with

$$\Pi_{i,n} = K_1\left(\frac{U-F_n(X_i)}{h_n}\right) K_2\left(\frac{V-G_n(Y_i)}{h_n}\right) - K_1\left(\frac{U-F(X_i)}{h_n}\right) K_2\left(\frac{V-G(Y_i)}{h_n}\right).$$

Let

$$Z_{i,n} = \begin{pmatrix} F_n(X_i) - F(X_i) \\ G_n(Y_i) - G(Y_i) \end{pmatrix}$$

$\|F_n(X_i) - F(X_i)\| \leq \|F_n - F\|_\infty$ and $\|G_n(Y_i) - G(Y_i)\| \leq \|Gn - G\|$ a.s. for every i = 1,2,..., n. Preceding Lemma thus entails that the norm of $Z_{i,n}$ is independent of i and such that

$$\|Z_{i,n}\| = O_P\left(\frac{1}{\sqrt{n}}\right), \quad i = 1, 2, \ldots, n, \tag{2.8}$$

$$\|Z_{i,n}\| = O_{a.s}\left(\frac{\ln\ln n}{n}\right) \quad i = 1, 2, \ldots, n. \tag{2.9}$$

Now, for every fixed $(u, v) \in [0, 1]^2$, since the kernel K is twice differentiable, there exists, by Taylor expansion, random variables $\tilde{U}_{i,n}$ and $\tilde{V}_{i,n}$ such that, almost surely,

$$\Pi = \frac{1}{nh_n^3}\sum_{i=1}^{n} Z_{i,n}^T \nabla\left(K_1\left(\frac{U-F_n(X_i)}{h_n}\right) K_2\left(\frac{V-G_n(Y_i)}{h_n}\right)\right)$$

$$+\frac{1}{2nh_n^4}\sum_{i=1}^{n} Z_{i,n}^T \nabla^2\left(K_1\left(\frac{U-\tilde{U}_{i,n}}{h_n}\right) K_2\left(\frac{V-\tilde{V}_{i,n}}{h_n}\right)\right) Z_{i,n} = \Pi_1 + \Pi_2,$$

where $Z_{i,n}^T$ denotes the transpose of the vector $Z_{i,n}$ and ∇K and $\nabla^2 K$ the gradient and the Hessian respectively of the multivariate kernel function K.

By centering at expectations, decompose further the first term Π_1 as,

$$\Pi_1 = \frac{1}{nh_n^3}\sum_{i=1}^n Z_{i,n}\nabla\left(K_1\left(\frac{U-F(X_i)}{h_n}\right)K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right)$$

$$-\mathbb{E}\nabla\left(K_1\left(\frac{U-F(X_i)}{h_n}\right)K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right)$$

$$+\frac{1}{nh_n^3}\sum_{i=1}^n Z_{i,n}^T\mathbb{E}\nabla\left(K_1\left(\frac{U-F(X_i)}{h_n}\right)K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right) = \Pi_{11}+\Pi_{12}$$

We again decompose one step further Π_{11}, Set

$$A_i = \nabla\left(K_1\left(\frac{U-F(X_i)}{h_n}\right)K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right) - \mathbb{E}\nabla\left(K_1\left(\frac{U-F(X_i)}{h_n}\right)K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right).$$

Then

$$|\Pi_{11}| \le \frac{\|Z_{i,n}\|}{nh_n^3}\sum_{i=1}^n (\|A_i\| - \mathbb{E}\|A_i\|) + \frac{\|Z_{i,n}\|}{nh_n^3}\sum_{i=1}^n \mathbb{E}\|A_i\| = \Pi_{111}+\Pi_{112}.$$

We now proceed to the study of the order of each terms in the previous decompositions.

• **Negligibility of** Π_2.

By the boundedness assumption on the second-order derivatives of the kernel, and equations (3.7) and (3.8),

$$\Pi_2 = O_P\left(\frac{1}{nh_n^4}\right), \text{ and } \Pi_2 = O_{a.s}\left(\frac{\ln\ln n}{nh_n^4}\right)$$

• **Negligibility of** Π_{12}.

Bias results on the bivariate gradient kernel estimator (See Scott [26] chapter 6) entail that

$$\mathbb{E}\nabla\left(K_1\left(\frac{U-F(X_i)}{h_n}\right)K_2\left(\frac{V-G(Y_i)}{h_n}\right)\right) = h_n^3\nabla c(u,v) + O(h_n^5)$$

Cauchy-Schwarz inequality yields that

$$|\Pi_{12}| \leq \frac{\|nZ_{i,n}\|}{nh_n^3} \left\| \mathbb{E}\nabla \left(K_1 \left(\frac{U - F(X_i)}{h_n} \right) K_2 \left(\frac{V - G(Y_i)}{h_n} \right) \right) \right\|$$

In turn, with equations (3.7) and (3.8),

$$\Pi_{12} = O_P \left(\frac{1}{\sqrt{n}} \right), \text{ and } \Pi_{12} = O_{a.s} \left(\frac{\ln \ln n}{n} \right)$$

• **Negligibility of Π_{11}**

• **Negligibility of Π_{111}.**

Boundedness assumption on the derivative of the kernel imply that $\|A_i\| \leq 2C$ a.s. We apply Hoeffding inequality for independent, centered, bounded by M, but non identically distributed random variables (η_j) (e.g. see [2]),

$$\mathbb{P}\left(\sum_{j=1}^{n} \eta_j > t \right) \leq exp \left(\frac{-t^2}{2nM^2} \right)$$

Here, for every $\epsilon > 0$, with $M = 2C$, $\eta_j = \|A_i\| - \mathbb{E}\|A_i\|$, $t = \epsilon\sqrt{\frac{1}{n} \ln \ln n}$, therefore,

$$\sum_{i=1}^{n} (\|A_i\| - \mathbb{E}\|A_i\|) = O_p(\sqrt{n \ln \ln n})$$

which is the definition of almost complete convergence (a.co.), see e.g. [12] definition A.3. p. 230. In turn, it means that

$$\sum_{i=1}^{n} (\|A_i\| - \mathbb{E}\|A_i\|) = O_{a.co} \left(\sqrt{n \ln n} \right)$$

and by the Borell-Cantelli lemma,

$$\sum_{i=1}^{n} (\|A_i\| - \mathbb{E}\|A_i\|) = O_{a.s}(\sqrt{n \ln n})$$

Therefore, using equations equations (3.7) and (3.8), we have that

$$\Pi_{111} = O_P\left(\frac{\sqrt{\ln\ln n}}{nh_n^3}\right) = O_{a.s}\left(\sqrt{\ln n}\frac{\sqrt{\ln\ln n}}{nh_n^3}\right)$$

- **Negligibility of** Π_{112}

The r.h.s. of the previous inequality is, after an integration by parts, of order a3 n by the results on the kernel estimator of the gradient of the density (See Scott [26] chapter 6). Therefore,

$$\sum_{i=1}^{n} \mathbb{E}\|A_i\| = O(nh_n^2)$$

$$\Pi_{112} = \frac{\|nZ_{i,n}\|}{nh_n^3}\sum_{i=1}^{n} \mathbb{E}\|A_i\| = O_P\left(\frac{1}{\sqrt{n}}\right) = O_{a.s}\left(\frac{\sqrt{\ln\ln n}}{n}\right)$$

by equations (3.7) and (3.8).

Recollecting all elements, we eventually obtain that

$$\Pi = \Pi_{111} + \Pi_{112} + \Pi_{12} + \Pi_2 = O_P\left(\frac{1}{\sqrt{n}}\right) + O_P\left(\frac{\ln\ln n}{nh_n^3}\right) + O_P\left(\frac{1}{nh_n^4}\right)$$

$$= O_{a.s}\left(\sqrt{\frac{\ln\ln n}{n}}\right) + O_{a.s}\left(\frac{\sqrt{\ln n}\sqrt{\ln\ln n}}{nh_n^3}\right) + O_{a.s}\left(\frac{\ln\ln n}{nh_n^4}\right).$$

By this last step we conclude the proof of our theorem.

After giving the proof of the convergence in probability, let us present the rate of convergence in the following corollary.

Corollary 2.4.1. *We get the rate of convergence, by choosing the bandwidth which balance the bias and variance trade-off: for an optimal choice of* $h_n \simeq n^{-1/6}$, *we get*

$$\widehat{r}_n(x) = r(x) + O_P(n^{-1/3})$$

Therefore, our estimator is rate optimal in the sense that it reaches the minimax rate $n^{-1/3}$ of convergence.

Now, Almost sure results can be proved in the same way: we have the following strong consistency result,

Theorem 2.4.2. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied.If the bandwidth h_n tends to zero as $n \to \infty$ in such a way that*

$$\frac{\sqrt{\ln n \ln \ln n}}{nh_n^3} \to 0, \frac{\ln \ln n}{nh_n^4} \to 0,$$

then,

$$\widehat{r}_n(x) = r(x) + O_{a.s}\left(h_n^2 + \sqrt{\frac{\ln \ln n}{nh_n^2}} + \frac{\ln \ln n}{nh_n^4} + \frac{\sqrt{\ln n \ln \ln n}}{nh_n^3}\right)$$

For the proof of this theorem, It is sufficient to follow the same lines as the preceding theorem , but uses the a.s. results of the consistency of the kernel density estimators of lemmas 3.13 and 3.15 and of the approximation propositions 3.16 and 3.17. It is therefore similar and omitted [7].

Corollary 2.4.2. *For $h_n \simeq (\ln \ln n/n)^{1/6}$ which is the optimal trade-off between the bias and the stochastic term, one gets the optimal rate*

$$\widehat{r}_n(x) = r(x) + O_{a.s}\left(\frac{\ln \ln n}{n}\right)^{1/3}.$$

For the he proof, we follow the same way given in [7]

2.5 Conclusion

In this paper we established the convergence Almost surely and in Probability (with rate) of regression model via copula function approach, it will be interesting in further work to study the asymptotic normality of such a model.

Bibliography

[1] Biau,G., AND Wegkamp,M.H.," A note on minimum distance estimation of copulas densities"., *Statist. Probab. Lett.*, No.73, (2006), pp. 105-114.

[2] Bosq, D., *Nonparametric statistics for stochastic processes*, second ed., Vol. 110 of Lecture Notes in Statistics. Springer-Verlag, New York, (1998). Estimation and prediction.

[3] Cherubini, U., Luciano, E., AND Vecchiato, W., *Copula Methods in Finance*. John Wiley and Sons, Chichester (2004).

[4] Deheuvels, P., "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance". *Acad. Roy. Belg. Bull. Cl. Sci.*, Vol.65,No.5 And 6, (1979),pp. 274-292.

[5] Deheuvels, P., "A Kolmogorov-Smirnov type test for independence and multivariate samples", *Rev. Roumaine Math. Pures Appl.*Vol 26,No.2, (1981), pp. 213-226.

[6] Embrechts, P., Lindskog, F., McNeil, A., *Modelling dependence with copulas and applications to risk management.*, Handbook of Heavey Tailed Distributions in Finance, S.T. Rachev (ed.). Elsevier, Amsterdam (2003).

[7] Faugeras, Olivier Paul., *Contributions à la prévision statistique*, Thèses de doctorat de l'université Pierre er Marie Curie, 28 Novembre 2008.

[8] Fermanian, J.D., "Goodness-of-fit tests for copulas.", *J. Multivariate Anal.* Vol.95, No.1, (2005), pp. 119-152.

[9] Fermanian, J.D., Radulovïc, D., AND Wegkamp, M., "Weak convergence of empirical copula processes". Bernoulli Vol.10,No.5, (2004), pp. 847-860.

[10] Fermanian, J.D., AND Scaillet, O., *Some statistical pitfalls in copula modelling for financical applications*, E. Klein (Ed.), Capital formation, Gouvernance and Banking. Nova Science Publishing, New York, (2005).

[11] Fermanian, J.D.,AND Scaillet, O., "Nonparametric estimation of copulas for time series", *Journal of Risk*, Vol.5,No. 4, (2003), pp. 25-54.

[12] Ferraty, F., AND Vieu, P., *Nonparametric functional data analysis.* Springer Series in Statistics. Springer, New York, (2006). Theory and practice.

[13] Fisher, N.I., *Copulas*, Encyclopedia of Statistical Sciences, Updated Volume 1, S. Kotz, C. B. Read, D. L. Banks (eds.), pp. 159-164. John Wiley and Sons, New York (1997).

[14] Gijbels, I., AND Mielniczuk, J., "Estimating the density of a copula function", *Comm. Statist. Theory Methods* ,Vol.19,No.2 (1990), pp.445-464.

[15] Hutchinson, T.P., Lai, C.D., *Continuous Bivariate Distributions: Emphasising Applications*,Rumsby Scientific Publishing, Adelaide (1990).

[16] Joe, H., *Multivariate Models and Dependence Concepts*, Chapman and Hall, London (1997).

[17] Joe, H., *Multivariate models and dependence concepts*, vol. 73 of Monographs on Statistics and Applied Probability. Chapman & Hall, London, (1997).

[18] Mikosch, T., "Copulas: Taies and facts", *Extrêmes*,Vol. 9,No.18,(2006), pp.3-22 .

[19] Nelsen, R.B., *An Introduction to Copulas*, 2nd edition. Springer-Verlag, New York (2006).

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

[21] Parzen, E., "On estimation of a probability density function and mode", *Ann. Math. Statist.33*, (1962) , pp.1065-1076.

[22] Pickands, III, J., "Statistical inference using extreme order statistics", *Ann. Statist.3*,(1975), pp.119-131.

[23] Rüschendorf, L.," Asymptotic distributions of multivariate rank order statistics", *Ann. Statist. 4*, 5 (1976), pp.912-923.

[24] Sancetta,A., "Nonparametric estimation of multivariate distributions with given marginals: L_2 theory",*Cambridge Working papers in Economics* No.0320, (2003).

[25] Sancetta, A., Satchell,S., "The Barnestien coplua and its application to modelling and approximation of multivariate distributions", *Econometric theory*, 20, (2004), pp.535-562.

[26] Scott, D. W., *Multivariate density estimation. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley & Sons Inc., New York, (1992). Theory, practice, and visualization, A Wiley-Interscience Publication.

[27] Shorack, G. R., AND Wellner, J. A., *Empirical processes with applications tostatistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* John Wiley & Sons Inc., New York, (1986).

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

[29] Van Der Vaart, A. W., AND Wellner, J. A., *Weak convergence and empirical processes*, Springer Series in Statistics. Springer-Verlag, New York, (1996). With applications to statistics.

Chapter 3

Asymptotic normality

This chapter is the subject of a paper submitted to Forum Proba Stat.

A note on Asymptotic normality of a copula function in regression model

Djamila BENNAFLA, Abbes RABHI, Amina Angelika BOUCHENTOUF

Stochastic Models, Statistics and Applications Laboratory, Moulay Tahar University, Saida
E-mail: dbennafla@yahoo.fr
Mathematics Laboratory, Djillali Liabes University, Sidi Bel Abbes
E-mail: rabhi_abbes@yahoo.fr
Mathematics Laboratory, Djillali Liabes University, Sidi Bel Abbes
E-mail: bouchentouf_amina@yahoo.fr

Abstract

Over the last decade, there has been significant and rapid development of the theory of copulas. Much of the work has been motivated by their applications to stochastic processes, economics, risk management, finance, insurance, the environment (hydrology, climate, etc.), survival analysis, and medical sciences. In many statistical models. The copula approach is a way to solve the difficult problem of finding the whole bivariate or multivariate distribution. The goal of this note is give the asymptotic normality of the copulae function in a regression model.

Keywords: *62N02, 62H12, 62J05.*
2000 MSC No: Copulas, nonparametric estimation, regression model.

3.1 Introduction

Copula models are becoming an increasingly tool for modelling the dependencies between random variables, especially in such fields as biostatistics, actuarial science, and finance. The construction and properties of copulas have been studied rather extensively during the last 15 years. Hutchinson and Lai (1990) [16] were among the early authors who popularized the study of copulas. Nelsen (1999) [21] presented a comprehensive treatment of bivariate copulas, while Joe (1997) [17] devoted a chapter of his book to multivariate copulas. Further authoritative updates on copulas are given in Nelsen (2006) [20]. Copula methods have many important applications in insurance and finance [Cherubini and al. (2004) [4] and Embrechts and al. (2003) [7]].

Briefly speaking, copulas are functions that join or "couple" multivariate distributions to their one-dimensional marginal distribution functions. Equivalently, copulas are multivariate distributions whose marginals are uniform on the interval $(0, 1)$. In this chapter, we restrict our attention to bivariate copulas. Fisher (1997) [14] gave two major reasons as to why copulas are of interest to statisticians: "Firstly, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions." Specifically, copulas are an important part of the study of dependence between two variables since they allow us to separate the effect of dependence from the effects of the marginal distributions. This feature is analogous to the bivariate normal distribution where the mean vectors are unlinked to the covariance matrix and jointly determine the distribution. Many authors have studied constructions of bivariate distributions with given marginals: This may be viewed as constructing a copula. There is a fast-growing industry for copulas. They have useful applications in econometrics, risk management, finance, insurance, etc. The commercial statistics software SPLUS provides a module in FinMetrics that include copula fitting written by Carmona (2004) [3]. One can also get copula modules in other major software packages such as R, Mathematica, Matlab, etc. The International Actuarial Association (2004) [30] in a paper on Solvency II,1 recommends using copulas for modeling dependence in insurance portfolios. Moodyïs uses a Gaussian copula for modeling credit risk and provides software for it that is used by many financial institutions. Basle II2 copulas are now standard tools in credit risk management. There are many other applications of copulas, especially the Gaussian copula, the extreme-value copulas, and the Archimedean copula. We now classify these applications into several categories Nonparametric estimators of copula densities have been suggested by Gijbels and Mielnicsuk [15] and Fermanian and Scaillet [11], who used kernel methods, Sancetta [25] and Sancetta and Satchell [26], who used techniques based on Bernstein polynomials. Biau and Wegkamp[1] proposed estimating the copula density through a minimum distance criterion. Faugeras [8] in his thesis studied the quantile copula approach to conditional density estimation.
the main goal of this note is devoted to the asymptotic normality of a regression model via a copulae function, for that; At first we introduce the model, then we make some notations and regularity assumptions for our main result given in the last part of this note.

3.2 The model

Let $((X_i;Y_i); i = 1, \ldots, n)$ be an independent identically distributed sample from real-valued random variables (X, Y) sitting on a given probability space. For predicting the response Y of the input variable X at a given location x, it is of great interest to estimate not only the conditional mean or regression function $E(Y/X = x)$, but the full conditional density $f(y/x)$. Indeed, estimating the conditional density is much more informative, since it allows not only to recalculate from the density the conditional expected value $E(Y/X)$, but also many other characteristics of the distribution such as the conditional variance. In particular, having knowledge of the general shape of the conditional density, is especially important for multi-modal or skewed densities, which often arise from nonlinear or non-Gaussian phenomena, where the expected value might be nowhere near a mode, i.e. the most likely value to appear.

A natural approach to estimate the conditional density $f(y/x)$ of Y given $X = x$ would be to exploit the identity

$$f(y/x) = \frac{f_{XY}(x, y)}{f_X(x)}, \quad f_X(x) \neq 0, \tag{3.1}$$

where f_{XY} and f_X denote the joint density of (X, Y) and X, respectively.

By introducing Parzen-Rosenblatt [22, 23] kernel estimators of these densities, namely,

$$\hat{f}_{n,XY}(x, y) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x)K_h(Y_i - y),$$

$$\hat{f}_{n,X}(x) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x),$$

where $K_h(.) = 1/hK(./h)$ and $K'_{h'}(.) = 1/h'K'(./h')$ are (rescaled) kernels with their associated sequence of bandwidth $h = h_n$ and $h' = h'_n$ going to zero as $n \to 1$, one can construct the quotient

$$\hat{f}_n(y/x) = \frac{\hat{f}_{n,XY}(x, y)}{\hat{f}_{n,X}(x)},$$

and obtain an estimator of the conditional density.

Formally, Sklar's theorem below elucidates the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals see Sklar[29].

Theorem 3.2.1 (Sklar 1959)**.** *For any bivariate cumulative distribution function $F_{X,Y}$ on $\mathbb{R}^2$, with marginal cumulative distribution functions F of X and G of Y, there exists some function $C : [0,1]^2 \to [0,1]$, called the dependence or copula function, such as*

$$F_{X,Y}(x, y) = C(F(x), G(y)), \;\; -\infty \leq x, y \leq +\infty. \tag{3.2}$$

If F and G are continuous, this representation is unique with respect to (F, G). The copula function C is itself a cumulative distribution function on $[0,1]^2$ with uniform marginals.

This theorem gives a representation of the bivariate c.d.f. as a function of each univariate c.d.f. In other words, the copula function captures the dependence structure among the components X and Y of the vector (X, Y), irrespectively of the marginal distribution F and G. Simply put, it allows to deal with the randomness of the dependence structure and the randomness of the marginals separately.

Copulas appear to be naturally linked with the quantile transform: in the case F and G are continuous, formula (3.2) is simply obtained by defining the copula function as $C(u, v) = F_{X,Y}(F^{-1}(u), G^{-1}(v)), \;\; 0 \leq u \leq 1, \;\; 0 \leq v \leq 1$. For more details regarding copulas and their properties, one can consult for example the book of Joe [18]. Copulas have witnessed a renewed interest in statistics, especially in finance, since the pioneering work of Räuschendorf [24] and Deheuvels [5], who introduced the empirical copula process. Weak convergence of the empirical copula process was investigated by Deheuvels [6], Van der Vaart and Wellner [31], Fermanian, Radulovic and Wegkamp [10]. For the estimation of the copula density, refer to Gijbels and Mielniczuk [15], Fermanian [9] and Fermanian and Scaillet [12].

From now on, we assume that the copula function $C(u, v)$ has a density $c(u, v)$ with respect to the Lebesgue measure on $[0,1]^2$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u, v)$ and $c(u, v)$ are then the cumulative distribution function (c.d.f.) and density respectively of the transformed variables $(U, V) = (F(x), G(y))$. By differentiating formula (3.2), we get for the joint density,

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} = \frac{\partial^2 C(F(x), G(y))}{\partial F(x) \partial G(y)} \frac{\partial F(x)}{\partial x} \frac{\partial G(y)}{\partial y} = f(x) g(y) c(F(x), G(y)),$$

where $c(u,v) := \frac{\partial^2 C(u,v)}{\partial u \partial v}$ is the above mentioned copula density. Eventually, we can obtain the following explicit formula of the conditional density

$$f(y/x) = \frac{f_{XY}(x,y)}{f(x)} = g(y)c(F(x), G(y)), \quad f(x) \neq 0. \tag{3.3}$$

Concerning the copula density $c(u,v)$, we noted that $c(u,v)$ is the joint density of the transformed variables $(U,V) := (F(x), G(y))$. Therefore, $c(u,v)$ can be estimated by the bivariate Parzen-Rosenblatt kernel type non parametric density (pseudo) estimator,

$$c_n(u,v) := \frac{1}{nh_nb_n} \sum_{i=1}^{n} K\left(\frac{u-U_i}{h_n}, \frac{v-V_i}{b_n}\right), \tag{3.4}$$

where K is a bivariate kernel and h_n, b_n its associated bandwidth. For simplicity, we restrict ourselves to product kernels, i.e. $K(u,v) = K_1(u)K_2(v)$ with the same bandwidths $h_n = b_n$.

Nonetheless, since F and G are unknown, the random variables (U_i, V_i) are not observable, i.e. c_n is not a true statistic. Therefore, we approximate the pseudo-sample (U_i, V_i), $i = 1, ..., n$ by its empirical counterpart $(F_n(X_i), G_n(Y_i))$, $i = 1, ..., n$. We therefore obtain a genuine estimator of $c(u,v)$.

$$\hat{c}_n(u,v) := \frac{1}{nh_n^2} \sum_{i=1}^{n} K_1\left(\frac{u - F_n(X_i)}{h_n}\right) K_2\left(\frac{v - G_n(Y_i)}{h_n}\right). \tag{3.5}$$

the empirical distribution functions $F_n(x)$ and $G_n(y)$ for $F(x)$ and $G(y)$ respectively,

$$F_n(x) = \sum_{j=1}^{n} 1_{X_j \le x} \text{ and } G_n(y) = \sum_{j=1}^{n} 1_{Y_j \le y}.$$

Our estimated model is given as follows: the regression function $r(x)$ is estimated by a function $\widehat{r}_n(x)$

$$r(x) = E(Y/X = x) = \int yf(y/x)dy = \int yg(y)c(F(x), G(y))dy = E(Yc(F(x), G(y))). \tag{3.6}$$

This regression function $r(x)$ is estimated by a function $\hat{r}_n(x) = \int y\hat{f}_n(y/x)dy$, thus, we obtain

$$\hat{r}_n(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i\hat{c}_n(F_n(x), G_n(y)) = Y\hat{c}_n(F_n(x), G_n(y))$$

For more detail see [8]

To state our result, we will have to make some regularity assumptions on the kernels and the densities which, although far from being minimal, are somehow customary in kernel density estimation.

3.3 Notations and Assumptions

We note the ith moment of a generic kernel (possibly multivariate) K as

$$m_i(K) := \int u^i K(u)du$$

and the L_p norm of a function h by $\|s\|_p := \int s^p$. We use the sign $\simeq$ to denote the order of the bandwidths. Set (u, v) fixed point in the interior of $supp(c)$. The support of the densitie function c is noted by $supp(c) = \overline{\{(u, v) \in \mathbb{R}^2; c(u, v) > 0\}}$ where $\bar{A}$ stands for the closure of a set A.Finally, $O_P(.)$ and $O_p(.)$ (respectively $o_{a.s}(.)$ and $O_{a.s}(.)$)will stands for convergence and boundedness in probability (respectively almost surely).

Assumptions

- (i) the c.d.f F of x and G of Y are strictly increasing and differentiable.

- (ii) the densitie c is twice continuously differentiable with bounded second derivatives on its support.

- (iii) the densitie c is uniformly continuous and non-vanishing almost everywhere on a compact set $D \subset (0, 1) \times (0, 1)$ included in the interior of $supp(c)$.

- (iv) K is of bounded support and of bounded variation.

- (v) $0 \leq K \leq \alpha$ and $0 \leq K_0 \leq \alpha$ for some constant α.

- (vi) K is second order kernels. $m_0(K) = 1, m_1(K) = 0$ et $m_2(K) < +\infty$

- (vii) K it is twice differentiable with bounded second partial derivatives.

Recall that $c_n(u, v)$ is the kernel copula (pseudo) density estimator from the unobservable, but fixed with respect to n, pseudo data $(F(X_i), G(Y_i))$, and that $\hat{c}_n(u, v)$ is its analogue made from the approximate data $(F_n(X_i), G_n(Y_i)$. The heuristic of the reason why our estimator works is that the $n^{-1/2}$ in probability rate of convergence in uniform norm of F_n and G_n to F and G is faster than the $1/\sqrt{na_n^2}$ rate of the non parametric kernel estimator c_n of the copula density c. Therefore, the approximation step of the unknown transformations F and G by their empirical counterparts F_n and G_n does not have any impact asymptotically on the estimation step of c by c_n. Put in another way, one can approximate $\hat{c}_n(F_n(x), G_n(y))$ by $c_n(F(x), G(y))$ at a faster rate than the convergence rate of $c_n(F(x), G(y))$ to $c(F(x), G(y))$.

3.4 Main Result

This part of the paper is devoted to the asymptotic study the convergence in probability and almost surely of our estimators introduced above. But at first let us present the rate convergence of the estimator.

Theorem 3.4.1. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied, if h_n tends to zero as $n \to \infty$ in such a way that*

$$nh_n^4 \to \infty, \frac{\sqrt{\ln \ln n}}{nh_n^3} \to 0,$$

then,

$$\hat{r}_n(x) = r(x) + O_P\left(h_n^2 + \frac{1}{\sqrt{nh_n^2}} + \frac{1}{nh_n^4} + \frac{\sqrt{\ln \ln n}}{nh_n^3}\right)$$

The main ingredient of the proof follows from the fact that

$$\hat{r}_n(x) - r(x) = Y(\hat{c}_n(F_n(x), G_n(y)) - c_n(F(x), G(y)))$$

On the one hand, convergence results for the kernel density estimators of what will follow entail that,

$$c_n(F(x), G(y)) - c(F(x), G(y)) = O_P(h_n^2 + 1/\sqrt{nh_n^2}).$$

Thus, by lemmas 3.13 and 3.4.1 respectively [8].

Corollary 3.4.1. *we get the rate of convergence, by choosing the bandwidth which balance the bias and variance trade-off: for an optimal choice of $h_n \simeq n^{-1/6}$, we get*

$$\hat{r}_n(x) = r(x) + O_P(n^{-1/3})$$

Therefore, our estimator is rate optimal in the sense that it reaches the minimax rate $n^{-1/3}$ of convergence.

Almost sure results can be proved in the same way: we have the following strong consistency result,

Theorem 3.4.2. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied.If the bandwidth h_n tends to zero as $n \to \infty$ in such a way that*

$$\frac{\sqrt{\ln n \ln \ln n}}{nh_n^3} \to 0, \frac{\ln \ln n}{nh_n^4} \to 0,$$

then,

$$\hat{r}_n(x) = r(x) + O_{a.s}\left(h_n^2 + \sqrt{\frac{\ln \ln n}{nh_n^2}} + \frac{\ln \ln n}{nh_n^4} + \frac{\sqrt{\ln n \ln \ln n}}{nh_n^3}\right)$$

For the proof of this theorem, It is sufficient to follow the same lines as the preceding theorem , but uses the a.s. results of the consistency of the kernel density estimators of lemmas 3.13 and 3.4.1 and of the approximation propositions 3.4.1 and 3.4.2[8]. It is therefore similar and omitted.

Corollary 3.4.2. *For $h_n \simeq (\ln \ln n/n)^{1/6}$ which is the optimal trade-off between the bias* *and the stochastic term, one gets the optimal rate*

$$\hat{r}_n(x) = r(x) + O_{a.s}\left(\frac{\ln\ln n}{n}\right)^{1/3}.$$

Let $\hat{r}_n(x) = Y\hat{c}_n(F_n(x), G_n(x))$, to demonstrate that $\hat{r}_n(x)$ converge to $r(x)$ it is sufficient to prove that $\hat{c}_n(u,v) \to c_n(u,v)$.

Recall a preliminary result that will be needed for the main result.

For $(X_i, i = 1, ..., n)$ an i.i.d. sample of a real random variable X with common c.d.f. F, the Kolmogorov-Smirnov statistic is defined as $D_n := \|F_n - F\|_\infty$. Glivenko-Cantelli, Kolmogorov and Smirnov, Chung, Donsker among others have studied its convergence properties in increasing generality (See e.g. [?] and [?] for recent accounts). For our purpose, we only need to formulate these results given in the Lemma 3.13 [8]:

Via Lemma 3.13 [8], we get naturally

$$|F(X_i) - F_n(X_i)| \leq \sup_{x\in R} |F(x) - F_n(x)| = \|F_n - F\|_\infty \quad a.s.$$

Apply this result to the estimator c_n. And for this let us introduce the following result Lemma 3.15 page 82 given in [8].

Lemma 3.4.1. *[8] With the previous assumptions, for $(u,v) \in (0,1)^2$, we have,*

- *for a bandwidth chosen such as $h_n \simeq n^{-1/6}$,*

$$|c_n(u,v) - c(u,v)| = O_P(n^{-1/3}),$$

- *for a point (u,v) where $c(u,v) > 0$, and $h_n = o(n^{-1/6})$,*

$$\sqrt{nh_n^2}\left(\frac{c_n(u,v) - c(u,v)}{\sqrt{c_n(u,v)\|K\|_2^2}}\right) \rightsquigarrow N(0,1)$$

- *for a bandwidth chosen of* $h_n \simeq (\ln\ln n/n)^{1/6}$,

$$|c_n(u,v) - c(u,v)| = O_{a.s}\left(\left(\frac{\ln\ln n}{n}\right)^{1/3}\right)$$

Now, we need two proposed approximation $\hat{c}_n$ par le c_n

Proposition 3.4.1. *Let* $(u,v) \in (0,1)^2$*. If the kernel* $K(u,v) = K_1(u)K_2(v)$ *is twice differentiable with bounded second derivative, then*

$$|\hat{c}_n(u,v) - c_n(u,v)| = O_P\left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\ln\ln n}}{nh_n^3} + \frac{1}{nh_n^4}\right)$$

$$|\hat{c}_n(u,v) - c_n(u,v)| = O_{a.s}\left(\sqrt{\frac{\ln\ln n}{n}} + \frac{\sqrt{\ln n}\sqrt{\ln\ln n}}{nh_n^3} + \frac{\ln\ln n}{nh_n^4}\right)$$

Proposition 3.4.2. *with the same assumptions as in the previous proposal was*

- *If* $h_n \to 0, nh_n^3 \to \infty$

$$|\hat{c}_n(F_n(x), G_n(y)) - c_n(F(x), G(y))| = O_P\left(\frac{1}{\sqrt{n}} + \frac{1}{nh_n^4}\right)$$

- *If* $h_n \to 0, nh_n^3/\ln\ln n \to \infty$

$$|\hat{c}_n(F_n(x), G_n(y)) - c_n(F(x), G(y))| = O_{a.s}\left(\sqrt{\frac{\ln\ln n}{n}} + \frac{\ln\ln n}{nh_n^4}\right)$$

Corollary 3.4.3. *[13] Consider the model (3.6). If the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied,then,*

1.

$$\sqrt{nh}\,(\hat{r}(x) - r(x)) \rightsquigarrow \mathcal{N}\left(0, \frac{(\phi(x) - r^2(x))}{f(x)}\|K\|_2^2\right)$$

2.

$$B_o = E(\hat{r}(x)) - r(x) = B(x)h^2 + o(h^2),$$

and

$$V_0 = Var(\hat{r}(x)) = V(x)\frac{1}{nh} + o\left(\frac{1}{nh}\right)$$

with

$$B(x) = \frac{\int t^2 K(t)dt(g^{(2)}(x) - r(x)f^{(2)})}{2f(x)},$$

$$V(x) = \int K^2(t)dt \frac{(\phi(x) - r^2(x))}{f(x)}$$

and $g(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h}\right), f(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right), \phi(x) = E(Y^2/X = x)$

Theorem 3.4.3. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied, if* $h_n \simeq (\ln n/n)^{1/6}$ *then,*

$$\sup_{x\in\mathbb{R}} |\hat{r}_n(x) - r(x)| = O_P\left(\left(\frac{\ln n}{n}\right)^{1/3}\right)$$

and

$$\sup_{x\in\mathbb{R}} |\hat{r}_n(x) - r(x)| = O_{a.s}\left(\left(\frac{\ln n}{n}\right)^{1/3}\right)$$

Proof theorem 3.4.3 The proof is identical to the ones of theorems (3.4.1) and (3.4.2), but uses propositions (3.4.4) and (3.4.3) below instead of propositions (3.4.1) and (3.4.2).

Proposition 3.4.3. *Let the regularity assumptions (i)-(vii) on the density and the kernel be satisfied,then, for a compact set* $D \subset (0,1)^2, h_n \to 0$ *and* $nh_n^3/\ln n \to \infty$ *entails*

$$\sup_{(x,y)\in D} |\hat{c}_n(F_n(x), G_n(y)) - c_n(F(x), G(y))| = O_P\left(\frac{1}{nh_n^4} + \frac{\ln n}{\sqrt{n}}\right)$$

$$\sup_{(x,y)\in D} |\hat{c}_n(F_n(x), G_n(y)) - c_n(F(x), G(y))| = O_{a.s}\left(\frac{\ln n\sqrt{\ln\ln n}}{\sqrt{n}} + \frac{\ln\ln n}{nh_n^4}\right)$$

Proposition 3.4.4. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied,then, for a compact set* $D \subset (0,1)^2$ *and* $h_n \simeq (\ln n/n)^{1/6}$ *, one has*

$$\sup_{(u,v)\in D} |\hat{c}_n(u,v) - c_n(u,v)| = O_P\left(\left(\frac{\ln n}{n}\right)^{1/3}\right) = O_{a.s}\left(\left(\frac{\ln n}{n}\right)^{1/3}\right)$$

Bibliography

[1] Biau,G., AND Wegkamp,M.H.," A note on minimum distance estimation of copulas densities"., *Statist. Probab. Lett.*, No.73, (2006), pp. 105-114.

[2] Bosq, D., *Nonparametric statistics for stochastic processes*, second ed., Vol. 110 of Lecture Notes in Statistics. Springer-Verlag, New York, (1998). Estimation and prediction.

[3] Carmona, R.A., *Statistical Analysis of Financial Data in S-PLUS*, Springer-Verlag, New York (2004).

[4] Cherubini, U., Luciano, E., AND Vecchiato, W., *Copula Methods in Finance.* John Wiley and Sons, Chichester (2004).

[5] Deheuvels, P., "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance". *Acad. Roy. Belg. Bull. Cl. Sci.*, Vol.65,No.5 And 6, (1979),pp. 274-292.

[6] Deheuvels, P., "A Kolmogorov-Smirnov type test for independence and multivariate samples", *Rev. Roumaine Math. Pures Appl.*Vol 26,No.2, (1981), pp. 213-226.

[7] Embrechts, P., Lindskog, F., McNeil, A., *Modelling dependence with copulas and applications to risk management.*, Handbook of Heavey Tailed Distributions in Finance, S.T. Rachev (ed.). Elsevier, Amsterdam (2003).

[8] Faugeras, Olivier Paul., *Contributions à la prévision statistique*, Thèses de doctorat de l'université Pierre er Marie Curie, 28 Novembre 2008.

[9] Fermanian, J.D., "Goodness-of-fit tests for copulas.", *J. Multivariate Anal.* Vol.95, No.1, (2005), pp. 119-152.

[10] Fermanian, J.D., Radulovïc, D., AND Wegkamp, M., "Weak convergence of empirical copula processes". Bernoulli Vol.10,No.5, (2004), pp. 847-860.

[11] Fermanian, J.D., AND Scaillet, O., *Some statistical pitfalls in copula modelling for financical applications*, E. Klein (Ed.), Capital formation, Gouvernance and Banking. Nova Science Publishing, New York, (2005).

[12] Fermanian, J.D.,AND Scaillet, O., "Nonparametric estimation of copulas for time series", *Journal of Risk*, Vol.5,No. 4, (2003), pp. 25-54.

[13] Ferraty, F., AND Vieu, P., *Nonparametric functional data analysis.* Springer Series in Statistics. Springer, New York, (2006). Theory and practice.

[14] Fisher, N.I., *Copulas*, Encyclopedia of Statistical Sciences, Updated Volume 1, S. Kotz, C. B. Read, D. L. Banks (eds.), pp. 159-164. John Wiley and Sons, New York (1997).

[15] Gijbels, I., AND Mielniczuk, J., "Estimating the density of a copula function", *Comm. Statist. Theory Methods* ,Vol.19,No.2 (1990), pp.445-464.

[16] Hutchinson, T.P., Lai, C.D., *Continuous Bivariate Distributions: Emphasising Applications*,Rumsby Scientific Publishing, Adelaide (1990).

[17] Joe, H., *Multivariate Models and Dependence Concepts*, Chapman and Hall, London (1997).

[18] Joe, H., *Multivariate models and dependence concepts*, vol. 73 of Monographs on Statistics and Applied Probability. Chapman & Hall, London, (1997).

[19] Mikosch, T., "Copulas: Taies and facts", *Extrêmes*,Vol. 9,No.18,(2006), pp.3-22 .

[20] Nelsen, R.B., *An Introduction to Copulas*, 2nd edition. Springer-Verlag, New York (2006).

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

[22] Parzen, E., "On estimation of a probability density function and mode", *Ann. Math. Statist.33*, (1962) , pp.1065-1076.

[23] Pickands, III, J., "Statistical inference using extreme order statistics", *Ann. Statist.3*,(1975), pp.119-131.

[24] Rüschendorf, L.," Asymptotic distributions of multivariate rank order statistics", *Ann. Statist. 4*, 5 (1976), pp.912-923.

[25] Sancetta,A., "Nonparametric estimation of multivariate distributions with given marginals: L_2 theory",*Cambridge Working papers in Economics* No.0320, (2003).

[26] Sancetta, A., Satchell,S., "The Barnestien coplua and its application to modelling and approximation of multivariate distributions", *Econometric theory*, 20, (2004), pp.535-562.

[27] Scott, D. W., *Multivariate density estimation. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley & Sons Inc., New York, (1992). Theory, practice, and visualization, A Wiley-Interscience Publication.

[28] Shorack, G. R., AND Wellner, J. A., *Empirical processes with applications tostatistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons Inc., New York, (1986).

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

[30] The International Actuarial Association, A global framework for insurer solvency assessment, Research Report. www.actuaries.org (2004)

[31] Van Der Vaart, A. W., AND Wellner, J. A., *Weak convergence and empirical processes*, Springer Series in Statistics. Springer-Verlag, New York, (1996). With applications to statistics.

Conclusion and Prospects

In this thesis we have introduced and studied a new approach to nonparametric estimation of regression. We modeled the regression by the transformed copula and estimated this model by the kernel method. This model is based on an efficient data transformation by transform quantile and the use of the copula representation it was found to have a remarkable product form.

The concept of copula was discussed, as their essential properties. Our study allowed us to establish convergence Probability and almost sure (with rates) of regression model through the copula function approach and asymptotic normality of the model.

Thus, it was found that K_{h_n} depends on the number of observations n. In some specific situations, the sample size is fluctuating, so the regression is estimated by our model, an increase of this size, even a few observations, leads to completely recalculate the estimator and this may be an additional computational load and lost considerable time even for powerful computers.

Therefore, this theory is not used to its full potential, it offers plenty of opportunities for research that should be explored, especially in the areas of: Modeling in dimension $n(n > 2)$, the survival function, hazard function and then it would be very interesting to extend our work to the recursive methods since they offer a kind of balance between accuracy and speed of calculation.

Appendix

Proof Propsition 3.4.1

Let

$$c_n(u,v) = \frac{1}{nh_n^2}\sum_{i=1}^{n} K_1\left(\frac{u-F(X_i)}{h_n}\right) K_2\left(\frac{v-G(Y_i)}{h_n}\right),$$

$$\hat{c}_n(u,v) = \frac{1}{nh_n^2}\sum_{i=1}^{n} K_1\left(\frac{u-F_n(X_i)}{h_n}\right) K_2\left(\frac{v-G_n(Y_i)}{h_n}\right).$$

So, we must show that $F_n(X_i)$ converge to $F(X_i)$ and $G_n(Y_i)$ converge to $G(Y_i)$.

$$\hat{c}_n(u,v) - c_n(u,v) = \frac{1}{nh_n^2}\sum_{i=1}^{n}\left[K_1\left(\frac{u-F_n(X_i)}{h_n}\right) K_2\left(\frac{v-G_n(Y_i)}{h_n}\right) - K_1\left(\frac{u-F(X_i)}{h_n}\right)\right.$$

$$\left. K_2\left(\frac{v-G(Y_i)}{h_n}\right)\right]$$

with

$$\Pi_{i,n} = K_1\left(\frac{u-F_n(X_i)}{h_n}\right) K_2\left(\frac{v-G_n(Y_i)}{h_n}\right) - K_1\left(\frac{u-F(X_i)}{h_n}\right) K_2\left(\frac{v-G(Y_i)}{h_n}\right)$$

Let

$$Z_{i,n} = \begin{pmatrix} F_n(X_i) - F(X_i) \\ G_n(Y_i) - G(Y_i) \end{pmatrix}$$

$|F_n(X_i) - F(X_i)| \leq \|F_n - F\|_\infty$ *and* $|G_n(Y_i) - G(Y_i)| \leq \|Gn - G\|_\infty$ *a.s. for every* $i = 1, ..., n$. *Preceding Lemma thus entails that the norm of* $Z_{i,n}$ *is independent of* i *and* *such that*

$$\|Z_{i,n}\| = O_P\left(\frac{1}{\sqrt{n}}\right), \quad i = 1, ..., n, \tag{3.7}$$

$$\|Z_{i,n}\| = O_{a.s}\left(\sqrt{\frac{lnlnn}{n}}\right) \quad i = 1, ..., n. \tag{3.8}$$

Now, for every fixed $(u, v) \in [0, 1]^2$, *since the kernel* K *is twice differentiable, there exists, by Taylor expansion, random variables* $\tilde{U}_{i,n}$ *and* $\tilde{V}_{i,n}$ *such that, almost surely,*

$$\begin{aligned} \Pi &= \tfrac{1}{nh_n^3}\textstyle\sum_{i=1}^n Z_{i,n}^T \nabla K\left(\frac{u-F(X_i)}{h_n}, \frac{v-G(Y_i)}{h_n}\right) + \tfrac{1}{2nh_n^4}\textstyle\sum_{i=1}^n Z_{i,n}^T \nabla^2 K\left(\frac{u-\tilde{U}_{i,n}}{h_n}, \frac{v-\tilde{V}_{i,n}}{h_n}\right) Z_{i,n} \\ &:= \Pi_1 + \Pi_2 \end{aligned}$$

where $Z_{i,n}^T$ *denotes the transpose of the vector* $Z_{i,n}$ *and* ∇K *and* $\nabla^2 K$ *the gradient and the Hessian respectively of the multivariate kernel function* K.

By centering at expectations, decompose further the first term Π_1 *as,*

$$\begin{aligned} \Pi_1 &= \tfrac{1}{nh_n^3}\textstyle\sum_{i=1}^n Z_{i,n}^T\left(\nabla K\left(\frac{u-F(X_i)}{h_n}, ...\right) - \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_n}, ...\right)\right) \\ &+ \tfrac{1}{nh_n^3}\textstyle\sum_{i=1}^n Z_{i,n}^T \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_n}, \frac{v-G(Y_i)}{h_n}\right) \\ &:= \Pi_{11} + \Pi_{12} \end{aligned}$$

We again decompose one step further Π_{11}, *Set*

$$h_i = \nabla K\left(\frac{u-F(X_i)}{h_n}, \frac{v-G(Y_i)}{h_n}\right) - \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_n}, \frac{v-G(Y_i)}{h_n}\right)$$

Then

$$|\Pi_{11}| \leq \frac{\|Z_{i,n}\|}{nh_n^3}\sum_{i=1}^n (\|h_i\| - \mathbb{E}\|A_i\|) + \frac{\|Z_{i,n}\|}{nh_n^3}\sum_{i=1}^n \mathbb{E}\|A_i\| = \Pi_{111} + \Pi_{112}.$$

We now proceed to the study of the order of each terms in the previous decompositions.

• **Negligibility of** Π_2.

By the boundedness assumption on the second-order derivatives of the kernel, and equations (3.7) and (3.8),

$$\Pi_2 = O_P\left(\frac{1}{nh_n^4}\right), \text{ and } \Pi_2 = O_{a.s}\left(\frac{\ln\ln n}{nh_n^4}\right)$$

• **Negligibility of** Π_{12}.

Bias results on the bivariate gradient kernel estimator (See Scott [?] chapter 6) entail that

$$\mathbb{E}\nabla\left(K_1\left(\frac{u-F(X_i)}{h_n}\right)K_2\left(\frac{v-G(Y_i)}{h_n}\right)\right) = h_n^3\nabla c(u,v) + O(h_n^5)$$

Cauchy-Schwarz inequality yields that

$$|\Pi_{12}| \leq \frac{n\|Z_{i,n}\|}{nh_n^3}\left\|\mathbb{E}\nabla\left(K_1\left(\frac{u-F(X_i)}{h_n}\right)K_2\left(\frac{v-G(Y_i)}{h_n}\right)\right)\right\|$$

In turn, with equations (3.7) and (3.8),

$$\Pi_{12} = O_P\left(\frac{1}{\sqrt{n}}\right), \text{ and } \Pi_{12} = O_{a.s}\left(\sqrt{\frac{\ln\ln n}{n}}\right)$$

• **Negligibility of** Π_{11}

• **Negligibility of** Π_{111}.

Boundedness assumption on the derivative of the kernel imply that $\|A_i\| \leq 2\alpha$ *a.s. We apply Hoeffding inequality for independent, centered, bounded by M, but non identically distributed random variables* (η_j) *(e.g. see [?]),*

$$\mathbb{P}\left(\sum_{j=1}^{n}\eta_j > t\right) \leq exp\left(\frac{-t^2}{2nM^2}\right)$$

Here, for every $\epsilon > 0$, *with* $M = 2\alpha$, $\eta_j = \|A_i\| - \mathbb{E}\|A_i\|$, $t = \epsilon\sqrt{\frac{1}{n}\ln\ln n}$, *Therefore,*

$$\sum_{i=1}^{n}(\|A_i\| - \mathbb{E}\|A_i\|) = O_P(\sqrt{n \ln \ln n})$$

which is the definition of almost complete convergence (a.co.), see e.g. [?] definition A.3. p. 230. In turn, it means that

$$\sum_{i=1}^{n}(\|A_i\| - \mathbb{E}\|A_i\|) = O_{a.co}\left(\sqrt{n \ln n}\right)$$

and by the Borell-Cantelli lemma,

$$\sum_{i=1}^{n}(\|A_i\| - \mathbb{E}\|A_i\|) = O_{a.s}(\sqrt{n \ln n})$$

Therefore, using equations (3.7) and (3.8), we have that

$$\Pi_{111} = O_P\left(\frac{\sqrt{\ln \ln n}}{nh_n^3}\right) = O_{a.s}\left(\frac{\sqrt{\ln n \ln \ln n}}{nh_n^3}\right)$$

- **Negligibility of** Π_{112}

The r.h.s. of the previous inequality is, after an integration by parts, of order a_n^3 *by the results on the kernel estimator of the gradient of the density (See Scott [?] chapter 6). Therefore,*

$$\sum_{i=1}^{n}\mathbb{E}\|A_i\| = O(nh_n^3)$$

and

$$\Pi_{112} = \frac{\|Z_{i,n}\|}{nh_n^3}\sum_{i=1}^{n}\mathbb{E}\|A_i\| = O_P\left(\frac{1}{\sqrt{n}}\right) = O_{a.s}\left(\sqrt{\frac{\ln \ln n}{n}}\right)$$

by equations (3.7) and (3.8) Recollecting all elements, we eventually obtain that

$$\begin{aligned}
\Pi &= \Pi_{111} + \Pi_{112} + \Pi_{12} + \Pi_2 \\
&= O_P(n^{-1/2}) + O_P(\tfrac{\sqrt{\ln\ln n}}{na_n^3}) + O_P(\tfrac{1}{na_n^4}) \\
or \quad &= O_{a.s}\left(\tfrac{\sqrt{lnn \ln\ln n}}{na_n^3}\right) + O_{a.s}\left(\sqrt{\tfrac{\ln\ln n}{n}}\right) + O_{a.s}\left(\tfrac{\ln\ln n}{na_n^4}\right)
\end{aligned}$$

By this last step we conclude the proof of our theorem.

Proof Proposition 3.4.2

We proceed as in the previous proposal. and we get

$$\begin{aligned}
\Pi'(x,y) &= O_P(\tfrac{1}{\sqrt{n}} + \tfrac{1}{nh_n^4}) \\
or \quad &= O_{a.s}(\sqrt{\tfrac{\ln\ln n}{n}} + \tfrac{n}{nh_n^4})
\end{aligned}$$

By this last step we conclude the proof of our theorems.

■

Proof Propostion 3.4.3

Set

$$W_{i,n} = \nabla K\left(\frac{u - F(X_i)}{h_n}, \frac{v - G(Y_i)}{h_n}\right)$$

By Taylor expansions, we still have the decomposition

$$\begin{aligned}
\Pi(x,y) &= \tfrac{Z_n^T(x,y)}{nh_n^3}\textstyle\sum_{i=1}^n W_{i,n}(F(x),G(y)) \\
&+ \tfrac{Z_n^T(x,y)}{2nh_n^4}\textstyle\sum_{i=1}^n Z_{i,n}^T(x,y)\nabla^2 K\left(\tfrac{u-F(X_i)}{h_n}, \tfrac{v-G(Y_i)}{h_n}\right) \\
&+ \tfrac{\|Z_n\|^2}{h_n^4}R_3
\end{aligned}$$

with the remainder term $R_3 = O_{a.s}(1)$ uniformly. By bounding the $\nabla^2 K$, and using the properties of the Kolmogorov-Smirnov statistic, the last two terms are of order

$O_P\left(\frac{1}{nh_n^4}\right)$, *or* $O_{a.s}\left(\frac{\ln\ln n}{nh_n^4}\right)$, *uniformly in* x, y. *For the first term, by Cauchy-Schwarz inequality,*

$$\sup_{(x,y)\in D}\left|\frac{Z_n^T(x,y)}{nh_n^3}\sum_{i=1}^n W_{i,n}(F(x),G(y))\right| \leq \|Z_n\| \sup_{(x,y)\in D}\left\|\frac{1}{nh_n^3}\sum_{i=1}^n W_{i,n}(F(x),G(y))\right\|$$

The convergence results of the kernel estimator $n^{-1}h_n^{-3}\sum_{i=1}^n W_{i,n}(u,v)$ *of the gradient of the density* $c(u,v)$ *can easily be derived from those of the kernel estimator (see Scott [?]). From the convergence results uniformly on a compact set of the latter obtained by e.g. Deheuvels [41] for the almost sure rates and Bickel and Rosenblatt [16] for the in probability rates, with the assumption that the gradient is uniformly bounded on* D *and that* $nh_n^3/\ln n \to \infty$, *one gets that the uniform norm of the estimator of the gradient is an* $O_P(\ln n)$ *or an* $O_{a.s}(\ln n)$. *In turn,* $\sup_{(x,y)\in D}|\Pi(x,y)| = O_P(\ln n/n^{-1/2})$ *or* $O_{a.s}(\ln n(\ln\ln n/n)^{1/2})$. *Thus the claimed result.*

■

Proof Propostion 3.4.4

For convenience, set $\|(x_1,...,x_d)\| = \max_{1\leq j\leq d}|x_j|$. *Set* $D = [u_0,u_\infty]\times[v_0,v_\infty] \subset (0,1)^2$ *a compact subset where* $0 < u_0 \leq u_\infty < 1$ *and* $0 < v_0 \leq v_\infty < 1$. *We mimic the proof of proposition 3.4.1. We still have the additive decomposition,*

$$\begin{aligned}\Pi(u,v) &= \Pi_1(u,v) + \Pi_2(u,v)\\ &= \Pi_{11}(u,v) + \Pi_{12}(u,v) + \Pi_2(u,v)\end{aligned}$$

with

$$\Pi_{11}(u,v) = \frac{1}{nh_n^3}\sum_{i=1}^n Z_{i,n}(W_{i,n}(u,v) - EW_{i,n}(u,v))$$

$$\Pi_{12}(u,v) = \frac{1}{nh_n^3}\sum_{i=1}^n Z_{i,n}EW_{i,n}(u,v)$$

and

$$W_{i,n} = \nabla K\left(\frac{u-F(X_i)}{h_n}, \frac{v-G(Y_i)}{h_n}\right)$$

• **Negligibility of** Π_2.

The proof remains the same:

$$\sup_{(x,y)\in D} |\Pi_2(u,v)| = O_P\left(\frac{1}{nh_n^4}\right), \text{ and } \Pi_2 = O_{a.s}\left(\frac{\ln\ln n}{nh_n^4}\right)$$

• **Negligibility of** Π_{12}.

Recall that in the Taylor's expansion of the bias of the kernel estimator, the $O(.)$ *is uniform in* (u,v)*, therefore one gets that*

$$\sup_{(u,v)\in D} \|EW_{i,n}(u,v) - h_n^3 \nabla c(u,v)\| = O(h_n^5)$$

Thus,

$$\sup_{(u,v)\in D} |\Pi_{12}(u,v)| = O_P(1/\sqrt{n}), \quad or \quad O_{a.s}((\ln\ln n/n)^{1/2}).$$

• **Negligibility of** Π_{11}.

Define a covering of D *by* M_n^2 *compact hypercubes* D_k *centered in* (u_k, v_k)*,*

$$D_k = \{(u,v) \in D : \|(u,v) - (u_k,v_k)\| \le 1/M_n\}, 1 \le k \le M_n^2$$

One can write

$$\begin{aligned}
\sup_{(x,y)\in D} |\Pi_{11}(u,v)| &\le \max_{1\le k\le M_n^2} \sup_{(x,y)\in D} |\Pi_{11}(u,v) - \Pi_{11}(u_k,v_k)| \\
&+ \max_{1\le k\le M_n^2} |\Pi_{11}(u_k,v_k)| \\
&:= (I) + (II)
\end{aligned}$$

• **Negligibility of** (I)

For (I)*, by boundedness and Lipshitz assumption on the product kernel* K*, there exists a constant* ζ *such that,*

$$\|\nabla K(u,v) - \nabla K(u_k,v_k)\| \le \zeta \|(u,v) - (u_k,v_k)\|$$

Therefore for $(u,v) \in D_k$,

$$\left\| \nabla K\left(\frac{u-F(X_i)}{h_n}, \frac{v-G(Y_i)}{h_n}\right) - \nabla K\left(\frac{u_k-F(X_i)}{h_n}, \frac{v_k-G(Y_i)}{h_n}\right)\right\| \leq \frac{\zeta}{M_n h_n}$$

since K is product-shaped. In turn, the same bound is valid by Jensen's inequality for the expectations of the difference, so that

$$(I) \leq \frac{2\zeta\|Z_n\|}{M_n h_n^4}$$

Setting $M_n = n^{1/2}h_n^3 \simeq n/\sqrt{\ln n}$ for $h_n \simeq (\ln n/n)^{1/6}$, one has that $(I) = O_{a.s}\left(\sqrt{\frac{\ln n}{nh_n^2}}\right)$ or $O_P((nh_n^2)^{-1/2})$.

• **Negligibility of** (II)

For the second term, set as before, $A_i(u,v) = W_{i,n}(u,v) - EW_{i,n}(u,v)$, and majorize, for each k,

$$\begin{aligned} |\Pi_{11}(u_k,v_k)| &\leq \tfrac{\|Z_n\|}{nh_n^3} \textstyle\sum_{i=1}^n \|A_i(u_k,v_k)\| \\ &\leq \tfrac{\|Z_n\|}{nh_n^3} \textstyle\sum_{i=1}^n (\|A_i(u_k,v_k)\| - E\|A_i(u_k,v_k) + E\|A_i(u_k,v_k)\|) \\ &\leq \tfrac{\|Z_n\|}{nh_n^3} \textstyle\sum_{i=1}^n \eta_i(u_k,v_k) + \tfrac{\|Z_n\|}{nh_n^3} \textstyle\sum_{i=1}^n E\|A_i(u_k,v_k)\| \end{aligned}$$

where we have set $\eta_i(u_k,v_k) = \|A_i(u_k,v_k)\| - E\|A_i(u_k,v_k)$. For the expectation term, as the product kernel is of finite variation, and with the assumption that the gradient of the copula density remains bounded on D, one has that $\max_{1\leq k\leq M_n^2} \sum_{i=1}^n E\|A_i(u_k,v_k)\| = O(h_n^3)$. This yields that

$$\max_{1\leq k\leq M_n^2} \frac{\|Z_n\|}{nh_n^3} \sum_{i=1}^n E\|A_i(u_k,v_k)\| = O_P(n^{-1/2}), \quad or \quad O_{a.s}\left(\left(\frac{\ln\ln n}{n}\right)^{1/2}\right)$$

It remains to deal with the deviation term

$$\frac{\|Z_n\|}{nh_n^3}\sum_{i=1}^{n}\eta_i(u_k,v_k)$$

We have

$$P\left(\max_{1\leq k\leq M_n^2}\left|\sum_{i=1}^{n}\eta_i(u_k,v_k)\right|>\epsilon\right)\leq\sum_{k=1}^{M_n^2}P\left(\left|\sum_{i=1}^{n}\eta_i(u_k,v_k)\right|>\epsilon\right)$$

and apply Hoeffding's inequality to the summand, to get that, for every $\epsilon>0$,

$$P\left(\left|\sum_{i=1}^{n}\eta_i(u_k,v_k)\right|>\epsilon\sqrt{n\ln n}\right)\leq M_n^2e^{-\frac{\epsilon^2\ln n}{\zeta}}\leq e^{\sqrt{2\ln M_n}-\frac{\epsilon^2\ln n}{\zeta}}$$

For $h_n\simeq(\ln n/n)^{1/6}$ *and* $M_n=n^{1/2}h_n^{-3}\simeq n/\sqrt{\ln n}$,

$$e^{\sqrt{2\ln M_n}-\frac{\epsilon^2\ln n}{\zeta}}\approx e^{-\frac{\epsilon^2\ln n}{\zeta}}=\frac{1}{n^{\epsilon^2/\zeta}}$$

which is absolutely summable for an ϵ *large enough. Therefore,*

$$\max_{1\leq k\leq M_n^2}\left|\sum_{i=1}^{n}\eta_i(u_k,v_k)\right|=O_{a.co}(\sqrt{n\ln n})$$

and eventually,

$$\frac{\|Z_n\|}{nh_n^3}\max_{1\leq k\leq M_n^2}\left|\sum_{i=1}^{n}\eta_i(u_k,v_k)\right|=O_{a.s}\left(\frac{\sqrt{\ln n\ln\ln n}}{nh_n^3}\right)$$

for the choice $h_n\simeq(\ln n/n)^{1/6}$.

Recollecting all elements gives the claimed result with the given choice of h_n.

General Bibliography

Beneš V, Štěpn J., eds *Distributions with Given Marginals and Moment Problems.* Kluwer, Dordrecht(1997)

Biau,G., AND Wegkamp,M.H.," A note on minimum distance estimation of copulas densities"., *Statist. Probab. Lett.*, No.73, (2006), pp. 105-114.

Bosq, D., *Nonparametric statistics for stochastic processes*, second ed., Vol. 110 of Lecture Notes in Statistics. Springer-Verlag, New York, (1998). Estimation and prediction.

Carmona, R.A., *Statistical Analysis of Financial Data in S-PLUS*, Springer-Verlag, New York (2004).

Cherubini, U., Luciano, E., AND Vecchiato, W., *Copula Methods in Finance.* John Wiley and Sons, Chichester (2004).

Conway. DA, *Multivariate Distributions with Specified Marginals.* (Technical Report 145, Stanford University)(1979)

Conway DA "Plackett family of distributions". *In: Kotz S, Johnson NL (eds) Encyclopedia of Statistical Sciences*, Vol 7. Wiley, New York, (1986) pp 1-5

Cuadras CM, Fortiana J, Rodríguez-Lallena JA eds *Distributions with Given Marginals and Statistical Modelling.* Kluwer, Dordrecht(2002)

Dall'Aglio, G., Fréchet classes: the beginnings". *In: Dall'Aglio G*, Kotz S, Salinetti, G.,(eds) "Advances in Probability Distributions with Given Marginals.", Kluwer, Dordrecht,(1991) pp 1-12

Daniels, HE., "Rank correlation and population models.", (1950), *J Roy Statist Soc Ser*

B, 12:171-181

Darsow, WF., Nguyen B, Olsen, ET., "Copulas and Markov processes.", *Illinois J Math*, (1992),36:600-642

Deheuvels, P., "Caractérisation complète des lois extrêmes multivariées et de la convergence des types extremes". *Publ Inst Statist Univ Paris* (1978) 23:1-37

Deheuvels, P., "La fonction de dépendence empirique et ses propriétés. Un test non paramétrique d'indépendence.", *Acad Roy Belg Bul Cl Sci* (5) (1979), 65:274-292

Deheuvels, P., "A Kolmogorov-Smirnov type test for independence and multivariate samples.", *Rev Roumaine Math Pures Appl*,(1981a), 26:213-226

Deheuvels, P., "A non parametric test for independence." *Publ Inst Statist Univ Paris*, (1981b),26:29-50

Deheuvels, P., "Multivariate tests of independence." In: Analytical Methods in Probability (Oberwolfach, 1980) *Lecture Notes in Mathematics 861. Springer*, Berlin,(1981c), pp 42-50

Deheuvels, P., "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance". *Acad. Roy. Belg. Bull. Cl. Sci.*, Vol.65,No.5 And 6, (1979),pp. 274-292.

Deheuvels, P., "A Kolmogorov-Smirnov type test for independence and multivariate samples", *Rev. Roumaine Math. Pures Appl.*Vol 26,No.2, (1981), pp. 213-226.

Durbin J, "Stuart A Inversions and rank correlations"., *J Roy Statist Soc Ser B*, (1951), 13:303-309

Embrechts, P., Lindskog, F., McNeil, A., *Modelling dependence with copulas and applications to risk management.*, Handbook of Heavey Tailed Distributions in Finance, S.T. Rachev (ed.). Elsevier, Amsterdam (2003).

Fan, J., and Yao, Q. *Nonlinear time series*, second ed. Springer Series in Statistics. Springer-Verlag, New York, (2005). Nonparametric and parametric methods.

Faugeras, Olivier Paul., *Contributions à la prévision statistique*, Thèses de doctorat de l'université Pierre er Marie Curie, 28 Novembre 2008.

Fermanian, J.D., "Goodness-of-fit tests for copulas.", *J. Multivariate Anal.* Vol.95, No.1, (2005), pp. 119-152.

Fermanian, J.D., Radulovïc, D., AND Wegkamp, M., "Weak convergence of empirical copula processes". Bernoulli Vol.10,No.5, (2004), pp. 847-860.

Fermanian, J.D., AND Scaillet, O., *Some statistical pitfalls in copula modelling for financical applications*, E. Klein (Ed.), Capital formation, Gouvernance and Banking. Nova Science Publishing, New York, (2005).

Fermanian, J.D.,AND Scaillet, O., "Nonparametric estimation of copulas for time series", *Journal of Risk*, Vol.5,No. 4, (2003), pp. 25-54.

Ferraty, F., AND Vieu, P., *Nonparametric functional data analysis.* Springer Series in Statistics. Springer, New York, (2006). Theory and practice.

Féron, R., " Sur les tableaux de corrélation dont les marges sont données, cas de l'espace à trois dimensions". *Publ Inst Statist Univ Paris.*, (1956), 5:3-12

Fisher NI, Sen PK eds *The Collected Works of Wassily Hoeffding.* Springer, New York, (1994)

Fisher, N.I., *Copulas*, Encyclopedia of Statistical Sciences, Updated Volume 1, S. Kotz, C. B. Read, D. L. Banks (eds.), pp. 159-164. John Wiley and Sons, New York (1997).

Galambos, J., *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York (1978)

Genest, C., MacKay, J., "Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données." *Canad J Statist* (1986a) 14:145-159

Genest, C., MacKay, J., "The joy of copulas: Bivariate distributions with uniform marginals." *Amer Statist*(1986b) 40:280-285

Gijbels, I., AND Mielniczuk, J., "Estimating the density of a copula function", *Comm. Statist. Theory Methods* ,Vol.19,No.2 (1990), pp.445-464.

Hoeffding, W., "Masstabinvariante Korrelationstheorie.", *Schriften des Matematischen Instituts und des Instituts für Angewandte Mathematik* der Universität Berlin 5 Heft 3:179-233 [Reprinted as Scale-invariant correlation theory. In: Fisher NI, Sen PK (eds) The Collected Works of Wassily Hoeffding. Springer, New York, (1940), pp 57-107]
Hoeffding, W., "Masstabinvariante Korrelationsmasse für diskontinuierliche Verteilungen."; *Arkiv für matematischen Wirtschaften und Sozialforschung* 7:49-70 [Reprinted as Scale-invariant correlation measures for discontinuous distributions. In: Fisher NI, Sen PK (eds) The Collected Works of Wassily Hoeffding. Springer, New York, (1941), pp 109-133]

Hollander,. M., Wolfe, DA., *Nonparametric Statistical Methods.* Wiley, New York (1973)

Hutchinson, T.P., Lai, C.D., *Continuous Bivariate Distributions: Emphasising Applications,*Rumsby Scientific Publishing, Adelaide (1990).

Joe, H., "Parametric families of multivariate distributions with given margins.", *J Multivariate Anal*, (1993), 46:262-282

Joe, H., *Multivariate Models and Dependence Concepts*, Chapman and Hall, London (1997).

Joe, H., *Multivariate models and dependence concepts*, vol. 73 of Monographs on Statistics and Applied Probability. Chapman & Hall, London, (1997).

Jogdeo, K., "Concepts of dependence". *In: Kotz S, Johnson NL (eds) Encyclopedia of Statistical Sciences*, Vol 1. Wiley, New York,(1982), pp 324-334

Kamiński, A., Sherwood, H., Taylor, MD., "Doubly stochastic measures with mass on the graphs of two functions.", *Real Anal Exchange*, (1987-88), 13:253-257

Kimeldorf, G., Sampson, A., "Uniform representations of bivariate distributions"., *Comm Statist A-Theory Methods*, (1975b),4:617-627

Kruskal, WH., "Ordinal measures of association.", *J Amer Statist Assoc*, (1958), 53:814-861

Lehmann, EL., "Some concepts of dependence.", *Ann Math Statist*, (1966), 37:1137-1153.

Lehmann, EL., *Nonparametrics: Statistical Methods Based on Ranks.*, Holden-Day, San Francisco, (1975)

Mardia, KV., *Families of Bivariate Distributions.*, Griffin, London, (1967)

Marshall, AW., Olkin, I., "Families of multivariate distributions.", *J Amer Statist Assoc*,(1988), 83:834-841

Menger, K., "Statistical metrics". *Proc Natl Acad Sci USA*, (1942),28:535-537

Mikusiń́ski, P., Sherwood, H., Taylor, MD., "Probabilistic interpretations of copulas and their convex sums". *In: Dall'Aglio G, Kotz S, Salinetti G* (eds) "Advances in Probability Distributions with Given Marginals". Kluwer, Dordrecht,(1991), pp 95-112

Mikusiń́ski, P., Sherwood, H., Taylor, MD., "The Fréchet bounds revisited.", *Real Anal Exchange*, (1991-92), 17:759-764

Mikusiń́ski, P., Sherwood, H., Taylor, MD., "Shuffles of Min". *Stochastica*, (1992),13:61-74

Mikosch, T., "Copulas: Taies and facts", *Extrêmes*,Vol. 9,No.18,(2006), pp.3-22 .

Moore, DS., Spruill, MC., "Unified large-sample theory of general chi-squared statistics for tests of fit". *Ann Statist*,(1975), 3:599-616

Nelsen, R.B., *An Introduction to Copulas*, 2nd edition. Springer-Verlag, New York (2006).

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

Parzen, E., "On estimation of a probability density function and mode", *Ann. Math. Statist.33*, (1962) , pp.1065-1076.

Pickands, III, J., "Statistical inference using extreme order statistics", *Ann. Statist.3*,(1975), pp.119-131.

Rüschendorf, L.," Asymptotic distributions of multivariate rank order statistics", *Ann. Statist. 4*, 5 (1976), pp.912-923.

Rüschendorf, L., Schweizer, B., Taylor, MD., eds *Distributions with Fixed Marginals and Related Topics.* Institute of Mathematical Statistics, Hayward, CA, (1996)

Sancetta,A., "Nonparametric estimation of multivariate distributions with given marginals: L_2 theory",*Cambridge Working papers in Economics* No.0320, (2003).

Sancetta, A., Satchell,S., "The Barnestien coplua and its application to modelling and approximation of multivariate distributions", *Econometric theory*, 20, (2004), pp.535-562.

Scott, D. W., *Multivariate density estimation. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley & Sons Inc., New York, (1992). Theory, practice, and visualization, A Wiley-Interscience Publication.

Scarsini, M., "On measures of concordance"., *Stochastica*, (1984), 8:201-218

Schweizer, B., Sklar, A., "Operations on distribution functions not derivable from operations on random variables". *Studia Math* (1974) 52:43-52

Schweizer, B., Sklar, A., *Probabilistic Metric Spaces.* North-Holland, New York (1983)

Schweizer, B., Wolff, EF., "On nonparametric measures of dependence for random variables"., *Ann Statist*, (1981),9:879-885

Schweizer, B., "Thirty years of copulas". *In: Dall'Aglio G, Kotz S, Salinetti G (eds) Advances in Probability Distributions with Given Marginals.* Kluwer, Dordrecht,(1991) pp 13-50

Seeley, RT., "Fubini implies Leibniz implies $F_{yx} = F_{xy}$". Am Math Monthly, (1961),68:56-57

Sherwood, H., Taylor, MD., "Doubly stochastic measures with hairpin support". *Probab*

Theory Related Fields,(1988), 78:617-626

Shorack, G. R., AND Wellner, J. A., *Empirical processes with applications tostatistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* John Wiley & Sons Inc., New York, (1986).

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

Sklar, A., "Random variables, joint distributions, and copulas". *Kybernetica*, (1973),9:449-460.

Sklar, A., "Random variables, distribution functions, and copulas a personal look backward and forward". In: Rüschendorf L, Schweizer B, Taylor MD (eds) Distributions with Fixed Marginals and Related Topics. *Institute of Mathematical Statistics, Hayward, CA*,(1996), pp 1-14

Sklar, A., personal communication,(1998)

The International Actuarial Association, A global framework for insurer solvency assessment, Research Report. www.actuaries.org (2004)

Van Der Vaart, A. W., AND Wellner, J. A., *Weak convergence and empirical processes*, Springer Series in Statistics. Springer-Verlag, New York, (1996). With applications to statistics.