UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN ECONOMICS AND FINANCE

TESI DI LAUREA

APPLYING COPULAS IN ECONOMETRICS: ESTIMATE OF PORTFOLIO VALUE AT RISK

RELATORE:

CH.MO PROF. NUNZIO CAPPUCCIO

LAUREANDA: ELEONORA MAGLIANI

MATRICOLA N. 1130641

ANNO ACCADEMICO 2017 – 2018

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell’elaborato sono stati indicati nel testo e nella sezione “Riferimenti bibliografici” e che le eventuali citazioni testuali sono individuabili attraverso l’esplicito richiamo alla pubblicazione originale.

Firma dello studente

____

ABSTRACT
Copula methods are spreading in finance, due to their capacity of handling co-movements in market factors and describe interdependent risk. Substantially, a copula is the joint distribution of a vector of uniform random variables. These random variables could be assets composing a portfolio. Copulas allow to analyse the type of dependence that exists among the assets, keeping individual asset characteristics separated from joint dependence. Moreover, copulas can be used to model joint extreme market realizations, where two assets, or more, jointly perform extremely well or extremely poorly. This is due to copula capacity of capturing assets interdependencies that are not encompassed by simple linear correlation. In the first part of this work we describe what a copula is and how it can be modelled, taking into account that different types of copulas exist, with their particular shape, behaviour and tail characteristics. These differences would allow us to fit empirical data to optimal copula, meaning the copula that best reflects data behaviour, especially behaviour in the tails. According to this, we are to take a set of empirical data: price time series of four financial traded indices: FTSE MIB, CAC All-Tradable, CDAX and IBEX35. From Eikon Reuters-Datastream, we download 20-years weekly price time series. We want to fit these data to various copulas and estimate copula parameters by Inference for Margins method. In the last chapter, we will avail of copula method in order to deduce Value at Risk for an imaginary portfolio, composed of our four financial indices. In the end, we will compare Values at Risk obtained by portfolios with the same weights, but applying different copulas. Theoretically, Values at Risk will show differences according to copula behaviour and tail dependence. All the analysis will be conducted with the statistic software R.

INDEX

INTRODUCTION

According to J.F. Jouanin, G. Riboulet and T. Roncalli (2011), a copula is both a powerful and simple tool to describe dependent risk. By the term "copula" we mean the joint distribution of a vector of uniform random variables. As an example, calculating the variance in the returns of a risky assets portfolio entails computing both individual assets variances and the type of dependence that exists among them. The latter element is captured by the copula: it allows to analyse the joint dependence separately from the single distributions.

In particular, S.T. Rachev, M. Stein and W. Sun (2015) suggest copulas can be useful in modelling extreme market events, like joint tail realizations, due to assets interdependencies that cannot be captured by simply using linear correlation.

A good example of useful copula application is provided by K. Aas (2004): he considers a portfolio composed of a stock market index and of an exchange rate. For what concerns single assets distributions, he has found that the Student t-distribution could provide a reasonable fit both to the univariate distribution of daily stock market index, and to exchange rate return. In this way, the obvious solution would be to model the joint distribution by a bivariate Student t-distribution. However, a standard bivariate Student t-distribution would force both assets distributions to have the same tail heaviness, while in reality it is not like this. On the other hand, decomposing the multivariate distribution between assets distributions on one side, and copula, on the other, would allow for the fitting of better models for each individual variable.

T. Schmidt (2006) offers an even easier explanation of the copula instrument: he proposes to consider two real-valued random variables, X_1 and X_2, where each could be the outcome of a simple experiment, like throwing a dice, or a more complex one. T. Schmidt says that we have to enter a bet on X_2, based on X_1, that is already known. A possible copula based on X_1 and X_2 would encompass the quantity of information deducible for X_2 by knowing X_1: the interrelation or dependence of these two random variables. Each random variable is fully described by its cumulative distribution function (cdf) $F_i(x) \coloneqq P(X_i \leq x)$, the so called marginal. In the case of throwing the dice twice we would have $F_1 = F_2 \eqqcolon F$. Here, we have an extreme case, as the two variables are independent, and the cumulative distribution functions give no information about the joint behaviour: in fact, the joint distribution function is simply the product of the marginal distributions:

$$P(X_1 \leq x_1, X_2 \leq x_2) = F(x_1) \cdot F(x_2)$$

However, the example is important as it shows the two ingredients to obtain a full description of X_1 and X_2 considered together: the marginal behaviours and the type of interrelation, in this case independence. Thanks to copulas, this kind of separation between margins and dependence can also be realized in a more general framework.

1. THE COPULA AS A DEPENDENCE FUNCTION

T. Schmidt (2006) settles the first goal into transforming random variables X_i into uniformly distributed random variables U_i. In this way, every random variable X with cumulative distribution function F can always be represented as $X = F^{\leftarrow}(U)$, where $F^{\leftarrow}$ denotes the generalized inverse of F. After having transformed marginal random variables distributions into uniform ones, we adopt the latter as the reference case. The copula is so expressed according to the reference case. Retrieving the independence case that we exposed above, the joint distribution function can be restated, by two standard uniform random variables U_1 and U_2, as:

$$P(F_1^{\leftarrow}(U_1) \leq x_1, F_2^{\leftarrow}(U_2) \leq x_2) = P(U_1 \leq F_1(x_1), U_2 \leq F_2(x_2))$$

$$C(u_1, u_2) = P\{U_1 \leq u_1, U_2 \leq u_2\}$$

P. Embrechts (2009) develops last formula step by step, like:

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$$
$$= P(F_1(X_1) \leq F_1(x_1), F_2(X_2) \leq F_2(x_2))$$
$$= P(U_1 \leq F_1(x_1), U_2 \leq F_2(x_2))$$
$$= C(F_1(x_1), F_2(x_2))$$

Where F is the joint distribution function and F_1 and F_2 are marginal distribution functions. The C above is the copula: the distribution function of the random vector (U_1, U_2), with standard uniform marginal distributions on $[0,1]^2$. The formula couples the continuous marginal distribution functions F_1, F_2 to the joint distribution function F via the copula C.

U. Cherubini, E. Luciano and W. Vecchiato (2004) state that, defining what a copula is, two conditions are substantially needed: groundedness and the 2-increasing property. If fulfilled, these two conditions allow copulas to respect properties of distribution functions.

Referring to a bivariate function $C: [0,1]^2 \longrightarrow [0,1]$, J.F. Jouanin, G. Riboulet and T. Roncalli (2011) explain that C is 2-increasing if, for $0 \leq u_1 \leq u_2 \leq 1$ and $0 \leq v_1 \leq v_2 \leq 1$, we have:

$$C([u_1, v_1]x[u_2, v_2]) \equiv C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$$

In order to be a copula function, the same function $C: [0,1]^2 \longrightarrow [0,1]$ must even be grounded: $C(1, u) = C(u, 1) = u$ and $C(0, u) = C(u, 0) = 0$ for all $u \in [0,1]$

By the notation $C(\boldsymbol{u}) = C(u_1, \dots, u_d)$ we can expand these considerations to the d-dimensional copula $C: [0,1]^d : \longrightarrow [0,1]$. In fact, according to U. Cherubini, E. Luciano and W. Vecchiato (2004), in the d-dimensional case, where d>2, notions of groundedness and n-increasing property are straightforward extensions of the two dimensional case.

Let the function $G: \mathfrak{R}^{*n} \longrightarrow \mathfrak{R}$ have a domain $Dom\ G = A_1 x A_2 x \dots x A_d$, where the non-empty sets A_i have a least element a_i, equal to zero. The function G is said to be grounded if and only if it is null for every $\boldsymbol{u} \in Dom\ G$, with at least one index k such that $v_k = a_k$:

$$G(\boldsymbol{u}) = G(u_1, u_2, \dots, u_{k-1}, a_k, u_{k+1}, \dots, u_d) = 0$$

The marginal in component "i" is obtained by setting $u_j = 1$ for all $j \neq i$ and, as it must be uniformly distributed:

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \forall u_i \in [0,1]$$

Similarly, $\forall u_i \in [0,1], C(u_1, \dots, u_d) = 0$ if at least one of the u_i equals zero.

For what concerns d-increasing property, for $a_i \leq b_i$ the probability $P(U_1 \in [a_1, b_1], \dots, U_d \in [a_d, b_d])$ must be non-negative: both T. Schmidt (2006) and Y. Malevergne and D. Sornette (2001) have efficiently defined it with the so-called rectangle inequality:

$$\sum_{i_1=1}^{2} \dots \sum_{i_d=1}^{2} (-1)^{i_1 + \dots + i_d} C(u_{1,i_1}, \dots, u_{d,i_d}) \geq 0$$

Where $u_{j,1} = a_j\ and\ u_{j,2} = b_j$

Every function which satisfies these properties is a copula.

2. DEFINITIONS

As a copula is substantially a dependence function, connecting random variables distribution functions, this relationship is condensed in Sklar's theorem, as reported by U. Cherubini, E. Luciano and W. Vecchiato (2004). The theorem states not only that copulas are joint distribution functions, but even that joint distribution functions can always be written in terms of uniform marginal distributions and a unique copula to entangle them. Therefore, every time we have to cope with joint distribution functions, we can easily avail ourselves of a copula. According to Sklar's theorem, we have to consider a probability space (Ω, $\mathcal{F}$, P), with Ω a non-empty set, $\mathcal{F}$ a sigma-algebra on Ω and P a probability measure on $\mathcal{F}$. Let X_1 and X_2 be two Borel-measurable random variables on (Ω, $\mathcal{F}$, P) with values in $\mathfrak{R}^*$, the extended real line. Let also F be a two-dimensional joint distribution function whose marginal distributions are F_1 and F_2. Then, F admits a copula representation:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$

The copula C is unique if its marginal distributions are continuous. Random variables are said to be continuous when their distribution functions are. We therefore have a canonical representation of the distribution: on the one hand, the marginal distributions F_1 and F_2, that is to say the one-dimensional directions; on the other hand, the copula, that links them. As such, the copula defines the dependence between the one-dimensional directions.

P. Embrechts (2009) extends Sklar's theorem from 2 to d>2 dimensions. We have just to suppose $X_1, \dots, X_d$ to be random variables with continuous distribution functions $F_1, \dots, F_d$ and one joint distribution function F. Then, there exists a unique copula C, on $[0,1]^d$, such that for all $\boldsymbol{x} = (x_1, \dots, x_d)^T \in \mathbb{R}^d$:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

From any multivariate distribution, F, we can extract the marginal distributions, F_i, and the copula, C. It is important to underline that marginal distributions do not need to be in any way similar to each other, nor the choice of copula is constrained by the choice of marginal distributions. This flexibility makes copula a potentially useful tool for building econometric models to analyse financial data.

As was stated above, marginal random variables X_i can be transformed into uniformly distributed random variables U_i. Knowing that $X = F^{\leftarrow}(U)$, P. Embrechts (2009) elaborates Sklar's theorem for $\boldsymbol{u} = (u_1, \dots, u_d)^T \in [0,1]^d$:

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

Where the F_i^{-1} are the quantile functions of the marginal distributions.

Always P. Embrechts (2009) reminds us that, if a joint bivariate distribution is continuous, it should even admit to a density like this:

$$f(x_1, x_2) = c\big(F_1(x_1), F_2(x_2)\big) \cdot f_1(x_1) \cdot f_2(x_2)$$

Where $c(u_1, u_2)$ is the density of the copula C.

To clarify the equation, C. Kharoubi- Rakotomalala proposes a practical example: imagine a portfolio composed of two risk factors: IBM (x_1) and Google (x_2) stocks. In this case:

- $f(x_1, x_2)$ represents the joint density of the portfolio: it encompasses the simultaneous behaviour of the two type of stocks.
- $\left.\begin{matrix} f_1(x_1) \\ f_2(x_2) \end{matrix}\right\}$ are the marginal densities
- $c\big(F_1(x_1), F_2(x_2)\big)$ stays for the copula density.

Calculating copula density in d-dimensions, A. Patton (2007) states that, if we have d>2 marginal distributions, and if the joint distribution function is d-times differentiable, then taking the d^{th} cross-partial derivative of equation

$$F(\boldsymbol{x}) = C\big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big)$$

$$\forall x \in R^d$$

We obtain: $f(\boldsymbol{x}) \equiv \frac{\delta^d}{\delta x_1, \delta x_2, \dots, \delta x_d} F(\boldsymbol{x})$

$$= \prod_{i=1}^{d} f_i(x_i) \cdot \frac{\delta^d}{\delta u_1, \delta u_2, \dots, \delta u_d} C\big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big)$$

$$\equiv \prod_{i=1}^{d} f_i(x_i) \cdot c\,\big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big)$$

Where

$$c(\boldsymbol{u}) \coloneqq \frac{\partial^{\mathrm{d}} C(u_1, \dots, u_d)}{\partial_{u_1} \dots \partial_{u_d}}$$

This last equation would be further clarified if we took Schmidt (2006) version:

$$c(\boldsymbol{u}) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1\big(F_1^{-1}(u_1)\big) \dots f_d(F_d^{-1}(u_d))}$$

Denoting the joint density by f and the marginal densities by $f_i, i = 1, \dots, d$, joint density is equal to the product of marginal densities and copula density, denoted with c.

Following T. Schmidt's (2006) treatise on copulas, we should now consider Hoeffding and Fréchet derivation that a copula always lies in between certain bounds. The reason is given by the existence of some extreme cases of dependence.

To make it more understandable, T. Schmidt (2006) proposes to start considering two uniform random variables, called U_1 and U_2. In the case $U_1 = U_2$, these two variables show extreme positive dependence on each other. In this case, the copula is given by:

$$C(u_1, u_2) = P(U_1 \leq u_1, u_2 \leq u_2) = \min(u_1, u_2)$$

This copula is always attained every time X_2 is a monotonic transformation of X_1. As a consequence, the two random variables are defined co-monotonic.

A strongly different case would be given by independence between the two random variables. In case of independence, the copula is equal to $C(u_1, u_2) = u_1 \cdot u_2$, that is just the case of the two dices thrown in the introduction.

However, independence is simply an intermediate step before the extreme that is opposite to co-monotonicity: counter-monotonicity. With uniform random variables, this case is due to $U_2 = 1 - U_1$. The related copula is:

$$C(u_1, u_2) = P(U_1 \leq u_1, 1 - U_1 \leq u_2)$$

$$= P(U_1 \leq u_1, 1 - u_2 \leq U_1) = u_1 + u_2 - 1$$

And zero otherwise.

To put everything together, U. Cherubini, E. Luciano and W. Vecchiato (2004) state that copulas are bounded by these extreme cases of dependence and have to satisfy the following inequality:

$$max(u_1 + u_2 - 1,0) \leq C(u_1, u_2) \leq \min(u_1, u_2)$$

For every point $(u_1, u_2) \in A \times B$, where A and B are non-empty subsets of I=[0,1], containing both 0 and 1; while the two-dimensional copula C is a real function defined on $A \times B$.

T. Schmidt (2006) advises that it would be even possible to draw these considerations to the multidimensional case, where dimensions are d>2. However, whereas a co-monotonic copula always exists in every d-dimension, there can be a problem with the counter-monotonic Hoeffding-Fréchet bound, if we consider more than two dimensions. To clarify this, consider three random variables: X_1, X_2, X_3. We are free to settle counter-monotonicity between X_1 and X_2 as well as between X_1 and X_3. Although, we get some restrictions when we have to carve out the relation between X_2 and X_3. In fact, if X_1 decreases, X_2 should increase, as it is counter-monotonic with respect to X_1. Even X_3 should increase, as X_3 is as well counter-monotonic with respect to X_1. As a consequence, X_2 cannot be counter-monotonic with respect to X_3, nor obviously vice versa. This ends to say that a perfect counter-monotonic copula cannot logically exists in more than two dimensions. Fortunately, the bond still holds, and this is all we have to care about.

To make Hoeffding-Fréchet bounds more understandable, U. Cherubini, E. Luciano and W. Vecchiato (2004) offer a nice graphical representation. Every copula has to lie inside of the pyramid shown in figures 1 and 2.

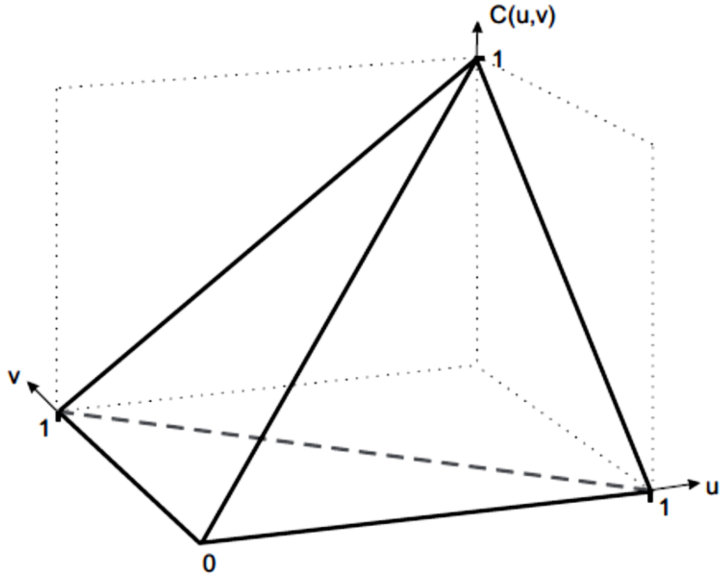


Figure 1

In fact, the graph of each copula can be defined as a continuous surface over the unit square that contains the skew quadrilateral whose vertices are (0,0,0), (1,0,0), (1,1,1) and (0,1,0). When $(u_1, u_2) \in I^2$, so that C becomes a copula, the bounds are copulas too.

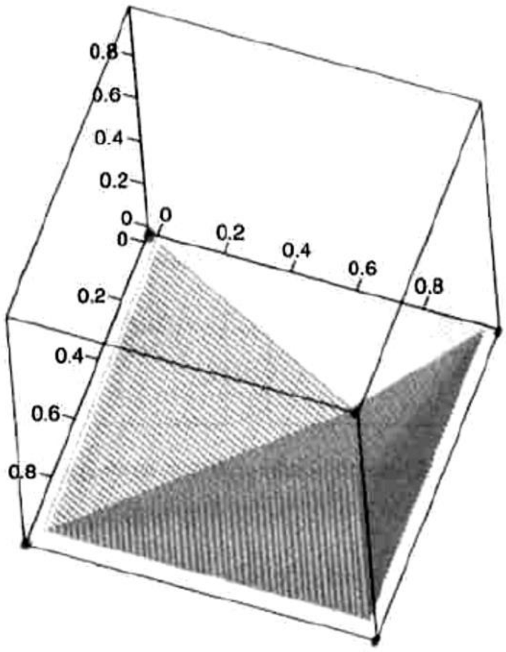


Figure 2

The surface given by the bottom and the back side of the pyramid represents the lower Hoeffding-Fréchet bound. The lower bound is denoted by C^{-}, and is called minimum copula: it is the counter-monotonicity copula $C(u_1, u_2) = max\{u_1 + u_2 - 1{,}0\}$. The upper bound is denoted by C^{+}, and called maximum copula: $C(u_1, u_2) = \min(u_1, u_2)$. Both minimum and maximum bounds are represented in the third figure, respectively at left and at right.

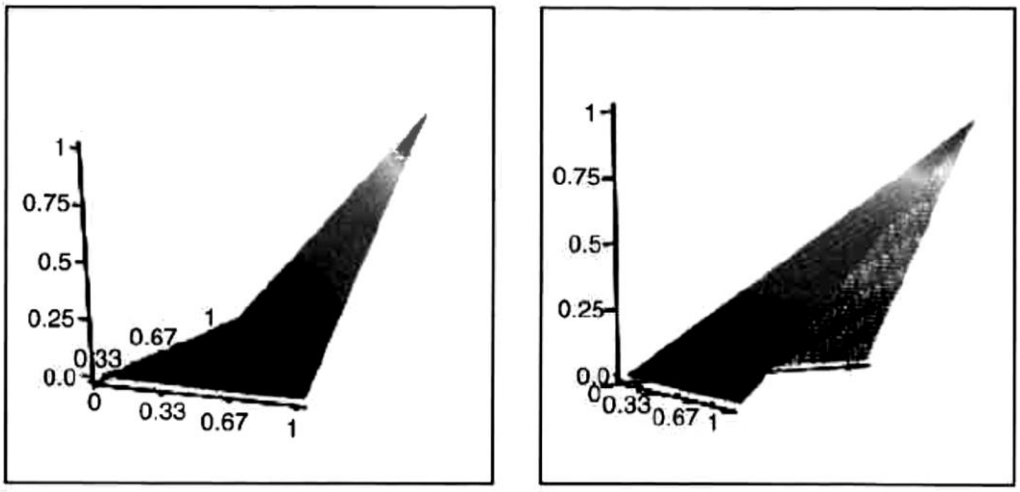


Figure 3

In order to resume what we said before, copulas have to satisfy the following inequality:

$$\max(u_1 + u_2 - 1{,}0) \leq c(u_1, u_2) \leq \min(u_1, u_2)$$

for every point $(u_1, u_2) \in$ A×B.

This theorem has consequences on the so-called level curves of the copula $\mathrm{C}(u_1, u_2)$: the set of points of I^2 such that $\mathrm{C}(u_1, u_2) = \mathcal{K}$, with $\mathcal{K}$ constant:

$$\{(u_1, u_2) \in I^2 \colon \mathrm{C}(u_1, u_2) = \mathcal{K}\}$$

The level curves of the minimum and maximum copula are respectively:

$$\{(u_1, u_2): \max(u_1 + u_2 - 1{,}0) = \mathcal{K}\}, \mathcal{K} \in I$$

$$\{(u_1, u_2): \min(u_1, u_2) = \mathcal{K}\}, \mathcal{K} \in I$$

They are represented in figure 4, by the courtesy of U. Cherubini, E. Luciano and W. Vecchiato (2004). In the plan (u_1, u_2) level curves of the minimum copula are characterized as segments parallel to the line u_2=u_1 . Level curves of the maximum copula are drawn as kinked lines instead.

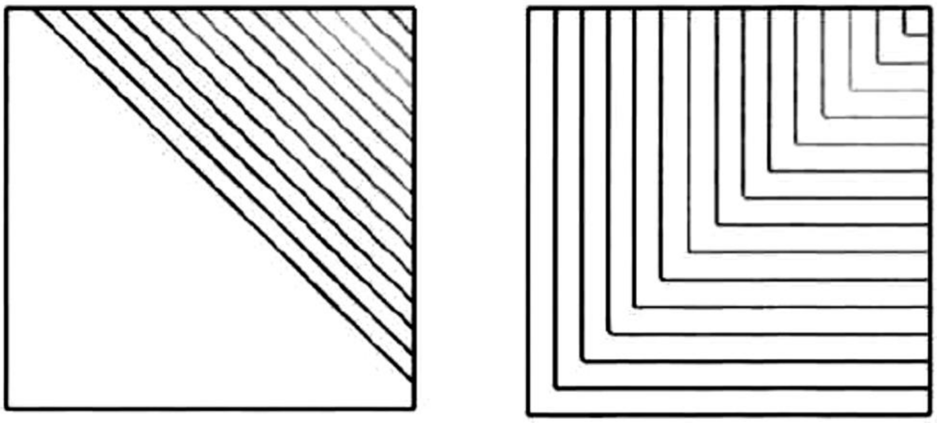


Figure 4

As $\mathcal{K}$ increases, the triangle is shifted upwards. The existence of the lower and upper bounds gives the possibility of defining a concordance order between copulas. In fact, we can say that the copula C_1 is smaller than the copula C_2 – written as $C_1 \prec C_2$ –if and only if

$$C_1(u_1, u_2) \leq C_2(u_1, u_2)$$

For every $(u_1, u_2) \in I^2$.

Naming Fréchet-Hoeffding lower and upper bound, respectively, C^- and C^+, U. Cherubini, E. Luciano and W. Vecchiato (2004) say that it is possible to avail oneself of Sklar's theorem in order to rewrite the inequality $C^- \leq C \leq C^+$ as:

$$\max(F_1(x_1) + F_2(x_2) - 1{,}0) \leq F(x_1, x_2) \leq \min(F_1(x_1), F_2(x_2))$$

Where the first member of the inequality is minimum copula, and the last is maximum one.

T. Schmidt (2006) provides the formulation even in d-dimensions, where d>2, for Fréchet-Hoeffding bounds: consider a copula $C(\boldsymbol{u}) = C(u_1, \dots, u_d)$. Then

$$max(u_1 + u_2 + \cdots + u_d - 1{,}0) \leq c(\boldsymbol{u}) \leq min(u_1, u_2, \dots, u_d) \quad \forall \boldsymbol{u} \in I^d$$

U. Cherubini, E. Luciano and W. Vecchiato (2004) remind that, in d-dimensions, the upper bound still satisfies the definition of copula, and is denoted by C^+ (the maximum copula). However, the lower bound never satisfies the definition of copula for d>2. Nonetheless, the bound is still the best possible: pointwise there always exists a copula that takes its value.

In order to further clarify co-monotonicity and counter-monotonicity, P. Embrechts (2009) translates Fréchet-Hoeffding bounds into the language of correlations. He states that, for any bivariate model F with F_1, F_2 as marginal distribution functions, the corresponding linear correlation coefficient ρ_F satisfies:

$$-1 \leq \rho_{min} \leq \rho_F \leq \rho_{max} \leq +1$$

Where all values in the closed interval $[\rho_{min}, \rho_{max}]$ can be achieved. One always has that $\rho_{min} \leq 0$ and $\rho_{max} > 0$ but it is possible that $\rho_{min} > -1$ and/or $\rho_{max} < +1$. ρ_{min} corresponds to counter-monotonicity, while ρ_{max} stays for co-monotonicity.

As co-monotonicity refers to perfect positive dependence and counter-monotonicity to perfect negative dependence respectively, the intermediate step of independence is to be settled between the extremes. The independence copula is:

$$\Pi(\boldsymbol{u}) = \prod_{\mathrm{i}=1}^{\mathrm{d}} \mathrm{u_i}$$

As T. Schmidt (2006) underlines, random variables are said to be independent if and only if their copula is the independence copula. The related copula density is simply constant.

Families of copulas which encompass product, minimum and even maximum copulas are called comprehensive.

It is worth noting that minimum and maximum copulas do not have a density as they both are represented by a kinked line and therefore cannot be differentiable. In the co-monotonic case, the distribution has mass only on the diagonal $u_2 = u_1$, while in the countermonotonic case there is mass only on $\{u_1 = 1 - u_2\}$. Because of this these two copulas cannot be described by a density.

The last property of copulas that is worth saying is that strictly increasing transformations do not change the dependence structure. On first sight, this seems to be counterintuitive: monotone transformations do change the dependence. Although, after removing the effects of the monotone transformation on the marginal distributions, we end up with the same dependence structure in the copula.

2.1 Survival copula and joint survival function

U. Cherubini, E. Luciano and W. Vecchiato (2004) propose us to consider the probability: $\bar{F}(\boldsymbol{x}) = Pr(X_1 > x_1, X_2 > x_2, \dots, X_d > x_d)$. It is defined as joint survival probability or joint survival function of the d random variables X_i, while the marginal survival probabilities or marginal survival functions are: $\bar{F}_i(x_i) = Pr(X_i > x_i)$. Since the probability $\bar{F}(x_1, x_2) = \Pr(X_1 > x_1, X_2 > x_2)$ represents the joint survival probability or joint survival function of X_1

and X_2, respectively beyond x_1 and x_2, the copula which represents it in terms of the marginal survival probabilities or survival distribution functions of the two agents or components separately, $\bar{F}_1(x_1)$ and $\bar{F}_2(x_2)$, is named survival copula.

Since $\overline{C}$ is a copula, it stays within the Fréchet bounds:

$$C^- \prec \bar{C} \prec C^+$$

In addition, it can be easily verified that in the minimum, product and maximum case, copulas and survival copulas coincide:

$$\bar{C}^- = C^-;\ \bar{C}^\perp = C^\perp;\ \bar{C}^+ = C^+$$

The copula that represents the joint survival probability in terms of the marginal survival probabilities of the d-components X_i is the survival copula. As we have seen above for copulas, uniqueness tout court holds true if every marginal survival probability is continuous.

3. MEASURES OF ASSOCIATION

From U. Cherubini, E. Luciano and W. Vecchiato (2004), association concepts, loosely speaking, aim at capturing whether the probability of having large or small values of both X_1 and X_2 is higher than the probability of having large values of X_1 together with small values of X_2, or vice versa. If we imagine it geometrically, it looks like the probability mass associated with the upper and lower quadrants, as opposite to the one associated with the rest of the Cartesian plane (x,y).

As T. Schmidt (2006) suggests, measures of association are of common usage when we need to summarize a complicated dependence structure. Substantially, they are three. The most classic one is linear correlation. However, this is suitable just to the class of elliptical distributions. In terms of copulas, this means that it is good only if we have a Gaussian or a t-Student copula, so not in the majority of cases. Outside the class of elliptical distributions, linear correlation causes fallacies. The second association measure is rank correlation, while the third is tail dependence. This last one is particularly useful in detecting dependence in the extremes. When it comes to rank correlation, instead, the most appropriate measures turn out to be Kendall’s tau and Spearman’s rho.

In order to compute every measure of association in a copula, it is a prerequisite that all marginal distributions involved are continuous.

3.1 Linear correlation

As T. Schmidt (2006) has said, linear correlation is a dependence measure applicable only in the case of elliptical distributions. An elliptical distribution can be obtained by an affine transformation like: $X = \mu + AY$, with $\mu \in \mathbb{R}^d, A \in \mathbb{R}^{dxd}$.

Taking two continuous random variables X_1 and X_2, the linear correlation coefficient $\rho_{X_1X_2}$ is:

$$\rho_{X_1X_2} = \frac{cov(X_1, X_2)}{\sqrt{var(X_1)var(X_2)}}$$

$\rho_{X_1X_2}$ is invariant under linear increasing transformations, but not under non-linear increasing transformations, like logarithmic transformations.

$\rho_{X_1X_2}$ is bounded as $\rho_l \leq \rho_{X_1X_2} \leq \rho_u$, where the bounds ρ_l and ρ_u are attained respectively when X_1 and X_2 are counter-monotonic and co-monotonic, so, when there is, respectively, perfect negative and positive dependence.

However, both T. Schmidt (2006) and U. Cherubini, E. Luciano and W. Vecchiato (2004) underline certain pitfalls that occur when linear correlation coefficient is used, outside the class of elliptical distributions, and that can seriously undermine the validity of the analysis. The first pitfall is that a linear correlation of 0 would mean independence for a normal distribution. Although, even for a Student t-distribution this is no longer true. The second pitfall is that linear correlation coefficient remains invariant under linear transformations, but not under general transformations: two log-normal random variables have a different linear correlation than the underlying normal random variables. The third problem is that it is not possible to elaborate a joint distribution for any couple of marginal distributions, given the correlation coefficient ρ. It is always feasible in the class of elliptical distributions, but not in general. As an example, in the case of log-normal marginal distributions, the interval of attainable linear correlation becomes smaller with increasing volatility. To illustrate this, consider two normal random variables X_1 and X_2, both with zero mean and variance $\sigma^2 > 0$. The linear correlation of the two log-normal random variables $Y_i = \exp(X_i)\, , i = 1{,}2$ equals

$$Corr(Y_1, Y_2) = \frac{e^{\rho\sigma^2} - 1}{\sqrt{\left(e^{\sigma^2} - 1\right)\left(e^{\sigma^2} - 1\right)}}$$

To make the example even more catchable, T. Schmidt (2006) provides a nice graphical representation in figure 5. The picture shows $\rho = corr(Y_1, Y_2)$ where $Y_i = \exp(X_i)$ with $X_i \sim \mathcal{N}(0, \sigma^2)$. Note that the smallest attained correlation is increasing with σ, so for σ=1 we have that $Corr(Y_1, Y_2) \geq -0.368$ and for σ=2 even $Corr(Y_1, Y_2) \geq -0.018$.

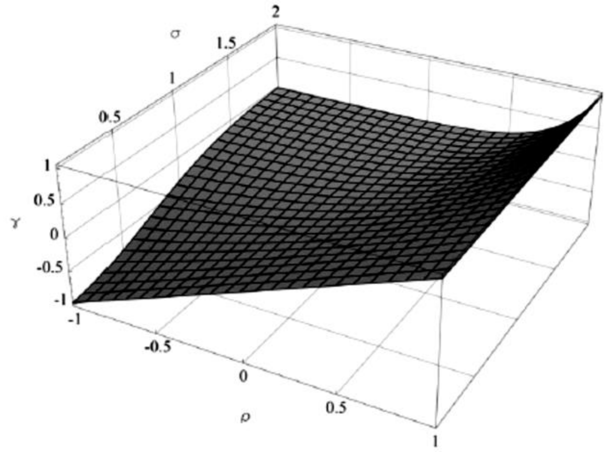


Figure 5

It is in general wrong to deduce a small degree of dependence from a small linear correlation as even perfectly related random variables can have zero linear correlation: consider $X_1 \sim \mathcal{N}(0,1)$ and $X_2 = X_1^2$. Then:

$$Cov(X_1, X_2) = E\left(X_1 \cdot (X_1^2 - 1)\right) = E(X_1^3) - E(X_1) = 0$$

Having covariance 0 implies of course zero linear correlation, while on the other side the observation of X_1 immediately yields full knowledge of X_2.

S.T. Rachev, M. Stein and W. Sun (2015) even notice an additional reason for which linear correlation wouldn't be a satisfactory measure of dependence. If we take as random variable the rate of return of a security, an index or a stock, linear correlation cannot keep track of higher variance in the returns, that is, when extreme events are observed more frequently than normal. Moreover, linear correlation coefficient only measures the degree of dependence, but does not clearly discover the structure of dependence.

3.2.Rank correlation

The most important rank correlation estimators are Kendall's tau and Spearman's rho. S.T. Rachev, M. Stein and W. Sun (2015) explain that the logic is to concentrate on the ranks of given data rather than on the data itself. Considering the ranks leads to scale invariant estimates, that is very pleasing when we have to work with copulas, as rank correlation measures allow to fit copulas to data.

In order to elaborate Kendall's tau, T. Schmidt (2006) suggests that we have to consider two random variables X_1 and X_2. For a comparison we take two additional random variables $\widetilde{X_1}$ and $\widetilde{X_2}$ into account, both being independent of X_1 and X_2, but with the same joint distribution. Now we plot a point in a graph from each couple of random variables, namely (x_1, x_2) and $(\widetilde{x_1}, \widetilde{x_2})$, and we connect them by a line. If we have positive dependence, we would expect that the line is increasing, and, otherwise, if there is negative dependence, the line is to be decreasing. Similarly, considering $\left(X_1 - \widetilde{X_1}\right) \cdot \left(X_2 - \widetilde{X_2}\right)$, a positive sign is indicative of the increasing case, while a negative sign would turn up into the decreasing case. As we have to express it in terms of expected value, we define Kendall's tau by:

$$\rho_\tau(X_1, X_2) \coloneqq \boldsymbol{E}\left[sign\left(\left(X_1 - \widetilde{X_1}\right) \cdot \left(X_2 - \widetilde{X_2}\right)\right)\right]$$

For a d-dimensional vector of random variables **X** and an independent copy $\widetilde{\boldsymbol{X}}$, but with the same joint distribution, we define Kendall's Tau by

$$\rho_\tau(\boldsymbol{X}) \coloneqq Cov\left[sign\left(\boldsymbol{X} - \widetilde{\boldsymbol{X}}\right)\right]$$

Alternatively, T. Schmidt (2006) writes the formula as:

$$\rho_\tau(X_1, X_2) = P\left(\left(X_1 - \widetilde{X_1}\right) \cdot \left(X_2 - \widetilde{X_2}\right) > 0\right) - P\left(\left(X_1 - \widetilde{X_1}\right) \cdot \left(X_2 - \widetilde{X_2}\right) < 0\right)$$

In the case both probabilities are the same, this means that upward slopes are to be expected with the same probability as downward slopes, and $\rho_\tau = 0$. Otherwise, if Kendall's tau is positive, there is a higher probability of upward slopes to occur. Similarly, if Kendall's tau is negative, we would expect rather downward sloping outcomes. As Kendall's tau is a measure with possible values in the interval [-1,1], when it takes a value of 0, this means that variables are independent. When it takes a value of 1, variables are co-monotonic: perfect positive dependence; while it is equal to -1 in case of perfect negative dependence: variables are counter-monotonic.

It is interesting to note that Kendall's tau of a copula and of its associated survival copula coincide: $\tau_C = \tau_{\bar{C}}$.

Now we can adapt Kendall's tau to the scope of our discussion, by fitting a copula to it: according to K. Aas (2004), Kendall's tau of two variables X_1 and X_2, jointly distributed, is:

$$\rho_\tau(X_1, X_2) = 4\int_0^1 \int_0^1 C(u_1, u_2)\, dC(u_1, u_2) - 1$$

Where C(u_1, u_2) is the copula of the bivariate distribution function of X_1 and X_2.

The double integral right above is the expected value of $C(u_1, u_2)$ where both u_1 and u_2 are standard uniforms and have joint distribution C: $\tau = 4E[C(u_1, u_2)] - 1$.

It follows that $-1 \leq 4E[C(u_1, u_2)] - 1 \leq 1$.

As K. Aas (2004) remembers, for elliptical copulas, like Gaussian and Student t-copulas, Kendall's tau can be included in the formulation of linear correlation coefficient:

$$cor(X_1, X_2) = sin\left(\frac{\pi}{2}\rho_\tau\right)$$

Where "cor" stays for the linear correlation coefficient.

When it comes to Archimedean copulas, B. Schweizer and E. Wolff (1981) establish that Kendall's tau could be related to the dependence parameter, that we will explain further. For the Clayton copula Kendall's tau is given by $\rho_\tau(X_1, X_2) = \frac{\alpha}{\alpha+2}$.

And for the Gumbel copula it is $\rho_\tau(X_1, X_2) = 1 - \frac{1}{\alpha}$

As written by K. Aas (2004), Spearman's rho of two variables X_1 and X_2 with copula C is given by:

$$\rho_s(X_1, X_2) = 12 \int_0^1 \int_0^1 C(u_1, u_2)\, du_1 du_2 - 3$$

$$= 12 \int_0^1 \int_0^1 u_1 u_2\, dC(u_1, u_2) - 3$$

Where C(u_1, u_2) is the copula of the bivariate distribution function of X_1 and X_2. Let X_1 and X_2 have distribution functions F_1 and F_2, respectively, Then, we have the following relationship between Spearman's rho and the linear correlation coefficient:

$$\rho_s(X_1, X_2) = cor(F(X_1), F(X_2))$$

T. Schmidt (2006) defines Spearman's rho by:

$$\rho_s \coloneqq Corr(F_1(X_1), F_2(X_2)) = \frac{cov\big(F_1(X_1), F_2(X_2)\big)}{\sqrt{var(F_1(X_1))var(F_2(X_2))}}$$

Even in this case, we note that Spearman's rho of a copula and of its associated survival copula coincide: $\rho_{SC} = \rho_{S\bar{C}}$

Also for Spearman's rho one could demonstrate that it reaches its minimum and maximum bounds if and only if X_1 and X_2 are respectively counter-monotonic and co-monotonic continuous random variables:

$$\rho_s = -1 \quad iff\ C = C^-$$
$$\rho_s = 1 \quad iff\ C = C^+$$

K. Aas (2004) even manages to demonstrate that, for the Gaussian and Student t-copulas, linear correlation coefficient and Spearman's rho are connected, in this way:

$$cor(X_1, X_2) = 2sin\left(\frac{\pi}{6}\rho_s\right)$$

Both $\rho_\tau(X_1, X_2)$ and $\rho_s(X_1, X_2)$ may be considered as measures of the degree of monotonic dependence between X_1 and X_2, whereas linear correlation measures the degree of linear dependence only. Moreover, these measures are invariant under monotone transformations, while the linear correlation generally isn't. Hence, according to P. Embrechts, A.J. McNeil and D. Straumann (1999) it is slightly better to use these measures than the linear correlation coefficient.

3.3.Tail dependence

The primary motivation for the use of copulas in finance comes from the growing empirical evidence that the dependence between many important assets returns is non-normal. K. Aas (2004) offers an evident example of this: in time of stress, correlation between assets returns tends to increase. One prominent example of non-normal dependence is where two assets returns exhibit greater correlation during market downturns than during market upturns. Bivariate tail dependence measures the amount of dependence in the upper and lower quadrant of a bivariate distribution. This is of great interest for the risk manager trying to guard against concurrent bad events.

Following U. Cherubini, E. Luciano and W. Vecchiato (2004), bivariate tail dependence refers to concordance in the tail: where extreme values of random variables X_1 and X_2 distributions are verified. These measures are independent of the univariate distributions of assets returns. Moreover, they are invariant under strictly increasing transformations of X_1 and X_2.

To better understand tail dependence, T. Schmidt (2006) proposes this example: consider two uniform random variables U_1 and U_2 with copula C. Upper tail dependence means, intuitively, that with large values of U_1 also large values of U_2 are to be expected. More precisely, the probability that U_2 exceeds a given threshold q, given that U_1 has already exceeded the same value q for $q \longrightarrow 1$, is considered. If this latter probability is smaller than of order q, then the random variables have no tail dependence, like for example in the independent case. Otherwise they have tail dependence. For our random variables X_1 and X_2 with distribution functions $F_i, i = 1,2$ we define the coefficient of upper tail dependence by:

$$\lambda_u \coloneqq \lim_{q \longrightarrow 1} P(X_2 > F_2^{\leftarrow}(q) | X_1 > F_1^{\leftarrow}(q))$$

The coefficient of lower tail dependence is defined analogously by:

$$\lambda_l \coloneqq \lim_{q \longrightarrow 0} P(X_2 \leq F_2^{\leftarrow}(q) | X_1 \leq F_1^{\leftarrow}(q))$$

U. Cherubini, E. Luciano and W. Vecchiato (2004) resume that copula C has upper tail dependence if and only if $\lambda_u \in (0,1]$, and no upper tail dependence if and only if $\lambda_u = 0$. If the coefficient of upper tail dependence is higher than 0, this means that large events tend to occur simultaneously. C is, otherwise, said to have lower tail dependence in the case $\lambda_L \in (0,1]$, and no lower tail dependence if $\lambda_L = 0$.

When it comes to elliptical distributions, like Gaussian copula and Student t-copula, it is important to remember that lower tail dependence is identical to upper tail dependence. As T. Schmidt (2006) formulates: $\lambda_u(X_1, X_2) = \lambda_l(X_1, X_2)$.

For the Gaussian copula, the coefficients of lower tail and upper tail dependence are

$$\lambda_l(X_1, X_2) = \lambda_u(X_1, X_2) = 2 \lim_{x \longrightarrow -\infty} \Phi\left(x \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right) = 0$$

where ρ is linear correlation coefficient and Φ denotes the standard Gaussian distribution function. Regardless of high correlation ρ we choose, extreme events appear to occur independently in X_1 and X_2, unless ρ=1.

For the Student t-copula, the coefficients of lower and upper tail dependence are

$$\lambda_l(X_1, X_2) = \lambda_u(X_2, X_2) = 2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$$

Where $t_{\nu+1}$ denotes the distribution function of a univariate Student t-distribution with $\nu + 1$ degrees of freedom. The stronger the linear correlation ρ and the fewer the degrees of freedom ν become, the stronger is the tail dependence. Surprisingly, perhaps, the Student t-copula gives asymptotic dependence in the tail, even when ρ is negative (> -1), or zero.

Just in order to resume, in elliptical copulas the coefficient of lower and higher tail dependence is identical, due to the radial symmetric shape of elliptical copulas. A Gaussian copula has both lower and higher tail dependence coefficients equal to 0. This is stemming from the fact that a multivariate Gaussian distribution is the n-dimensional version of a Gaussian distribution, which assigns too low probabilities to extreme outcomes.

Now T. Schmidt (2006) considers Clayton copula. The coefficient of lower tail dependence equals:

$$\frac{(2q^{-\alpha}-1)^{-\frac{1}{\alpha}}}{q} = (2-q^{\alpha})^{-\frac{1}{\alpha}} \longrightarrow 2^{-\frac{1}{\alpha}} = \lambda_l\,, for\ q \to 0$$

Thus, for $\alpha > 0$, the Clayton copula has lower tail dependence. Furthermore, for $\alpha \longrightarrow \infty$ the coefficient converges to 1. This is because the Clayton copula tends to the co-monotonicity copula as α goes to infinity. The coefficient of upper tail dependence is zero.

Following T. Schmidt (2006), it is a little more complicated to show that for the Gumbel copula $\lambda_u = 2 - 2^{\frac{1}{\alpha}}$, thus the Gumbel copula exhibits upper tail dependence for $\alpha > 1$. The coefficient of lower tail dependence is zero instead.

No matters which copula we choose, if $\bar{C}$ is the survival copula associated with C, then

$$\bar{\lambda}_U = \lambda_L, \bar{\lambda}_L = \lambda_U$$

In A. Patton's (2007) opinion, the first area of application of copulas and of association measures in finance should be risk management. In fact, as fat tails or excess kurtosis in random variables distributions increase the likelihood of extreme events, the presence of positive tail dependence increases the likelihood of joint extreme events. To take this into account, risk managers need to focus on Value at Risk and other measures designed to estimate the probability of portfolio losses beyond a certain threshold.

4.COPULAS DERIVED FROM DISTRIBUTIONS

Generally, a bivariate copula can be represented by its distribution function, as K. Aas (2004) depicts, like this:

$$C(u_1, u_2) = P(U_1 \leq u_1, U_2 \leq u_2) = \int_{-\infty}^{u_1} \int_{-\infty}^{u_2} c(F_1(x_1), F_2(x_2))\, du_1 du_2$$

Where $c(F_1(x_1), F_2(x_2))$ is the density of the copula. Otherwise, this is a general framework. If we want to get more specific, we must distinguish between two parametric families of copulas: implicit and explicit. The so-called implicit copulas owe their name to the double integral at the right-hand side of equation, that is implied by a well-known distribution function. For explicit copulas, instead, this double integral has a simple closed form. Before analysing specifically distribution functions of the best known implicit and explicit copulas, we avail ourselves of a nice graphical representation provided by A. Patton (2007). Here behind level curves of some bivariate copula densities are shown, constructed using Sklar's theorem. Different parametric copulas are drawn, while all have marginal distributions $F_1 = F_2 = N(0,1)$, and linear correlation is constrained to be 0.5 in all cases. In the upper left there are the elliptical contours of a bivariate Normal copula, where both margins and copula are meant to be Normal. The scope of the figure is offering a rapid idea of what was said before. As an example, we can compare what was written above, about different coefficients of tail dependence, with corner shapes that level curves assume. As we noticed previously, elliptical copulas have identical lower and upper tail dependence. For both Normal and Student t-copulas, the shape of level curves in the upper right corner and downward left corners, that are respectively higher and lower tail dependence, is symmetrical, so identical. However, tails of Student t-copula are both much more slanted than tails of Normal copula, as Normal copula has zero tail dependence, apart from the case where ρ=1, while for Student t-copula it is in both case positive. For what concerns Archimedean copulas, we easily see the negative tail dependence of Clayton copula and positive tail dependence of Gumbel copula. Archimedean copulas are, with respect to elliptical copulas, asymmetric.

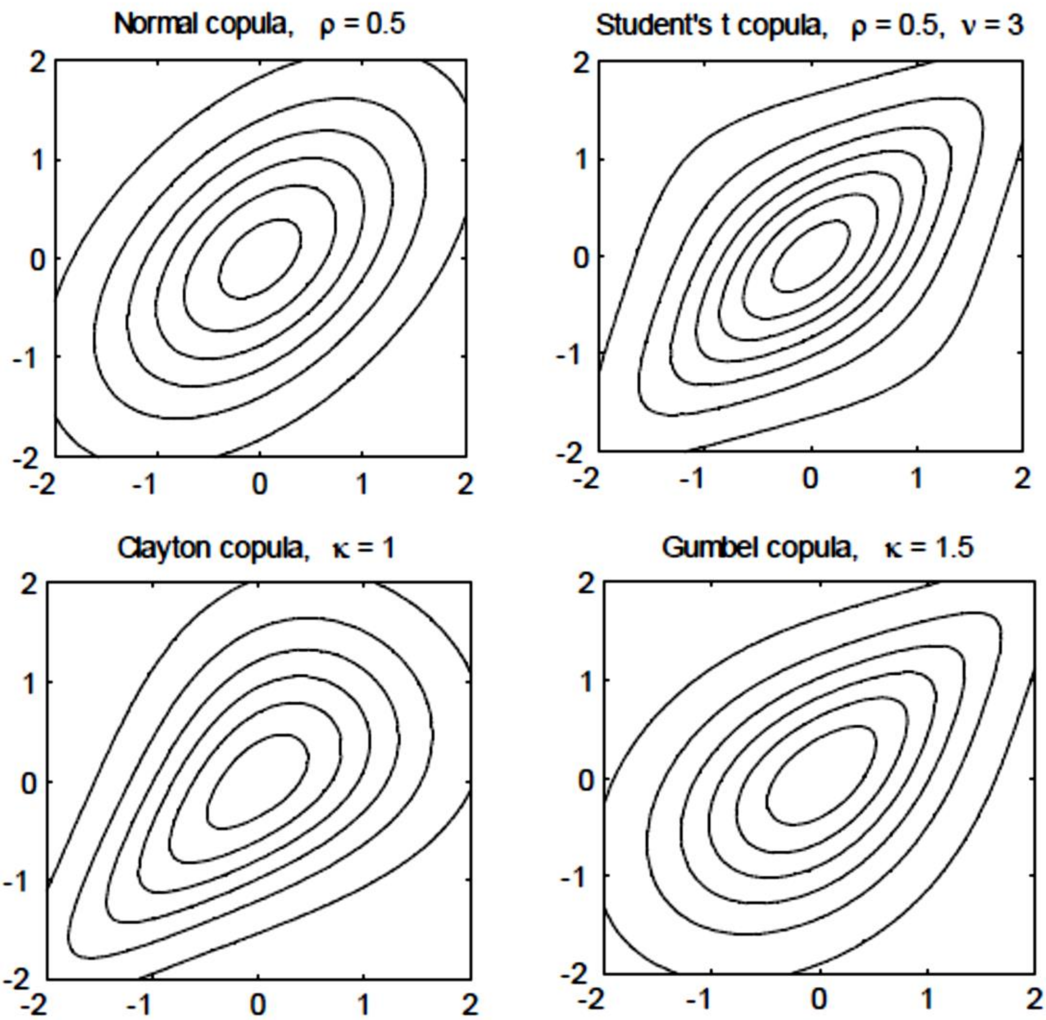


Figure 6

4.1.Implicit copulas

As now we are to start exposing the best known implicit copulas, it is important to underline that they do not have a simple closed form, but they are implied by multivariate distribution functions. A multivariate normal distribution function will lead to a Gaussian copula, while a multivariate Student t-distribution function will lead to a t-copula. In order to state that the joint distribution function of a random vector (X_1, X_2) constitutes a Gaussian copula, we should be sure that the univariate marginal distributions are both Gaussian. These margins, then, must be linked by a unique Normal copula function C. J.F. Jouanin, G. Rapuch, G. Riboulet and T. Roncalli (2001) define the bivariate Gaussian copula C as follows:

$$C_\rho^{Ga}(u_1, u_2) = \phi_\Sigma(\phi^{-1}(u_1), \phi^{-1}(u_2))$$

Where ρ is the parameter of the copula: the linear correlation coefficient in the case of a Normal copula. Σ is the 2x2 matrix with 1 on the diagonal and ρ otherwise. ϕ_Σ is the joint bivariate distribution function with zero mean and correlation matrix Σ and $\Phi^{-1}(\cdot)$ is the inverse of the standard univariate Gaussian distribution function.

Therefore, $\phi_\Sigma(\phi^{-1}(u_1), \phi^{-1}(u_2)) =$

$$\int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho_{X_1X_2}^2}} exp\left(\frac{2\rho_{X_1X_2}st - s^2 - t^2}{2\left(1-\rho_{X_1X_2}^2\right)}\right) dsdt$$

For normal and elliptical distributions independence is equivalent to zero linear correlation. Hence for ρ= 0, the Gaussian copula equals the independence copula. On the other side, if ρ=1 we obtain the co-monotonicity copula, while for ρ= -1 the counter-monotonicity copula is got. Gaussian copula interpolates between these three fundamental dependency structures via one simple parameter: correlation coefficient ρ.

The following representation has been proved by T. Roncalli (2002) to be equivalent to the previous one:

$$C^{Ga}(u_1, u_2) = \Phi_\Sigma\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right) = \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho_{X_1X_2}\Phi^{-1}(t)}{\sqrt{1-\rho_{X_1X_2}^2}}\right) dt$$

The density of the Gaussian copula is:

$$\frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{\zeta_1^2 + \zeta_2^2}{2} + \frac{2\rho\zeta_1\zeta_2 - \zeta_1^2 - \zeta_2^2}{2(1-\rho^2)}\right)$$

Where $\zeta_1 \coloneqq \Phi^{-1}(u_1), \zeta_2 \coloneqq \Phi^{-1}(u_2)$

As the copula is absolutely continuous, we can integrate the density into the expression of the copula, obtaining this:

$$C^{Ga}(u_1, u_2) = \int_0^{u_1} \int_0^{u_2} \frac{1}{\sqrt{1-\rho_{X_1X_2}^2}} \exp\left(\frac{2\rho_{X_1X_2}x_1x_2 - x_1^2 - x_2^2}{2\left(1-\rho_{X_1X_2}^2\right)} + \frac{x_1^2 + x_2^2}{2}\right) dsdt$$

Where $x = \Phi^{-1}(s), y = \Phi^{-1}(t)$.

As ρ is the unique parameter of the copula, and it represents linear correlation between marginal distributions, Y. Malevergne and D. Sornette (2001) conclude that the Gaussian copula is completely determined by the knowledge of the correlation matrix.

Linear correlation is expressed as:

$$Corr(X_1, X_2) \coloneqq \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1) \cdot Var(X_2)}}$$

And it fully describes the dependence structure. This remains true in the whole family of elliptical distributions, while it is totally wrong outside this family and risks to produce many fallacies in the dependence analysis. Specifically, note that matrix Σ is a correlation matrix, obtained from the covariance matrix by scaling each component by variance.

The covariance matrix:

$$\tilde{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

Leads to the correlation matrix Σ:

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

From the bivariate case U. Cherubini, E. Luciano and W. Vecchiato (2004) easily deduce the multivariate case: the Gaussian copula for a correlation matrix R is given by

$$C_R^{Ga}(\boldsymbol{u}) = \Phi_R(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n))$$

Where R is a symmetric, positive definite matrix with diagonal(R)=$(1,1,\dots,1)^T$ and ϕ_R is the standardized multivariate normal distribution with correlation matrix R. Φ^{-1} is the inverse of the standard univariate normal distribution function Φ.

As in the bivariate case, the Gaussian copula generates the standard Gaussian joint distribution function whenever the marginal distributions are standard normal. U. Cherubini, E. Luciano and W. Vecchiato (2004) advise that, for any other marginal choice, the Gaussian copula does not give a standard jointly normal vector. In order to have a visual representation of the phenomenon, and more generally of the effect of "coupling" the same copula with different marginal distributions, let us consider the joint density functions in the following figures. Figures 7.a and 7.b show respectively density and level curves of the distribution obtained coupling a Gaussian copula with two standard normal marginal distributions. Figure 8.a and figure 8.b are referred to a Gaussian copula with two three-degrees of freedom Student t-marginal distributions. Both for figures 7 and 8 is considered ρ=0.2. Figures 9.a and 9.b illustrate density and level curves of a Gaussian copula with standard normal marginal distributions and ρ=0.9; while figures 10.a and 10.b are referred to a Gaussian copula with two three-degrees of freedom Student t-marginal distributions. Even in this last case ρ=0.9. It does not depend on the correlation coefficient we choose: in every case, the same copula, with different marginal distributions, presents a different graphical joint behaviour, that indicates that marginal choice influences the density. As we could expect, the effect of Student marginal distributions is increasing tail probabilities.

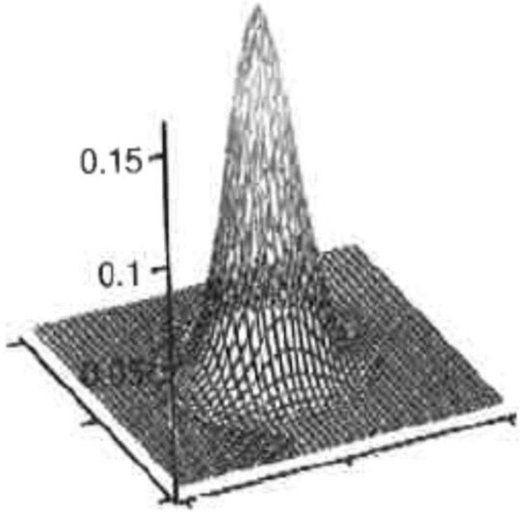


Figure 7.a

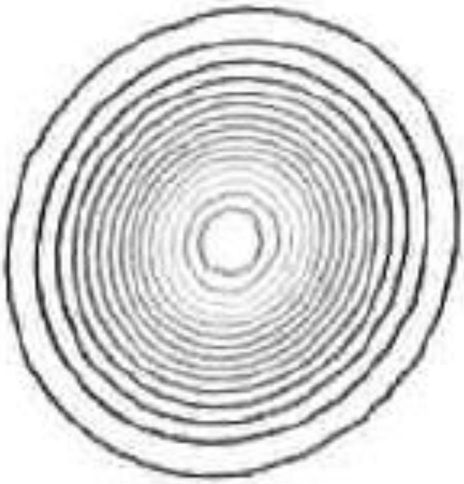


Figure 7.b

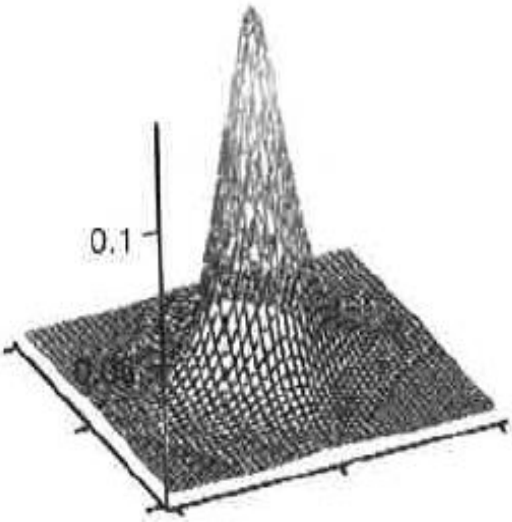


Figure 8.a

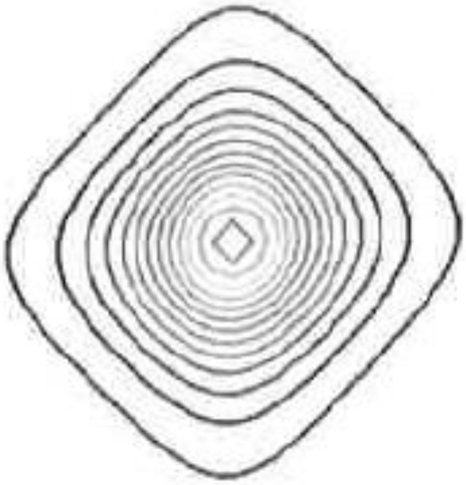


Figure 8.b

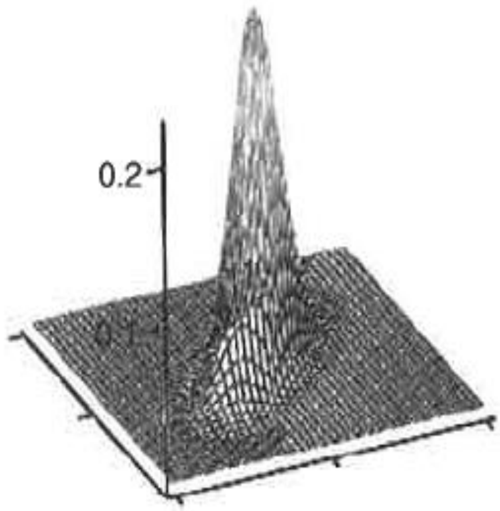


Figure 9.a

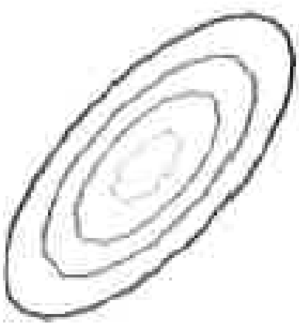


Figure 9.b

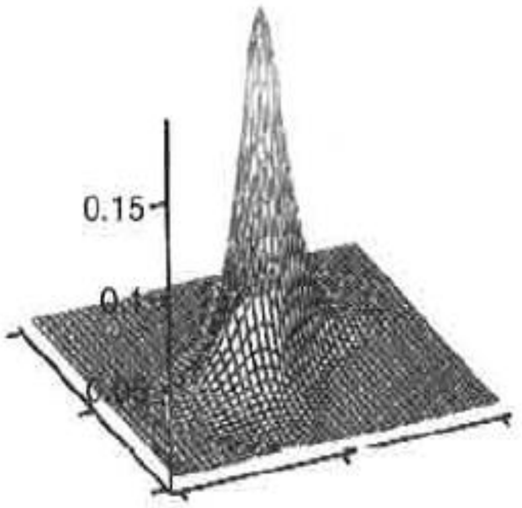


Figure 10.a

Figure 10.b

Now, U. Cherubini, E. Luciano and W. Vecchiato (2004) say that it is possible to easily determine the density of the multivariate Gaussian copula:

$$\frac{1}{(2\pi)^{\frac{n}{2}}|R|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\boldsymbol{x}^T R^{-1}\boldsymbol{x}\right) = c_R^{Ga}(\Phi(x_1),\Phi(x_2),\dots,\Phi(x_n)) \times \prod_{j=1}^{n}\left(\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x_j^2)\right)$$

Where $|R|$ is the determinant of R. we deduce that:

$$c_R^{Ga}\big(\Phi(x_1),\Phi(x_2),\dots,\Phi(x_n)\big) = \frac{\frac{1}{(2\pi)^{\frac{n}{2}}|R|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\boldsymbol{x}^T R^{-1}\boldsymbol{x}\right)}{\prod_{j=1}^{n}\left(\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x_j^2)\right)}$$

Let $u_j = \Phi(x_j)$, so that $x_j = \Phi^{-1}(u_j)$. The density can be rewritten as follows:

$$C_R^{Ga}(u_1,u_2,\dots,u_n) = \frac{1}{|R|^{\frac{1}{2}}} exp\left(-\frac{1}{2}\boldsymbol{\varsigma}^T(R^{-1}-I)\boldsymbol{\varsigma}\right)$$

Where $\varsigma = (\Phi^{-1}(u_1),\Phi^{-1}(u_2),\dots,\Phi^{-1}(u_n))^T$".

As Gaussian copula is parametrized by linear correlation coefficient, and it respects concordance order, Gaussian copula can be positively ordered with respect to ρ:

$$C_{\rho=-1}^{Ga} \prec C_{\rho<0}^{Ga} \prec C_{\rho=0}^{Ga} \prec C_{\rho>0}^{Ga} \prec C_{\rho=1}^{Ga}$$

Also, Gaussian copula is comprehensive: in fact it encompasses all the range of dependence, starting from counter-monotonic copula till co-monotonic and passing through the independence copula:

$$C_{\rho=-1}^{Ga} = C^- \; and \; C_{\rho=1}^{Ga} = C^+$$

In addition, $C_{\rho=0}^{Ga} = C^{\perp}$. As we have already stated, Gaussian copula does not show tail dependence: the unique exception is given in the case ρ=1:

$$\lambda_U = \lambda_L = \begin{cases} 0 & iff\ \rho < 1 \\ 1 & iff\ \rho = 1 \end{cases}$$

Here behind T. Schmidt (2006) offers a nice picture of a bivariate Gaussian copula on the left, and of a bivariate Student t-copula on the right. Both copulas have correlation coefficient ρ=0.3 and the t-copula has 2 degrees of freedom. It is worth remarking that the behaviour at the four corners is different, while in the centre they are quite similar.

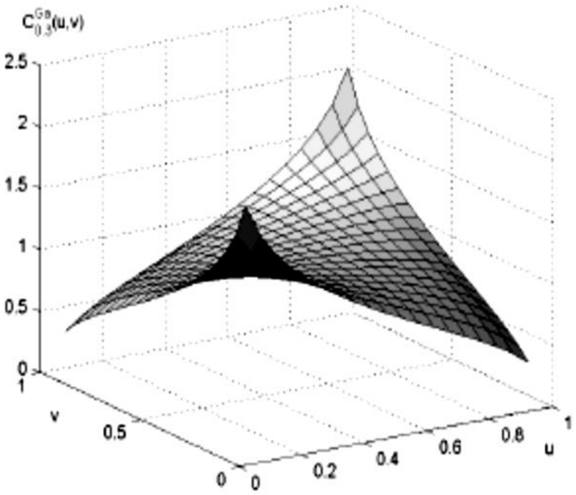


Figure 11.a

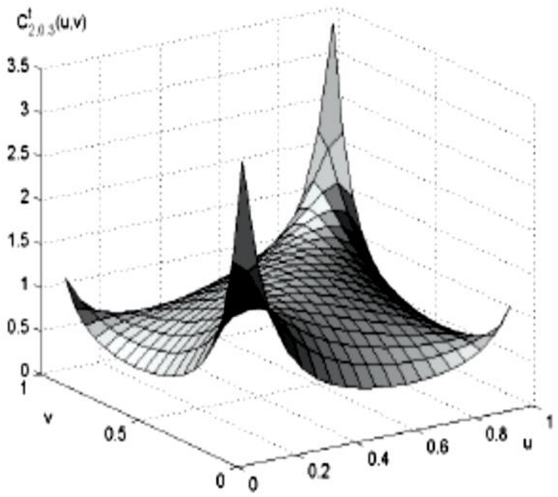


Figure 11.b

Although having the same correlation, extreme cases, represented by the corners, are much more pronounced in the t-copula. This gets particularly evident in (0,0) and (1,1) corners, that refer to the possibility that two very negative or very positive events occur simultaneously. Student t-copula is able to describe extreme cases duly to tail dependence. Anyway, we can even notice that t-copula shows peaks at the (0,1) and (1,0) corners. The peaks in these corners stem from a negative value in X_1 and a positive value in X_2, and vice versa. If we have an independent copula, density should rise up at all four corners symmetrically. When we start introducing some correlation, like 0.3 in previous figures, probabilities change and it is more likely having values with the same sign. As a consequence, peaks in (0,0) and (1,1) corners are higher than others.

Now we are to pass to the Student t-copula. U. Cherubini, E. Luciano and W. Vecchiato (2004) start developing Student t-copula from the univariate Student t-distribution function. Let $t_\nu\colon \Re \rightarrow \Re$ be the central univariate Student t-distribution function, with ν degrees of freedom:

$$t_\nu(x) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{(\nu+1)}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{s^2}{\nu}\right)^{-\frac{\nu+1}{2}} ds$$

Where Γ is the Euler function.

Let $\rho \in I$ *and* $t_{\rho,\nu}$ be the bivariate distribution corresponding to t_ν:

$$t_{\rho,\nu}(x_1,x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi\sqrt{1-\rho^2}}\left(1+\frac{s^2+t^2-2\rho st}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}} dsdt$$

The bivariate Student t-copula, $T_{\rho,\nu}$, is defined as:

$$T_{\rho,\nu}(u_1,u_2) = t_{\rho,\nu}\left(t_\nu^{-1}(u_1)t_\nu^{-1}(u_2)\right)$$

$$= \int_{-\infty}^{t_\nu^{-1}(u_1)}\int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}}\left\{1+\frac{s^2-2\rho st+t^2}{\nu(1-\rho^2)}\right\}^{-\frac{\nu+2}{2}} dsdt$$

Where ρ, that is linear correlation coefficient, and ν are the parameters of the copula, and t_ν^{-1} is the inverse of the standard univariate Student t-distribution with ν degrees of freedom, expectations 0 and variance $\frac{\nu}{\nu-2}$.

The Student t-dependence structure introduces an additional parameter compared with the Gaussian copula, namely the degrees of freedom ν. Increasing the value of ν decreases the tendency to exhibit extreme co-movements. As Y. Malevergne and D. Sornette (2001) resume, since the Student t-distribution tends to the normal distribution when ν goes to infinity, the Student t-copula tends to the Gaussian copula as ν⟶+∞.

As U. Cherubini, E. Luciano and W. Vecchiato (2004) formulate, bivariate Student t-copula density is:

$$c_{\nu,\rho}^{s}(u_1,u_2) = \rho^{-\frac{1}{2}}\frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)^2}\,\frac{\left(1+\frac{\varsigma_1^2+\varsigma_2^2-2\rho_{\varsigma_1\varsigma_2}}{\nu(1-\rho^2)}\right)^{\frac{-(\nu+2)}{2}}}{\prod_{j=1}^{2}\left(\frac{1+\varsigma_j^2}{\nu}\right)^{\frac{-(\nu+1)}{2}}}$$

Where $\varsigma_1 = t_\nu^{-1}(u_1), \varsigma_2 = t_\nu^{-1}(u_2)$ and the copula itself is absolutely continuous.

From the bivariate case, it is easy to expand to the multivariate Student t-copula. Let R be a symmetric, positive definite matrix with $diag\,(R) = (1,1,\dots,1)^T$ and $t_{R,\nu}$ the standardized multivariate Student t-distribution with correlation matrix R and ν degrees of freedom:

$$t_{R,\nu}(x_1,x_2,\dots,x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\dots\int_{-\infty}^{x_n} \frac{\Gamma\left(\frac{\nu+n}{2}\right)|R|^{-\frac{1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{\frac{n}{2}}}\left(1+\frac{1}{\nu}\boldsymbol{x}^T R^{-1}\boldsymbol{x}\right)^{-\frac{\nu+n}{2}} dx_1 dx_2 \dots dx_n$$

The correlation matrix is obtained from an arbitrary covariance matrix by scaling each component to variance 1. The multivariate Student t-copula is then defined as follows:

$$T_{R,\nu}(u_1,u_2,\dots,u_n) = t_{R,\nu}\left(t_\nu^{-1}(u_1),t_\nu^{-1}(u_2),\dots,t_\nu^{-1}(u_n)\right)$$

$$= \int_{-\infty}^{t_\nu^{-1}(u_1)}\int_{-\infty}^{t_\nu^{-1}(u_2)}\dots\int_{-\infty}^{t_\nu^{-1}(u_n)} \frac{\Gamma\left(\frac{\nu+n}{2}\right)|R|^{-\frac{1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{\frac{n}{2}}}\left(1+\frac{1}{\nu}\boldsymbol{x}^T R^{-1}\boldsymbol{x}\right)^{-\frac{\nu+n}{2}} dx_1 dx_2 \dots dx_n$$

Where t_ν^{-1} is the inverse of the univariate Student t-distribution function with ν degrees of freedom. Using the canonical representation, it turns out that the copula density for the multivariate Student t-case is:

$$c_{R,\nu}(u_1,u_2,\ldots,u_n) = |R|^{-\frac{1}{2}}\frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\right)^n \frac{\left(1+\frac{1}{\nu}\varsigma^T R^{-1}\varsigma\right)^{-\frac{\nu+n}{2}}}{\prod_{j=1}^n\left(1+\frac{\varsigma_j^2}{\nu}\right)^{-\frac{\nu+1}{2}}}$$

Where $\varsigma_j = t_\nu^{-1}(u_j)$.

4.2.Explicit copulas: Archimedean copulas

T. Schmidt (2006) says that there is a class of copulas, Archimedean copulas typically, that can be stated directly and have quite a simple form, in contrast to copulas derived from distributions, known as implicit copulas. According to U. Cherubini, E. Luciano and W. Vecchiato (2004), Archimedean copulas can be constructed using a function $\varphi: I \longrightarrow \mathfrak{R}^{*+}$, continuous, decreasing, convex and such that φ(1) =0. A similar function φ is called a generator. It becomes a strict generator whenever φ(0) = +∞.

The pseudo-inverse of φ is defined, as follows:

$$\varphi^{[-1]}(u_1) = \begin{cases} \varphi^{-1}(u_1) & 0 \le u_1 \le \varphi(0) \\ 0 & \varphi(0) \le u_1 \le +\infty \end{cases}$$

This pseudo-inverse is such that, if composed with the generator, it gives the identity, as ordinary inverses do for functions with domain and range $\mathfrak{R}$:

$$\varphi^{[-1]}\big(\varphi(u_1)\big) = u_1, \forall u_1 \in I$$

In addition, it coincides with the usual inverse if φ is a strict generator.

Revisiting examples above more closely, we can realize that the bivariate implicit copula itself was always in the form:

$$C(u_1,u_2) = \Phi_\Sigma(\Phi^{-1}(u_1),\Phi^{-1}(u_2))$$

Similarly, in more than 2 dimensions, with the condition that ϕ^{-1} is completely monotonic on [0,∞], the function $C: [0,1]^n \longrightarrow [0,1]$ defines an implicit copula as:

$$C_R^{Ga}(\boldsymbol{u}) = \Phi_R(\Phi^{-1}(u_1),\Phi^{-1}(u_2),\ldots,\Phi^{-1}(u_n))$$

Otherwise, Cherubini, Luciano and Vecchiato (2004) state that, given a generator and its pseudo-inverse, an Archimedean copula C^A is generated as follows:

$$C^A(u_1,u_2) = \varphi^{[-1]}\big(\varphi(u_1)+\varphi(u_2)\big)$$

In addition to this, R.B. Nelsen (1999) proves that level curves of an Archimedean copula are convex, and that the density of an Archimedean copula is:

$$C^A(u_1,u_2) = \frac{-\varphi''\big(C(u_1,u_2)\big)\varphi'(u_1)\varphi'(u_2)}{\Big(\varphi'\big(C(u_1,u_2)\big)\Big)^3}$$

For what concerns dependence, U. Cherubini, E. Luciano and W. Vecchiato (2004) say that Archimedean copulas can be easily related to measures of association. C. Genest and J. MacKay (1986) demonstrate that Kendall's tau is given by:

$$t = 4\int_I \frac{\varphi(u_1)}{\varphi'(u_1)} du_1 + 1$$

Where $\varphi'(u_1)$ exists since the generator is convex. In addition to this, Genest and MacKay (1986) guarantee that conditions on the generators of two Archimedean copulas φ_1 and φ_2 can be given, and this assures that the corresponding generated copulas are to be ordered in the same way as their association parameters. If we denote by C_i the copula that corresponds to $\varphi_i, i = 1{,}2$, then

$$C_1 \prec C_2 \leftrightarrow \tau_{C_1} \leq \tau_{C_2}$$

or, equivalently,

$$C_1 \prec C_2 \leftrightarrow \rho_{SC1} \leq \rho_{SC2}$$

Where τ is Kendall's tau and ρ is Spearman's rho. This means that the order between copulas can be resumed by just an association measure like rank correlation: Kendall's tau and Spearman' rho. This result has been demonstrated by H. Joe (1997).

Later, we will analyse specifically the best known examples of Archimedean copulas, like Gumbel, Clayton and Frank copulas. By now, we take the general definition of upper and lower tail dependence given by U. Cherubini, E. Luciano and W. Vecchiato (2004). If Archimedean copula C has upper tail dependence, then the coefficient of upper tail dependence is $\lambda_U = 2 - 2\lim_{s\to 0+} \frac{\varphi'(s)}{\varphi'(2s)}$

While coefficient of lower tail dependence is $\lambda_L = 2\lim_{s\to +\infty} \frac{\varphi'(s)}{\varphi'(2s)}$.

Now, for the scope of our analysis, among Archimedean copulas, we are to choose one-parameter copulas. By one-parameter we mean copulas that are based on a generator $\varphi_\alpha(t)$, indexed by a unique real parameter α. By choosing the generator, one obtains a different type of copula.

We will start with Gumbel copula: Gumbel family has been introduced by Gumbel in 1960. Since it has been analysed by P. Hougaard, it is also known as the Gumbel-Hougaard family. By T. Schmidt (2006), the bivariate Gumbel copula is given in the following form:

$$C_\alpha^{Gu}(u_1,u_2) = exp\left[-((-ln\, u_1)^\alpha + (-\ln u_2)^\alpha)^{\frac{1}{\alpha}}\right]$$

Where $\alpha \in [1, \infty)$. For α=1 we have the product copula, while for $\alpha \longrightarrow \infty$ the Gumbel copula tends to co-monotonicity copula; so that Gumbel copula interpolates between independence and perfect positive dependence. This is a perfect example of a copula with tail dependence in just one corner, the one corresponding to joint extreme positive behaviour. In figure 12 we can take a glimpse of Gumbel copula positive tail behaviour and its level curves, in correspondence to α=1.5.

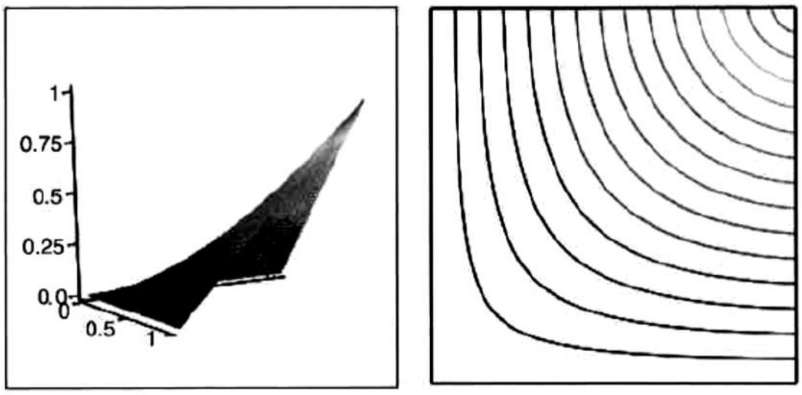


Figure 12

Passing to Clayton family, U. Cherubini, E. Luciano and W. Vecchiato (2004) remind that it is a comprehensive copula: it encompasses counter-monotonicity, independence and even co-monotonicity. Product copula is due to α=0, the lower Fréchet bound to α= -1 and upper Fréchet bound to $\alpha \to +\infty$. As we previously did for Gumbel copula, in figure 13 Clayton positive tail behaviour and corresponding level curves are presented, in correspondence of α=6.

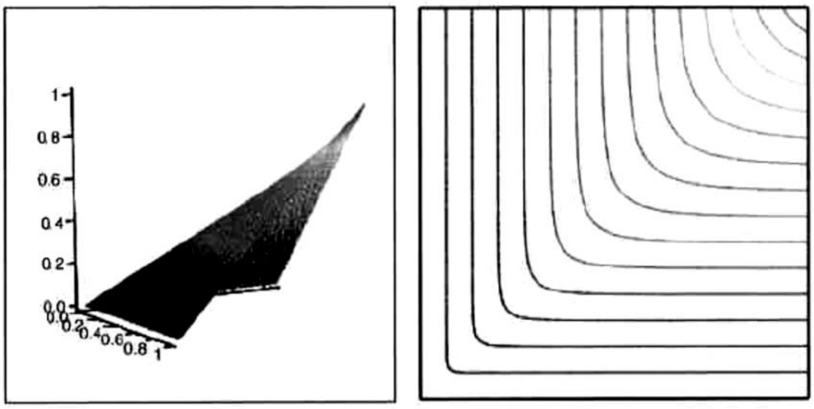


Figure 13

We are to end with Frank copula. It reduces to product copula if α=0, and reaches lower and upper Fréchet bounds for $\alpha \to -\infty$ and $\alpha \to +\infty$, respectively. In figure 14 we show its behaviour in the tail and level curves in correspondence to α=0.5.

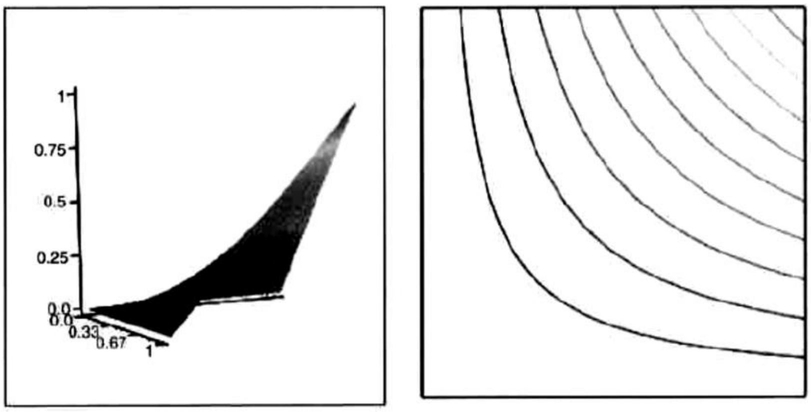


Figure 14

In the following table, by the courtesy of U. Cherubini, E. Luciano and W. Vecchiato (2004), we can resume some well-known families of bivariate Archimedean copulas and their respective generators:

Gumbel (1960)	
$\varphi_\alpha(t)$	$(-\ln t)^\alpha$
Range for α	[1, +∞)
$C(u_1, u_2)$	$exp\left\{-[(-\ln u_1)^\alpha + (-\ln u_2)^\alpha]^{1/\alpha}\right\}$
Clayton (1978)	
$\varphi_\alpha(t)$	$\frac{1}{\alpha}(t^{-\alpha} - 1)$
Range for α	$[-1,0)U(0,+\infty)$
$C(u_1, u_2)$	$max\left[(u_1{}^{-\alpha} + u_2{}^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0\right]$
Frank (1979)	
$\varphi_\alpha(t)$	$-\ln\frac{\exp(-\alpha t) - 1}{\exp(-\alpha) - 1}$
Range for α	(-∞,0)U(0,+∞)
$C(u_1, u_2)$	$-\frac{1}{\alpha} ln\left(1 + \frac{(exp(-\alpha u_1) - 1)(exp(-\alpha u_2) - 1)}{exp(-\alpha) - 1}\right)$

In the following table, instead, C. Kharoubi-Rakotomalala and F. Maurer (2013) give us the possibility to compare tail dependence of Archimedean copulas:

	Upper tail dependence	Lower tail dependence	Conditions
	λ_U	λ_L	
Gumbel	$2 - 2^{\frac{1}{\alpha}}$	0	Asymmetric upper dependence if $\alpha > 1$
Clayton	0	$2^{-\frac{1}{\alpha}}$	Asymmetric lower dependence if α>0
Frank	0	0	Asymptotic independence

And here, finally, U. Cherubini, E. Luciano and W. Vecchiato (2004) summarize association measures:

Family	Kendall's tau	Spearman's rho
Gumbel (1960)	$1-\alpha^{-1}$	No closed form
Clayton (1978)	$\frac{\alpha}{(\alpha+2)}$	Complicated expressions
Frank (1979)	$1+4\frac{[D_1(\alpha)-1]}{\alpha}$	$1-12\frac{[D_2(-\alpha)-D_1(-\alpha)]}{\alpha}$

U. Cherubini, E. Luciano and W. Vecchiato (2004) make us to remark that concordance measures of Frank copula require the so-called "Debye" function, defined as:

$$D_k(\alpha)=\frac{k}{\alpha^k}\int_0^{\alpha}\frac{t^k}{\exp(t)-1}dt,\quad k=1,2$$

When it comes to define an Archimedean copula in more than two dimensions, we have to consider a strictly decreasing and continuous generator function anyway, like this:

$$\varphi(u):[0,1]\rightarrow[0,\infty]$$

Let φ be a strict generator, C.H. Kimberling (1974) says that the function $C:[0,1]^d\rightarrow[0,1]$, defined by: $C(u_1,u_2,\dots,u_d)=\varphi^{-1}\big(\varphi(u_1)+\varphi(u_2)+\cdots+\varphi(u_d)\big)$ is a copula if and only if φ^{-1} is completely monotonic on [0,∞]. Retrieving Gumbel copula, in the d-dimensional case, with d>2, the generator is given by $\varphi(u)=(-\ln(u))^{\alpha}$, hence $\varphi^{-1}(t)=exp\left(-t^{\frac{1}{\alpha}}\right)$; it is completely monotonic if α>1. The Gumbel d-copula is therefore:

$$C(u_1,u_2,\dots,u_d)=exp\left\{-\left[\sum_{i=1}^{d}(-\ln u_i)^{\alpha}\right]^{\frac{1}{\alpha}}\right\}\quad with\ \alpha>1$$

For what concerns Clayton copula, K. Aas (2004) compares it to a Student t-copula. Even Student t-copula allows for joint extreme events, but it is symmetric: it gives the same probabilistic weight to extreme negative and extreme positive events. However, as in economics and finance extreme negative events are more probable than extreme positive events, for the scope of our further analysis, we could recur to a Clayton copula, that is asymmetric: it exhibits greater dependence in the negative tail than in the positive. The generator of the Clayton copula is given by $\varphi(u)=u^{-\alpha}-1$, hence $\varphi^{-1}(t)=(t+1)^{-\frac{1}{\alpha}}$; it is completely monotonic if α >0. The Clayton d-copula is therefore:

$$C(u_1,u_2,\dots,u_d)=\left[\sum_{i=1}^{d}u_i^{-\alpha}-d+1\right]^{-\frac{1}{\alpha}}\quad \text{With } \alpha>0$$

Last in the list, the generator of the Frank n-copula is given by $\varphi(u) = \ln\left(\frac{exp(-\alpha u)-1}{exp(-\alpha)-1}\right)$ hence $\varphi^{-1}(t) = -\frac{1}{\alpha}\ln\left(1 + e^{t}(e^{-\alpha} - 1)\right)$ is completely monotonic if α>0.

We can deduce that Frank d-copula is given by:

$$C(u_1, u_2, \dots, u_d) = -\frac{1}{\alpha}\ln\left\{1 + \frac{\prod_{i=1}^{d}(e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{d-1}}\right\}$$

With α>0 when $d \geq 3$.

In order to get a graphical resume of Archimedean copulas, we will avail ourselves of the following picture, from T. Schmidt (2006):

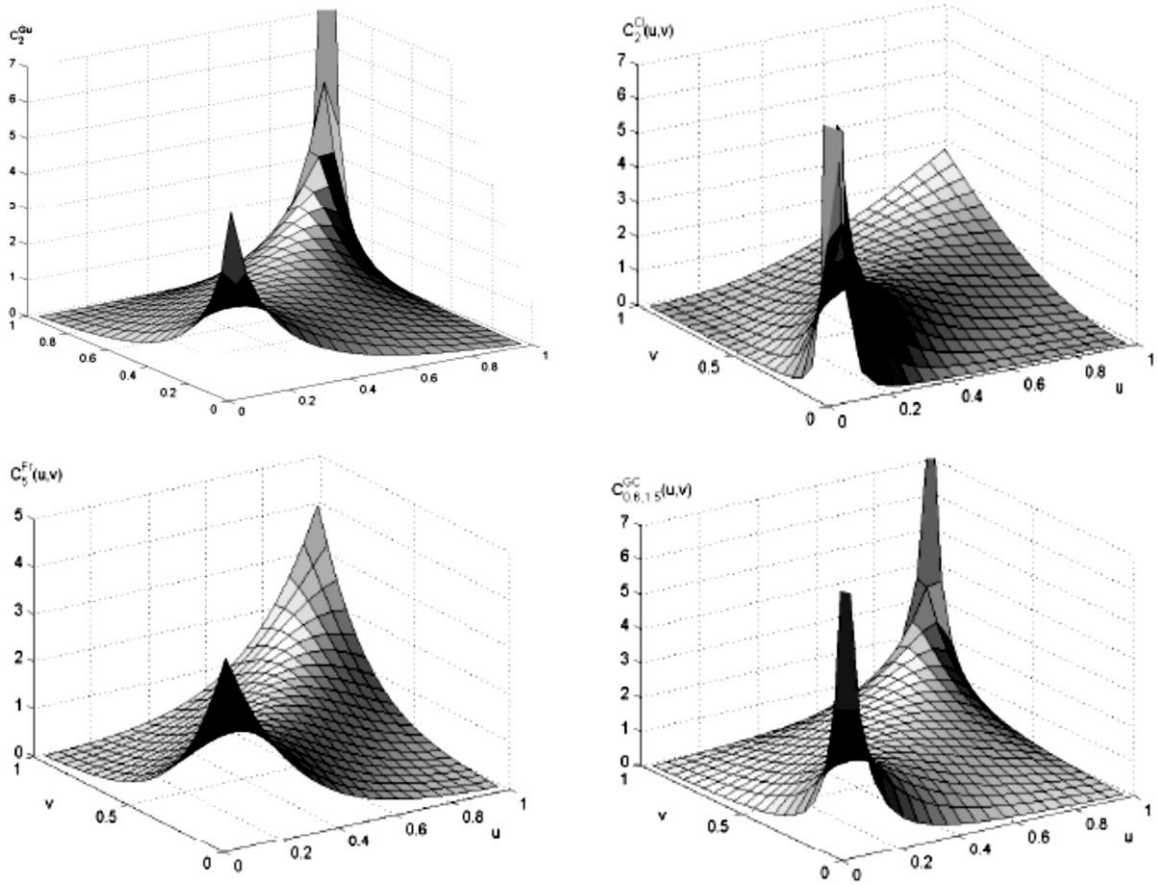


Figure 15

It shows densities of Gumbel copula in upper left, Clayton copula in upper right, Frank copula in lower left and generalized Clayton copula in lower right. In all cases, α=2, and, for generalized Clayton copula, we have an additional parameter to be taken into account, that is δ, equal to 2. Note that for δ=1 the standard Clayton copula is attained. All copulas, with the exception of Frank copula, have been cut at a level of 7. It may be spotted at first sight, that these copulas have different behaviours at the lower and upper corners, that are respectively the points (0,0) and (1,1). As we have already noticed before, Gumbel copula shows an extremely uprising peak at (1,1), while a less pronounced behaviour at (0,0), if compared with other copulas. By this, we want to say that Gumbel copula has strong upper tail dependence. For what concerns Clayton copula, the situation is reversed: low tail dependence in the positive tail, but much more evident negative peak. For Frank copula there is no upper nor lower strong tail dependence. It may be glimpsed that standard Clayton copula differs quite dramatically from the generalized one in the behaviour at the corners: generalized Clayton copula shows strong tail behaviour at both corners in contrast to the standard one.

5.STATISTICAL INFERENCE FOR COPULAS: ESTIMATING PARAMETERS

As U. Cherubini, E. Luciano and W. Vecchiato (2004) advise, similarly to most multivariate statistical models, much of the classical statistical inference theory is not applicable for copulas. Fortunately, K. Aas (2004) suggests that there are mainly two ways to infer copula parameters: a fully parametric method and a semi-parametric method. The semi-parametric method is the asymptotic Maximum Likelihood Estimation (MLE). This method does not take into account any parametric assumption for marginal distributions. A possible expansion of Maximum Likelihood Estimation technique is provided by A. Patton (2007): if the model is such that the parameters of the marginal distributions can be separated from each other and from those of the copula, then Multi-Stage Likelihood (ML) estimation is an option. This method is the fully parametric one and can be even known as the “Inference Functions for Margins” (IFM) method, due to H. Joe (1997). It involves estimating all parameters of the marginal distributions separately one from the other in the first step, via univariate maximum likelihood. Then, second step corresponds to plugging each parametric margin into the copula likelihood function, and this likelihood function is maximized with respect to copula parameters. Both estimation techniques require a numerical optimization of an objective function, as likelihood of a multivariate model substantially involves mixed derivatives.

5.1.Maximum Likelihood Estimation

J. Myung (2002) demonstrates that Maximum Likelihood Estimation guarantees many optimal properties in estimation. To begin with, it is sufficient, in the sense that it gives complete information about parameters of interest. Secondly, it is consistent: the true value of the parameters is recovered asymptotically for sufficiently large samples. Thirdly, it is efficient: it achieves asymptotically the lowest possible variance in parameter estimation. Last but not least, the same Maximum Likelihood Estimation is obtained independently of the parametrization used. The principle of Maximum Likelihood Estimation is to find out the value of the parameters vector that maximizes likelihood function. In order to apply Maximum Likelihood Estimation techniques, we are to refer to R. Lucchetti’s instructions. He starts by extracting a sample constituted of n random variables X_i, independently and identically distributed, taken from a population X with probability function $f(x,\theta)$. With this sample, we are to build likelihood function, that represents probability function of the sample itself. We hypothesize that this probability function is written in function of parameters vector θ, while sample realizations are fixed. Analytically, we have

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

Statistical function $\hat{\theta} = t(x_1, x_2, \dots, x_n)$ is called maximum likelihood estimation if, in correspondence to every extracted sample, it assigns to one parameter vector θ a value that maximizes likelihood function. By symbols:

$$\max L(x,\theta) = L(x,\hat{\theta})$$

Maximum likelihood estimation is defined as:

$$\hat{\theta} = \arg\max L(x,\theta)$$

in order to calculate maximum likelihood estimator we have to recur to log-likelihood function, that is obtained applying natural logarithm. So, it results:

$$l(x,\theta) = \ln L(x,\theta)$$

Given that logarithmic function is an increasing monotone transformation, when we pass to log-likelihood we are not going to lose $L(x,\theta)$ function characteristics; moreover, we get a simpler analytical expression to work with. If we have independently, identical distributed random variables, joint density function of the sample can be expressed as marginal product. By logarithmic properties, from this $L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^{n} f(x_i, \theta)$ we can carve out log-likelihood function as summation, in fact:

$$l(x,\theta) = \ln\left[\prod_{i=1}^{n} f(x_i,\theta)\right] = \sum_{i=1}^{n} \ln f(x_i,\theta)$$

An important log-likelihood property is the following:

$$\frac{\partial l(x,\theta)}{\partial\theta} = [L(x,\theta)]^{-1}\frac{\partial L(x,\theta)}{\partial\theta}$$

Otherwise, the most important property of log-likelihood function is the one that constitutes principal reason to use this technique. As realizations x_i of n random variables are implied, log-likelihood is a random function of unknown parameters vector. This means that, given θ, log-likelihood function gives back a random variable. Alternatively, we can think that, for each possible sample realization, a different θ is to be associated. If this function has an expected value, it will be a non-stochastic function of parameters vector. Now we can observe the figure: dotted lines correspond to every log-likelihood function that can be observed for each sample realization $f(x_1, \dots, x_n, \theta_0)$, while the continue line represents function $E[l(x, \theta)]$. This function assumes its maximum just in correspondence of θ_0, that is, the true value of x density function.

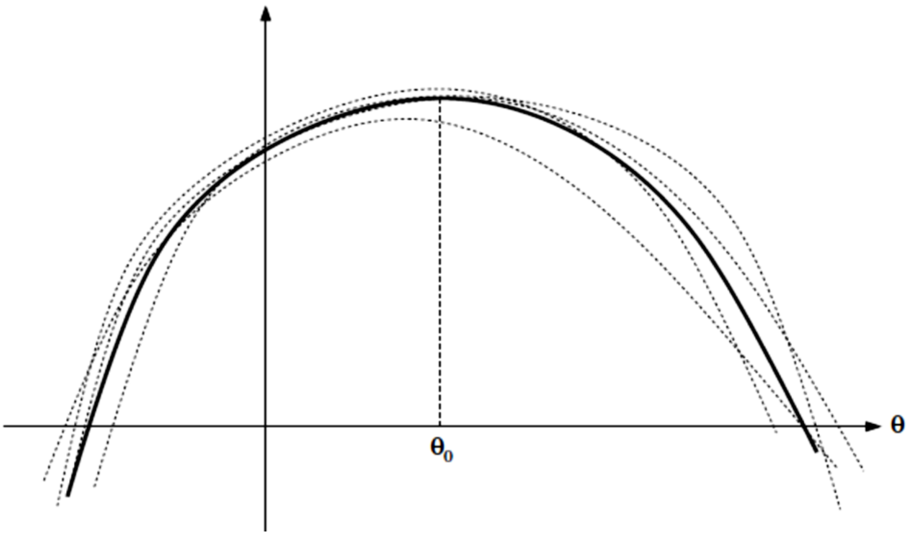


Figure 16

Looking at the figure, it comes logic to think that the maximum point of function $l(x, \theta)$ could be used as estimator for θ_0. R. Lucchetti grants us that the so obtained estimator will be consistent. For maximizing $l(x, \theta)$ estimator, R. Lucchetti sets the following first order conditions:

$$s(x, \theta) = \frac{\partial l(x, \theta)}{\partial \theta} = 0$$

Where function $s(x, \theta)$ is named score. In case we have a sample composed of n random variables, independent and identically distributed, score can be defined as:

$$s(x, \theta) = \sum_{i=1}^{n} s(x_i, \theta) = \sum_{i=1}^{n} \frac{\partial \ln f(x_i, \theta)}{\partial \theta}$$

The score is the gradient vector that contains partial derivatives of log-likelihood equation, calculated with respect to parameter θ. Being expressed in function of random samples, the score is a random variable itself. When $\theta = \theta_0$, its first and second order moments are, respectively:

$$E[s(x, \theta_0)] = 0$$

$$Var[s(x, \theta_0)] = I(\theta_0)$$

Where $I(\theta_0)$ is Fisher information matrix, valued in correspondence of the true parameter θ_0. $I(\theta)$ is defined as the opposite of log-likelihood Hessian matrix expected value, so it results as:

$$I(\theta) = -E\left[\frac{\partial^2 l(x, \theta)}{\partial \theta \partial \theta'}\right] = -E[H(x, \theta)]$$

When $\theta = \theta_0, I(\theta_0) = Var[s(x, \theta_0)] = -E[H(x, \theta_0)]$

In order to resume, we can say that Maximum Likelihood estimate is that value of θ for which score is zero and for which log-likelihood function is maximized. In the graph of figure 17, maximum likelihood estimation is point $\hat{\theta}$.

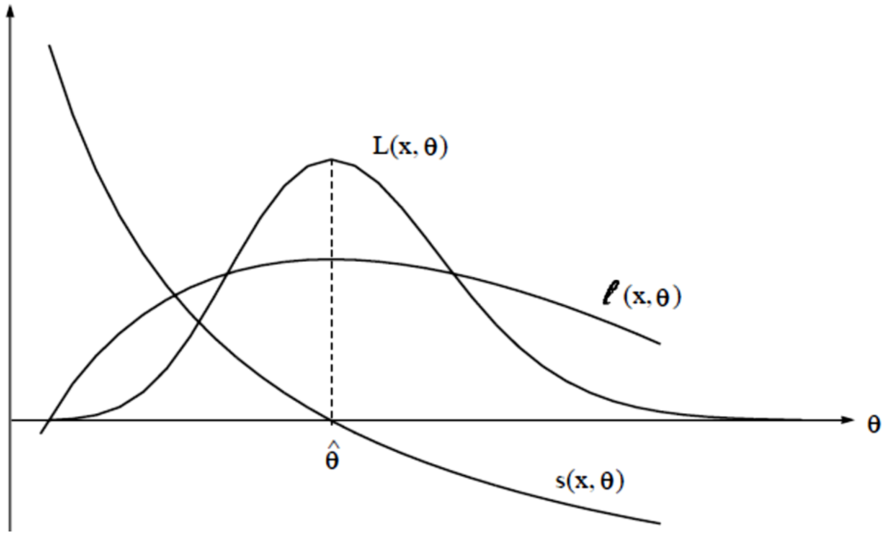


Figure 17

5.2.Inference for Margins Method

However, U. Cherubini, E. Luciano and W. Vecchiato (2004) advise us that Maximum Likelihood method is, unfortunately, computationally intensive, especially in the case of high dimensions. In fact, we would have to estimate jointly marginal distributions parameters and joint distribution parameters. Anyway, U. Cherubini, E. Luciano and W. Vecchiato (2004) propose even a nicer solution. First of all, they want us to remind canonical representation for a multivariate density function:

$$f(x_1, x_2, \dots, x_n) = c\big(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\big) \cdot \prod_{j=1}^{n} f_j(x_j)$$

Where $c\big(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\big) = \frac{\partial^n \left(C\left(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\right)\right)}{\partial F_1(x_1) \partial F_2(x_2) \dots \partial F_n(x_n)}$ is the n-th mixed partial derivative of the copula C, c is the copula density and f_j is standard univariate probability density function. Let $\mathfrak{X} = \{x_{1t}, x_{2t}, \dots, x_{nt}\}_{t=1}^{T}$, where t indicates the time, be the sample data matrix. Thus, we can redefine log-likelihood function as:

$$l(\boldsymbol{\theta}) = \sum_{t=1}^{T} \ln c\big(F_1(x_{1t}), F_2(x_{2t}), \dots, F_n(x_{nt})\big) + \sum_{t=1}^{T} \sum_{j=1}^{n} \ln f_j(x_{jt})$$

Where **θ** is the set of all parameters of both marginal distributions and the copula. If we look at log-likelihood function, we can notice that it is composed of two terms, both positive. The first involves copula density and copula parameters, while the second involves marginal distributions parameters. Taking this into account, H. Joe and J.J. Xu (1996) propose to estimate parameters in two separate steps. In the first step, we have to estimate marginal parameters vector, that will be called $\boldsymbol{\theta}_1$. We are to perform estimation of the univariate marginal distributions:

$$\widehat{\boldsymbol{\theta}}_1 = Arg\ Max_{\boldsymbol{\theta}_1} \sum_{t=1}^{T} \sum_{j=1}^{n} \ln f_j\left(x_{jt}; \boldsymbol{\theta}_1\right)$$

In the second step, given $\widehat{\boldsymbol{\theta}}_1$, we can pass to perform the estimation of the copula parameters vector $\boldsymbol{\theta}_2$:

$$\widehat{\boldsymbol{\theta}}_2 = Arg\ Max_{\boldsymbol{\theta}_2} \sum_{t=1}^{T} \ln c\left(F_1(x_{1t}), F_2(x_{2t}), \dots, F_n(x_{nt}); \boldsymbol{\theta}_2, \widehat{\boldsymbol{\theta}}_1\right)$$

This method is called Inference for the Margins or IFM. The IFM estimator is defined as the vector: $\widehat{\boldsymbol{\theta}}_{IFM} = \left(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2\right)'$. Alternatively, we could say that Inference for Margins estimator is the solution of:

$$\left(\frac{\partial l_1}{\partial \theta_{11}}, \frac{\partial l_2}{\partial \theta_{12}}, \dots, \frac{\partial l_n}{\partial \theta_{1n}}, \frac{\partial l_c}{\partial \boldsymbol{\theta}_2}\right) = \mathbf{0}'$$

Where l is the entire log-likelihood function, l_j is the log-likelihood of the jth marginal, and l_c the log-likelihood for the copula itself. On the other hand, Maximum Likelihood Estimation comes from solving:

$$\left(\frac{\partial l}{\partial \theta_{11}}, \frac{\partial l}{\partial \theta_{12}}, \dots, \frac{\partial l}{\partial \theta_{1n}}, \frac{\partial l}{\partial \boldsymbol{\theta}_2}\right) = \mathbf{0}'$$

Generally, the two estimators are not equivalent. Since it is much more easier applying Inference for Margins estimator, we think we would prefer it to Maximum Likelihood Estimator. Otherwise, before application, we need proof that IFM is asymptotic efficient with respect to MLE. U. Cherubini, E. Luciano and W. Vecchiato (2004) suggest to compare asymptotic covariance matrix of the two estimators. In IFM we use a set of inference equations to estimate a vector of parameters. In this case each equation is a score function: its left side is the partial derivative of the log-likelihood of each marginal density. As it was proven by H. Joe (1997), like the MLE, the IFM estimator turns out to be, under regular conditions, asymptotically normal:

$$\sqrt{T}\left(\widehat{\boldsymbol{\theta}}_{IFM} - \boldsymbol{\theta}_0\right) \longrightarrow N\left(0, Ģ^{-1}(\boldsymbol{\theta}_0)\right)$$

With $Ģ(\boldsymbol{\theta}_0)$ the Godambe information matrix.

Thus, if we define a score function

$$s(\boldsymbol{\theta}) = \left(\frac{\partial l_1}{\partial \theta_{11}}, \frac{\partial l_2}{\partial \theta_{12}}, \dots, \frac{\partial l_n}{\partial \theta_{1n}}, \frac{\partial l_c}{\partial \boldsymbol{\theta}_2}\right)'$$

We split log-likelihood in two parts $l_1, l_2, \dots, l_n$ for each margin and l_c for the copula. Godambe information matrix takes the following form:

$$Ģ(\boldsymbol{\theta}_0) = D^{-1}V(D^{-1})'$$

With $D = E\left[\frac{\partial s(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]$ and $V = E[s(\boldsymbol{\theta})s(\boldsymbol{\theta})']$.

After covariance matrix estimation, H. Joe (1997) assures that Inference for Margins method is highly efficient with respect to Maximum Likelihood Method.

6.PRACTICAL APPLICATION

In previous pages we have described what a copula is and how it can be modelled. Moreover, we noticed that different types of copulas exist, with their particular shape, behaviour and tail characteristics. We have even underlined that these differences allow us to fit empirical data to optimal copula, meaning the copula that best reflects data behaviour, especially behaviour in the tails, and carve out a nice analysis. Now, we want to apply what we learnt and transform theory into practice. We are to take a set of empirical data: price time series of four financial traded indices. We want to fit these data to various copulas, both implicit and Archimedean. When estimating copula parameters, we will use Inference for Margins method: firstly we will estimate conditional distributions parameters of each price time series. Secondly, we will estimate copula parameters. In the last chapter, we will avail of copula method in order to deduce Value at Risk for an imaginary portfolio, composed of our four financial indices. For a more refined analysis, we will vary the weights in the portfolio. In the end, we will compare Values at Risk obtained by portfolios with the same weights, but applying different copulas. Theoretically, Values at Risk will show differences according to copula behaviour and tail dependence.

6.1.First step: data collection and observation

We are to analyse four financial indices, traded in stock markets. They are: FTSE MIB, Italian, CDAX, German, CAC All-Tradable, French, and IBEX35, Spanish.

For simplicity, from now we will rename CAC All-Tradable with the simpler notation of CACT. From Eikon Reuters-Datastream, we download price time series for each index. We are to use weekly data. All time series are referred to a 20 years-time span: from December 1997 to August 2017, resulting in something like one thousand of observations for each index. All the analysis will be conducted on the statistic software R. We are to start by loading on R our data in the form of four data frames and, then, we will convert them into four time series.

Before proceeding into the analysis, we think it would be nice getting a glimpse at price evolutions during our time span. They are shown in figure 18.a, 18.b, 18.c and 18.d.

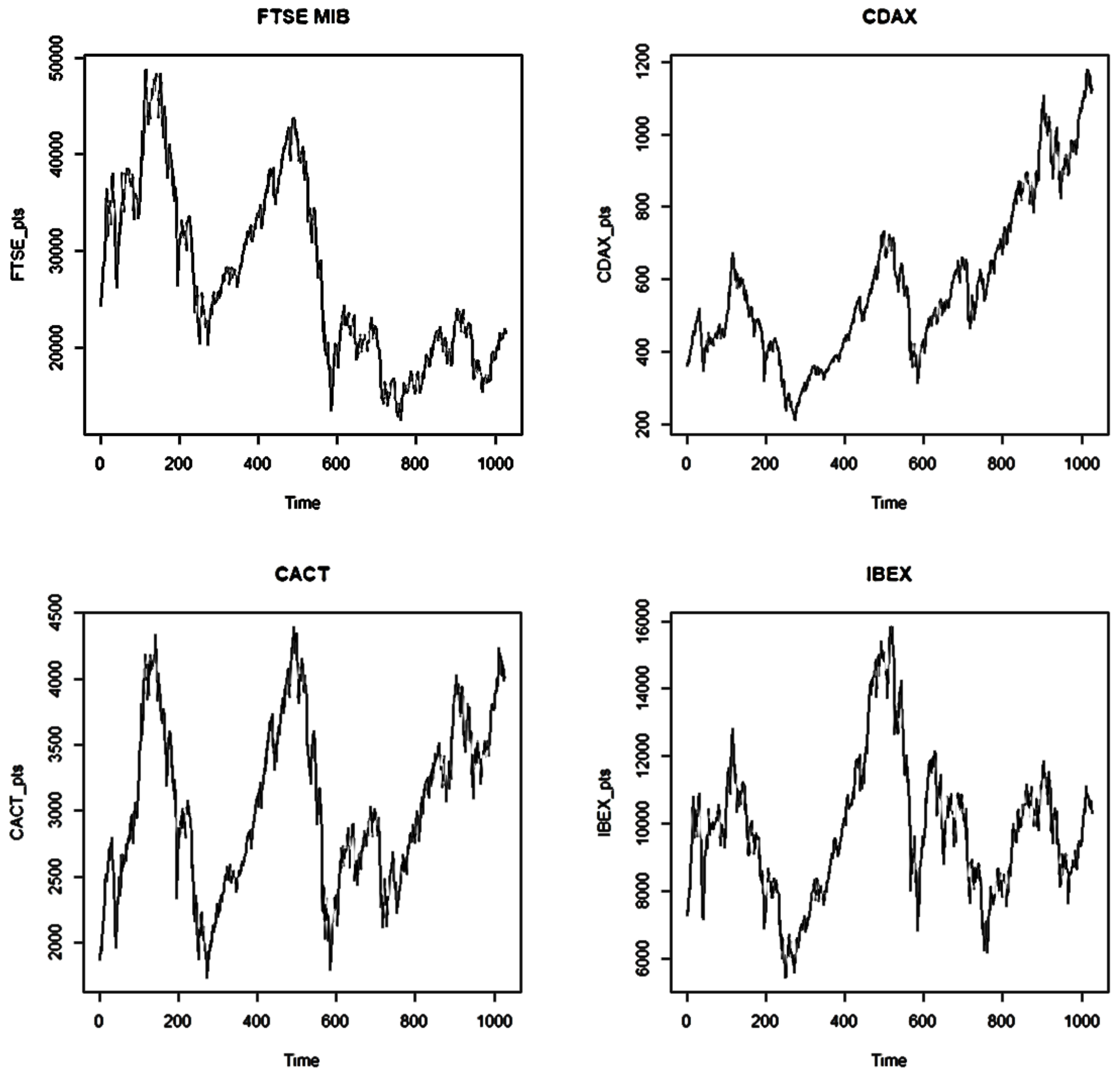


Figure 18

However, we are not to analyse prices, but returns. The reason for this is exposed by K. Aas and X.K. Dimakos (2004): if we try to directly analyse financial prices, we encounter many difficulties, as consecutive prices are highly correlated, and the variance of prices often increases with time. It is much more convenient to analyse changes in prices. According to P. Jorion (1997), we can choose between two main type of price changes: arithmetic or geometric returns. He reminds us that the formula for arithmetic return is:

$$\frac{P_t}{P_{t-1}} - 1 \qquad with\ \frac{P_t}{P_{t-1}} \sim \exp\{N(0, \sigma_r^2)\}$$

Where P_t stays for price in time t, while P_{t-1} is price in time t-1. Geometric returns are instead defined by:

$$\ln\left[\frac{P_t}{P_{t-1}}\right] \sim N(0, \sigma_r^2)$$

There is substantially one advantage of working with the log-scale: if geometric returns are normally distributed, prices will never be negative. In contrast, assuming that arithmetic returns are normally distributed may lead to negative prices, which is economically meaningless. According to this notation of K. Aas and X.K. Dimakos (2004) we elaborate geometric returns: firstly we apply logarithmic transformation to all our time series data. As a consequence, data width is much more restricted because of logarithm.

Then, we derive returns as the difference between logarithmic price recorded in time t and logarithmic price recorded in the previous period. In the following figures we can get a glimpse at graphical representation of geometric returns. As it was expected, every return time series evolves around an expected value of zero.

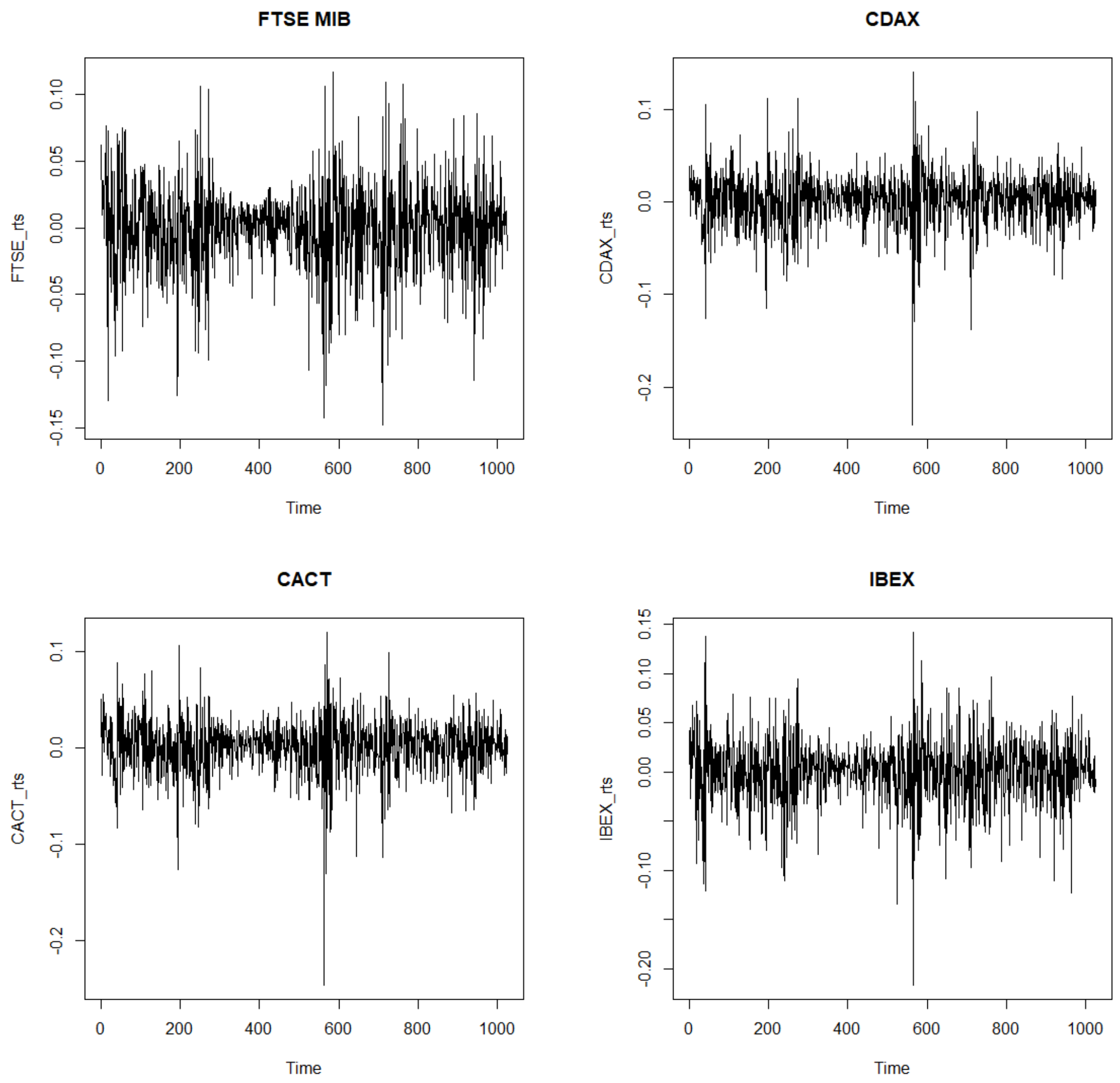


Figure 19

Before proceeding any further with the analysis, we dedicate to same descriptive statistics. For each financial index return time series we calculate mean, median, standard deviation (Sd), variance, asymmetry and kurtosis. The results are exposed below:

	FTSE MIB	CDAX	CACT	IBEX
Mean	-0.0001232285	0.001107212	0.0007440668	0.0003372297
Median	0.00195045	0.004107951	0.003105453	0.002378105
Sd	0.03349165	0.03074855	0.02901188	0.03382556
Variance	0.001121691	0.0009454734	0.0008416894	0.001144169
Skew	-0.4030825	-0.7565778	-0.8943171	-0.4706912
Kurtosis	4.767812	8.309947	9.339892	5.946987

As we have already seen from the graphs, indices returns seem to be mean reverting and they should float around an expected value of zero. According to this, all means calculated in the previous table have small values, very next to zero. There is another thing we would like to take into account: kurtosis. Kurtosis is referred to the shape of a distribution, and constitutes a measure of tail thickness of a density function. Specifically, kurtosis coefficient measures how much our distribution seems to be far from a Normal distribution. If the coefficient is bigger than zero, our distribution is defined leptokurtic, and it is sharper, more poignant than a Normal distribution. If the coefficient is equal to zero, our distribution is as flat as a Normal distribution. Lastly, if coefficient is smaller than zero, our distribution is platikurtic: it is flatter than a Normal distribution. In our case, all our four indices present a strong positive kurtosis coefficient. This was to be expected, as time series returns tend to have ticker tails than a Normal distribution. As a consequence, instead of fitting empirical data to a simple Normal distribution, we could use a conditional t-Student distribution for each index. This would allow us to better take into account extreme phenomena, that are registered in negative tail behaviour, on which we have to focus for Value at Risk calculation. Later, we could even adopt a Normal conditional distribution to describe indices returns, with the aim of comparing results dependent of different assumptions.

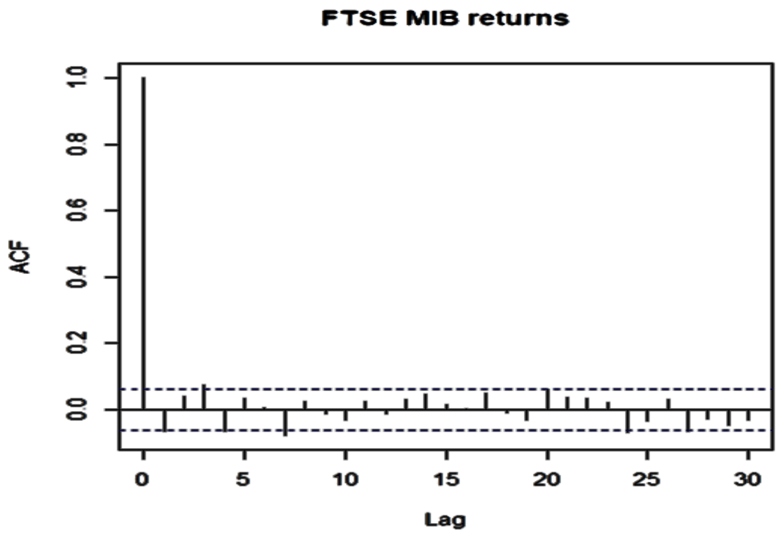
Anyway, having to work with empirical data, kurtosis coefficient by itself is not sufficient to assure us that conditional Normal distribution is not a good choice. In order to have an additional proof that we would do better adopting a conditional Student t-distribution, we will perform even a Jarque-Bera normality test. We will assume a significance value of 0.05. Null hypothesis is Normal distribution. Here are the results:

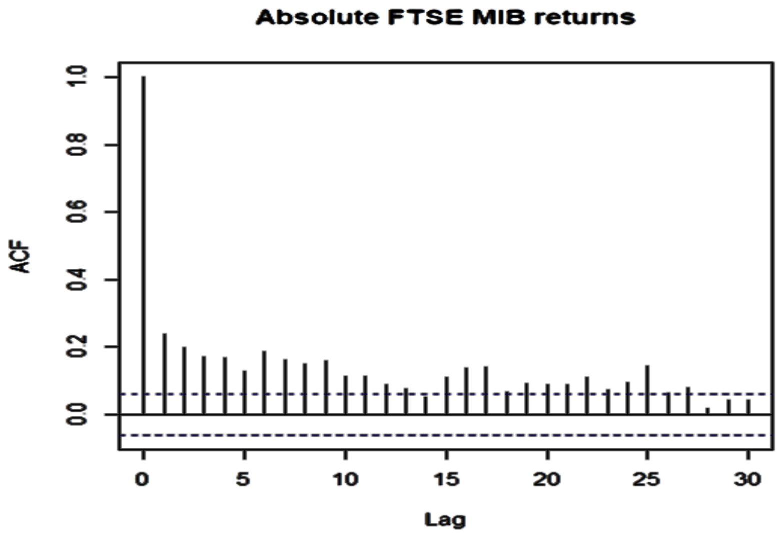
FTSE MIB	p-value = 0
CDAX	p-value = 0
CACT	p-value = 0
IBEX	p-value = 0

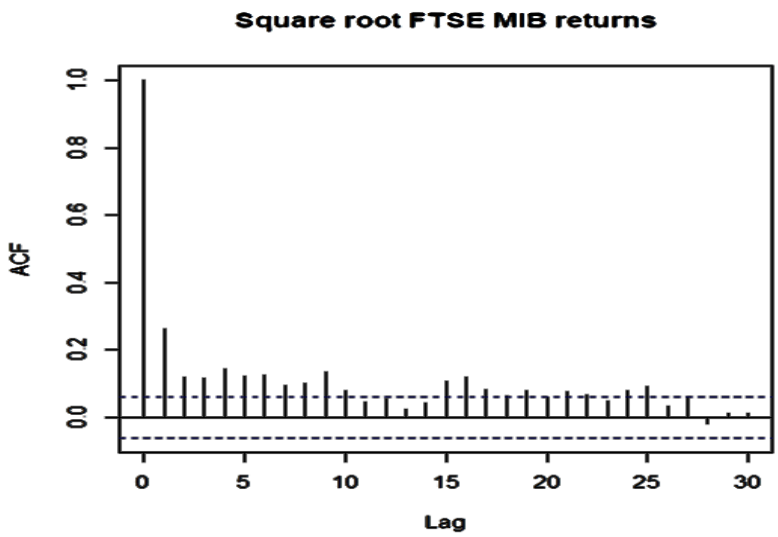
In all cases we refuse null hypothesis of normality in distribution: we may adopt a conditional Student t-distribution for each index considered. If possible, we would even add an asymmetry option, in order to take into account negative asymmetry that appear in all indices descriptive statics.

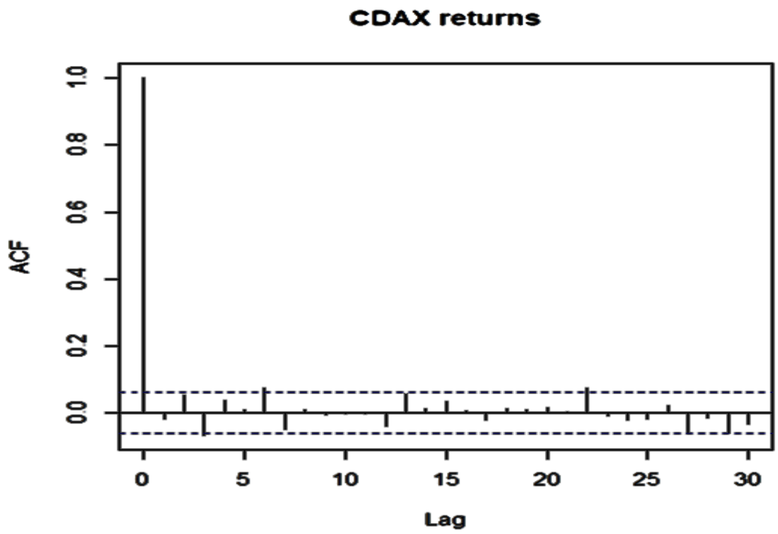
6.2.Second step: handling autocorrelation by GARCH

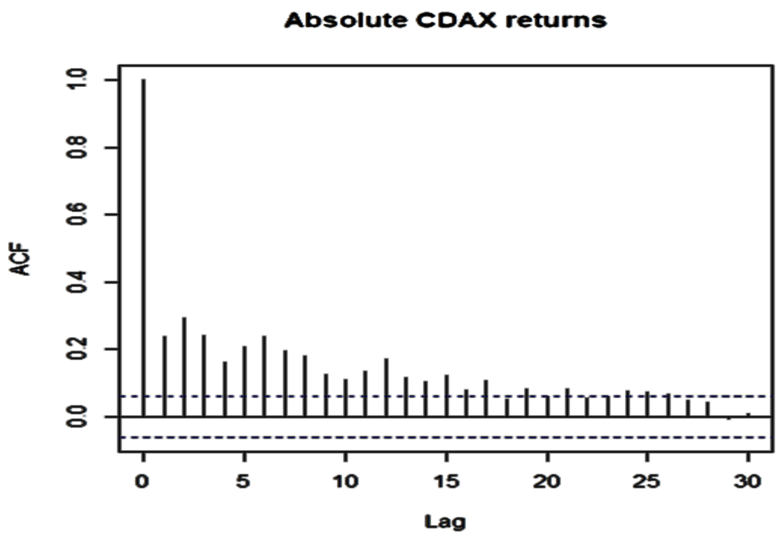
Before proceeding further with the analysis, we want to verify if our data present autocorrelation. In case we detect autocorrelation, there would be some additional consideration to be made. We are to check not only autocorrelation in return time series by themselves, but even in return time series in absolute value and at the square root. Firstly we will show some graphs, in order to get a rapid idea, and then we will perform Ljung-Box test for autocorrelation. Here are the autocorrelograms:

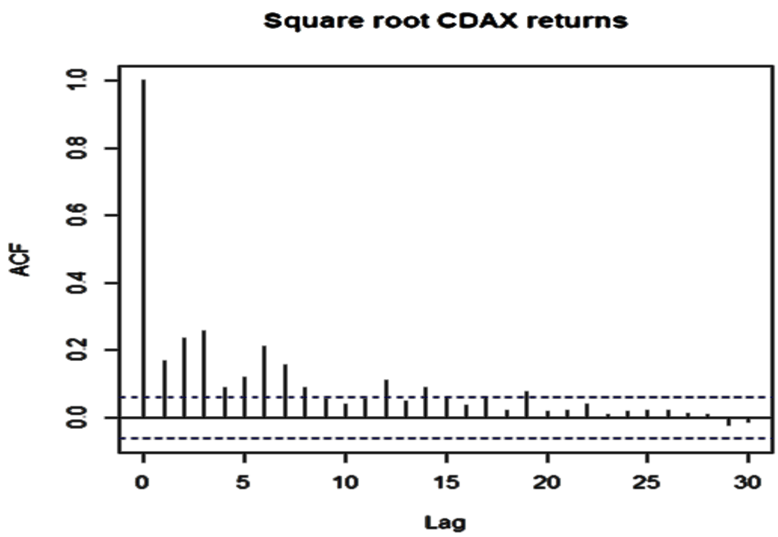


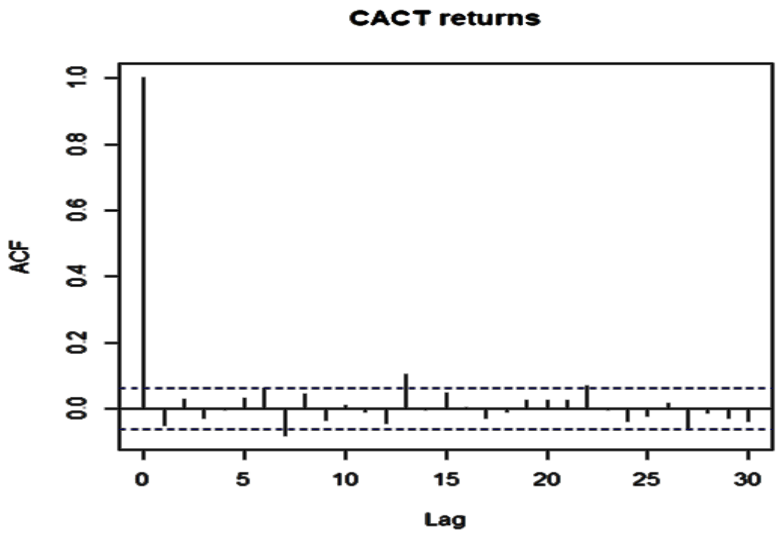


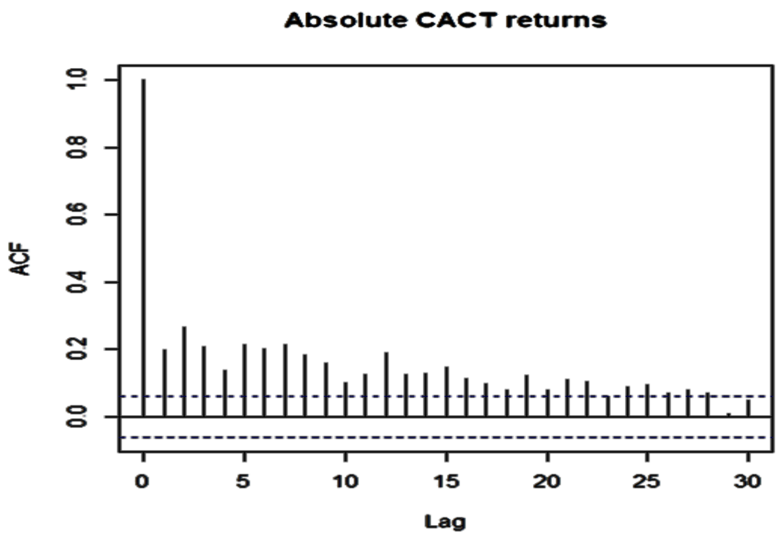

Square root FTSE MIB returns
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

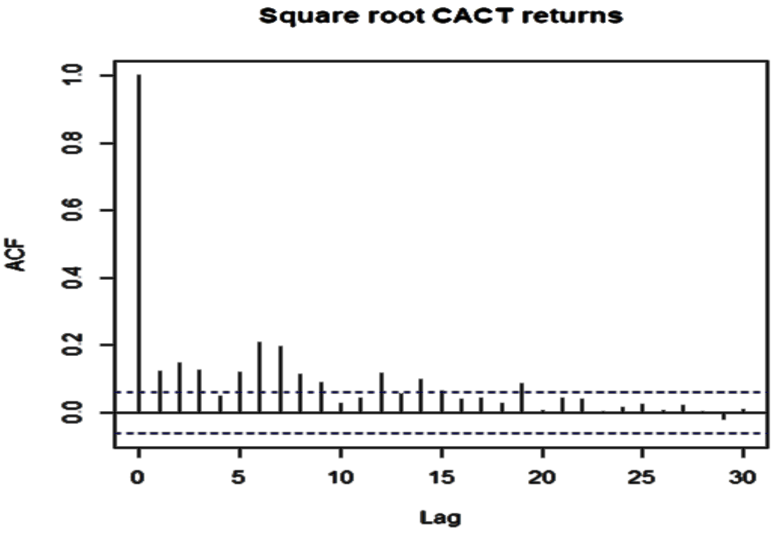

CDAX returns
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

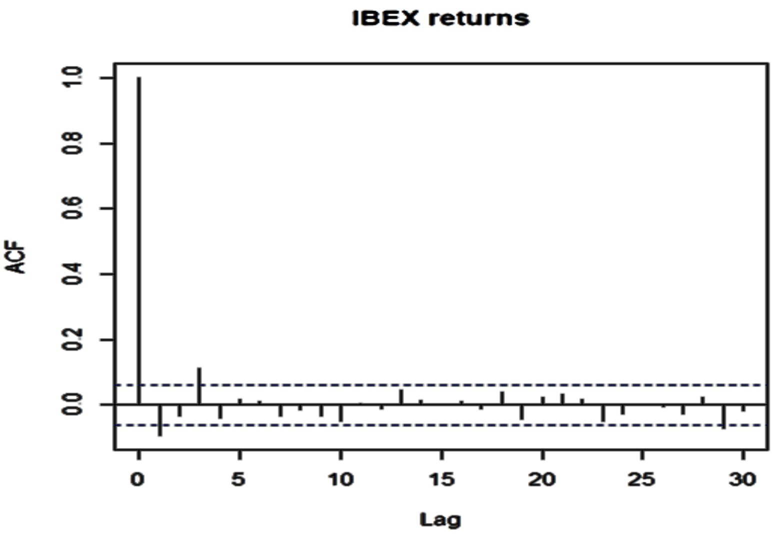

Absolute CDAX returns
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

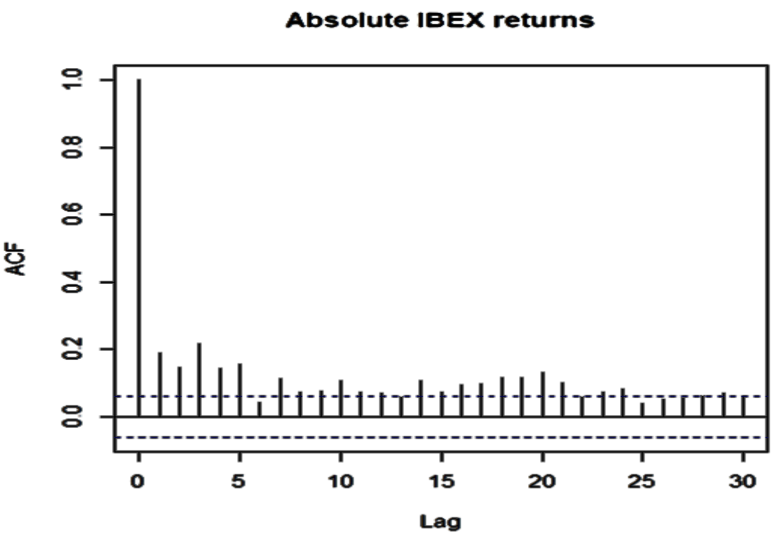

Square root CDAX returns
ACF
Lag


CACT returns
ACF
Lag


Absolute CACT returns
ACF
Lag


Square root CACT returns
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


IBEX returns
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


Absolute IBEX returns
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

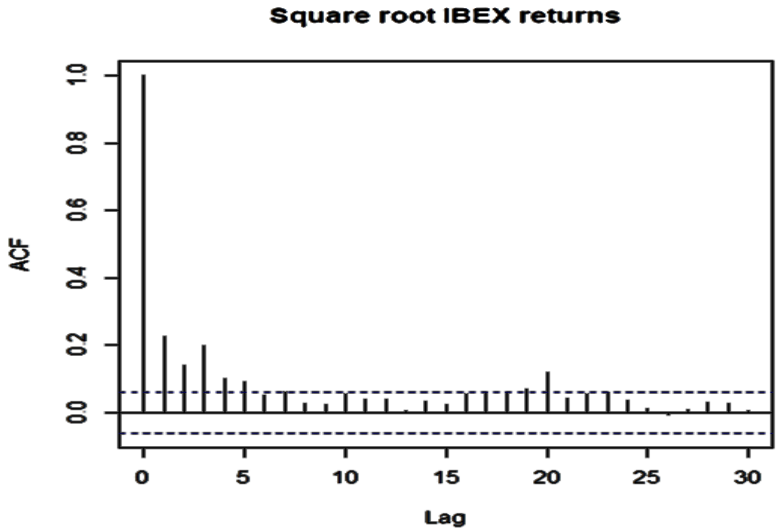


Figure 20

As we can infer graphically, it seems there is no strong autocorrelation in simple time series returns, with the exception of IBEX, that presents autocorrelation not only in the first lag, that is normal, but even in second and fourth lags. As, from both theory and empirical literature, we would expect returns autocorrelograms not to show significant lags but the first, we could wander whether IBEX behaves someway unusually. A possible explanation could be due to a contamination of the first graph of returns by the ones of absolute returns and of square root returns. As it is to be expected that every index absolute and square root returns present strong autocorrelation even in other lags than the first, it could be that IBEX returns have been a bit stained by other data.

As an additional exam, we even perform Ljung-Box test for autocorrelation. We are going to execute it on time series returns, on absolute value time series returns and on time series returns at square root. According to the null hypothesis, data are not autocorrelated. As usual, we take a p-value of 0.05. In the following table we report p-value for every test:

	Returns	Absolute Returns	Square Root Returns
FTSE MIB	p-value = 0.03607	p-value = 0	p-value =0
CDAX	p-value = 0.5569	p-value = 0	p-value = 0
CACT	p-value = 0.1109	p-value = 0	p-value = 0.0001185
IBEX	p-value = 0.003236	p-value = 0	p-value = 0

While we expected to refuse null hypothesis of no autocorrelation for both absolute returns and square root returns, we cannot ignore autocorrelation in IBEX returns. Our hope is that, even in Ljung- Box test, return results have been drawn by strongly autocorrelated absolute and square root returns. In order to prove that IBEX behaves like other indices and we can proceed with our analysis, we will try to fit IBEX returns dataset to an autoregressive model of order four. If all autoregressive estimated coefficients turn out to be non-significant, we will deduce that IBEX returns are not influenced by previous four lags. Checking all outcomes of Ljung- Box test, we will fit an autoregressive model of order four even to FTSE MIB, as a p-value of 0.03607 could seem borderline between refusing or not null hypothesis. We will integrate autoregressive model with GARCH model, in order to take into account everything together. However, we are not to explain GARCH immediately, as, before, we just want to make it clear why autocorrelation is so important for us.

It is clear that we cannot ignore autocorrelation in absolute returns and in returns at square root. To understand exactly what does this autocorrelation mean and why we really need to cope with it for the scope of our analysis, we refer to an article of P. Posedel (2005). As we could already know, she reminds that financial markets react nervously to stress, independently of the reason of the shock: political, economic, natural… During stress periods, prices of financial assets tend to fluctuate much more than normally. Statistically, this means that we have heteroscedasticity: among random variables there are sub-populations that have different variance from others. Posedel writes that prices have been always believed to be non-stationary. As a consequence, till 1940's, economists resorted to log-returns, that were supposed to be stationary instead, at least in periods of time that were not too long. Log-returns were referred to as if they represented a sequence of independent, identically distributed random variables. It was thought log-returns evolve like a random walk and that they could have been modelled in continuous time by a geometric Brownian motion. Discretization of such a model leads to a random walk with independent, identically distributed Gaussian log-returns in discrete time. However, this hypothesis was rejected in the 1960's, thanks to some empirical studies based on the log-return time series data of US stocks. They demonstrated that serial dependence is present in the data and that volatility changes in time. The latter point means that we have volatility clustering and, as we said previously, as volatility changes data present heteroscedasticity. Moreover, the same empirical studies demonstrated that distribution of the data is heavy tailed, asymmetric and, so, tricky to describe with a Gaussian. All those considerations seem to coincide with what we could infer from our data: even our log-return series are heavy-tailed and seem not to correspond so much to a Gaussian distribution. As we have already said, we could try to fit each of our time series to a Student t-distribution. However, for taking heteroscedasticity into account, we will avail ourselves of a discrete model found by R. Engle, that is both meticulous in description, practical to use and stationary, so the inference is possible. R. Engle calls this model ARCH, that stays for Autoregressive Conditional Heteroskedastic, because it takes into account that conditional variance is not constant over time and shows an autoregressive structure, due to the clustering. Some years later, T.P. Bollerslev generalized the model and introduced the Generalized Autoregressive Conditionally Heteroskedastic model, the GARCH, that we are going to use. Generally, we can say ARCH model is appropriate when the error variance in a time series follows an autoregressive AR model. If we assume, for the error variance, an autoregressive moving average model instead, we can adopt GARCH.

Now we are to give definition of a general GARCH(p,q) model, where p is the order of the GARCH terms $\sigma^2$and q is the order of the GARCH terms ε^2. To begin with, we state that our returns are defined as:

$$r_t = \alpha_0 + \varepsilon_t$$

Where r_t is return in time t, α_0 is a constant, whose expected value is zero, as our returns are mean reverting, and ε_t is the error term. Returns are distributed like a t-Student, with an expected value of zero. For what concerns variance of the returns, things get a bit more complicated: the variance of the returns is the variance of the error term and it is described by GARCH(p,q) as:

$$\sigma_t^2(\varepsilon_t) = \beta_0 + \gamma_1 \varepsilon_{t-1}^2 + \cdots + \gamma_q \varepsilon_{t-q}^2 + \delta_1 \sigma_{t-1}^2 + \cdots + \delta_p \sigma_{t-p}^2$$

Where ε_t is distributed as a t-Student, with the same degrees of freedom of r_t. ε_t has mean equal to zero and variance equal to σ_t^2. As in our case α_0 is equal to zero, because time series of index log-returns are mean reverting and expected value is zero, we could even write:

$$\sigma_t^2(\varepsilon_t) = \beta_0 + \gamma_1 r_{t-1}^2 + \cdots + \gamma_q r_{t-q}^2 + \delta_1 \sigma_{t-1}^2 + \cdots + \delta_p \sigma_{t-p}^2$$

In order to treat our data for heteroscedasticity, we think a GARCH(1,1) model would be sufficient. It should be done like this:

$$\sigma_t^2(\varepsilon_t) = \beta_0 + \gamma_1 \varepsilon_{t-1}^2 + \delta_1 \sigma_{t-1}^2$$

As we need to correct all our data for GARCH(1,1), we can create a general univariate GARCH model specification before, and then fit each of our time series to it. Firstly, we ask for “rugarch” library on R and we resort to command “ugarchspec”. By ugarchspec, we can create our model by specifying variance, mean and shape of every conditional distribution. We want to conduct this marginal estimation separately for each of our return time series. We indicate that variance is described by a GARCH(1,1), and mean by an ARMA model. While we are to adopt a GARCH(1,1) for all our four indices, the choice of ARMA depends on what we observed previously in the autocorrelograms of simple returns. For what concerns CDAX and CACT we just have to tackle heteroscedasticity in the error term, that we will treat with a GARCH(1,1), as previously said. Apart from this, for these series we will not need to handle autoregressive behaviour of returns. As a consequence, we will apply an autoregressive moving average model, ARMA, of order (0,0). Things get a bit more complicated for FTSE MIB and IBEX series of returns. As we spotted from both the graphs of autocorrelations and Ljung- Box tests, there seems to be autocorrelation both in FTSE MIB and IBEX returns until lag four. Before we suggested the hypothesis of a possible contamination of returns by absolute value returns and by returns at square root. We want to verify whether this supposition is true: if it really is, we will be able to treat these series as the others. We will start our check from FTSE MIB. We are to construct a model with GARCH(1,1), similar to the one we spoke about before. However, for FTSE MIB and IBEX, we are to add an ARMA(4,0) model, instead of an ARMA(0,0). According to this, software R won't only estimate GARCH parameters, but even ARMA ones. If autocorrelation is to be detected into returns, corresponding estimated parameters will be significant. Although, whether we will be likely to not refuse null hypothesis, that parameters are not significant, we will finally adopt a GARCH(1,1) ARMA(0,0) model even for FTSE MIB and IBEX, as we did for CDAX and CACT. We estimate ARMA(4,0)- GARCH(1,1) FTSE MIB and IBEX parameters in Appendix A.

As we have proved, in Appendix A, that FTSE MIB does not need an ARMA model different from (0,0), we could exploit just a GARCH(1,1) ARMA(0,0) model, whose codes are exposed in the following:

```
gspec.ru.std<-
ugarchspec(variance.model=list(model="sGARCH",garchOrder=c(1,1)),
mean.model=list(armaOrder=c(0,0)),distribution.model="std")
FTSEgarch.std<-ugarchfit(gspec.ru.std,FTSE_rts)
FTSEgarch.std
```

	Estimate	Std. Error	t value	Pr(>|t|)
Mu	0.002183	0.000778	2.8039	0.005048
omega	0.000015	0.000007	2.2264	0.025990
alpha1	0.128956	0.027331	4.7183	0.000002
beta1	0.867858	0.023854	36.3824	0.000000
shape	6.636685	1.394077	4.7606	0.000002

Here we start by fitting FTSE MIB observations by ugarchspec to a Student t-conditional distribution option, without taking asymmetry into account. Mind that there is no need to indicate degrees of freedom, as the program will estimate them: they are returned by the parameter "shape". According to what we observed before, we should adopt a "std" or a "sstd" option in order to describe all our conditional distributions of returns. Although, we decide to check all the four options for each index, in order to compare results. We will have the possibility to choose conditional distribution that we find more convincing and control whether it gets along with what we observed before: whether, empirically, a Student t-distribution would suit data better. We are to fit our FTSE MIB returns time series to a GARCH(1,1)- ARMA(0,0) model, with all possible options of shape. We expose all FTSE MIB estimated parameters in the following table. In each cell, we put standard error under the estimate. Here are the parameters:

FTSE MIB

	Mu	Omega	alpha1	beta1	shape	skew
Non- asymmetric	0.002183	0.000015	0.128956	0.867858	6.636685	
t- Student	0.000778	0.000007	0.027331	0.023854	1.394077	
Asymmetric	0.000996	0.000017	0.141791	0.849844	8.433522	-0.208433
t- Student	0.000787	0.000007	0.027968	0.026563	2.183227	0.034870
Non- asymmetric	0.001246	0.000019	0.136018	0.855609		
Normal	0.000799	0.000009	0.025229	0.025237		
Asymmetric	0.000812	0.000018	0.147333	0.842713		-0.236278
Normal	0.000773	0.000007	0.025123	0.024987		0.000000

Omega, alpha1 and beta1 are the coefficients in the variance of the error term:

$$\sigma_t^2(\varepsilon_t) = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

All these coefficients are significant. However, what we are more interested into is degrees of freedom estimation: here, GARCH(1,1)- ARMA(0,0) estimates 7 degrees of freedom for t-Student option with no asymmetry, and this estimate is significant. We want to underline that this result is quite important for us, as it justifies Student t-distribution as a correct choice to depict conditional behaviour of FTSE MIB. If we pass to second row of estimated parameters, we can see that, for what concerns omega, alpha1 and beta1 estimates, they are extremely similar in the two cases: with and without asymmetry. We can so compare the models by focusing on parameter shape, that indicates degrees of freedom. Shape is significant and equal to 8 here, not so far from 7 we got by Student t-hypothesis with no asymmetry before. Moreover, we have an additional parameter now, that is in fact skewness estimation: -0.21. Being this last parameter significant, we can infer it would be nice to take even asymmetry into account in FTSE MIB returns distribution. From what we could deduce till now, it seems legit to adopt the hypothesis of the asymmetric Student t-distribution. As a consequence, we will assume 8 degrees of freedom for FTSE MIB, that is the unique parameter we need to give to R in building copulas. Anyway, we decide to take 8 just because it is coherent with model choice: if we had assumed just one degree of freedom less, copula estimate would not have changed for that.

Till now, everything seems to be quite linear, and there appears to be no need to fit data even to normality option, neither to asymmetric normality. However, we will try even that possibilities, just in order to see if error term coefficients are similarly estimated. If so, we can deduce that the unique discriminants are degrees of freedom and asymmetry. We want to check if these two elements demonstrate being significant every time tested.

From the third row of estimated parameters in the table of FTSE MIB, we check that omega, alpha1 and beta1 are very similar to the corresponding parameters estimates we did under t-Student assumption, and always significant. We are not to extract one value or the other, as we will not need them in copula construction. However, we wanted to verify whether all these parameters were similarly estimated, independently of conditional distribution choice. As a consequence, we can discriminate by two last parameters: degrees of freedom and asymmetry. Obviously no normal distribution encompasses degrees of freedom, but the asymmetric one has a skew parameter. Being both parameters always significant in our estimates, we are to choose the conditional FTSE MIB distribution that allows both, so, an asymmetric Student-t distribution.

Now, we will show rapidly the corresponding tables for the other three indices, based on all available ugarchspec options. We start with CDAX index:

CDAX

	mu	omega	alpha1	beta1	shape	skew
Non- asymmetric	0.003393	0.000048	0.158009	0.792104	8.828732	
t- Student	0.000758	0.000018	0.039073	0.048652	1.913535	
Asymmetric	0.002566	0.000043	0.143073	0.806081	9.994852	-0.237774
t- Student	0.000774	0.000015	0.033132	0.043135	2.457459	0.036243
Non- asymmetric	0.003525	0.000077	0.248341	0.688159		
Normal	0.000744	0.000022	0.045070	0.051397		
Asymmetric	0.003128	0.000058	0.201068	0.736903		-0.26019
Normal	0.000746	0.000017	0.039252	0.047911		0.032277

As we checked before for FTSE MIB, omega, alpha1 and beta1 CDAX estimates are always significant and display very similar values. Finally, even for degrees of freedom and skewness, when encompassed by distributional choice, CDAX estimated parameters are pretty similar to the ones of FTSE MIB. Being both degrees of freedom and skewness significant, we are driven to think adopting a GARCH- ARMA model allowing for both would be a nice match. As a conclusion, an asymmetric t-Student conditional distribution could be our first choice for CDAX, as it was already declared for FTSE MIB. However, even if these options seem the best to fit our data, we will later even try the others in copula construction: non-asymmetric t-Student, normal and asymmetric normal. This is to be done for the sake of exploration: we would like to investigate whether and in which measure does Value at Risk estimation change, based on conditional distribution assumptions.

Now we are to pass to CACT and IBEX index, respectively:

CACT

	mu	omega	alpha1	beta1	shape	skew
Non- asymmetric	0.002493	0.000021	0.117755	0.858205	9.230062	
t- Student	0.000685	0.000012	0.030341	0.038195	2.841978	
Asymmetric	0.001851	0.000021	0.115467	0.857328	10.680108	-0.213481
t- Student	0.000707	0.000009	0.025739	0.031515	2.783579	0.036256
Non- asymmetric	0.002329	0.000025	0.148009	0.831674		
Normal	0.000716	0.000011	0.030104	0.034963		
Asymmetric	0.001977	0.000023	0.129459	0.843473		-0.242361
Normal	0.000704	0.000009	0.025021	0.030618		0.031552

IBEX

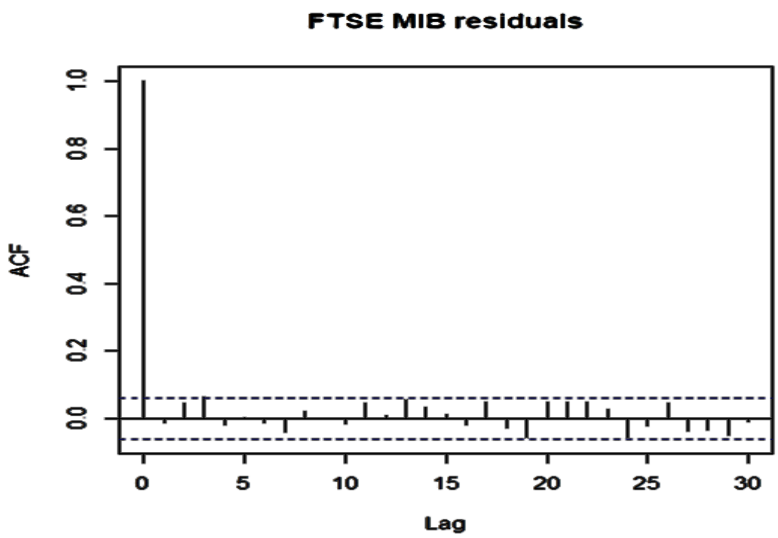
	mu	omega	alpha1	beta1	shape	skew
Non- asymmetric	0.002610	0.000019	0.110307	0.880349	5.984010	
t- Student	0.000819	0.000012	0.027636	0.029824	1.034413	
Asymmetric	0.001555	0.000021	0.113048	0.872798	6.987641	-0.167795
t- Student	0.000847	0.000012	0.027085	0.030659	1.403147	0.037090
Non- asymmetric	0.001543	0.000074	0.146449	0.793014		
Normal	0.000886	0.000030	0.034594	0.053162		
Asymmetric	0.001055	0.000042	0.125826	0.838502		-0.220117
Normal	0.000857	0.000019	0.027681	0.039410		0.032029

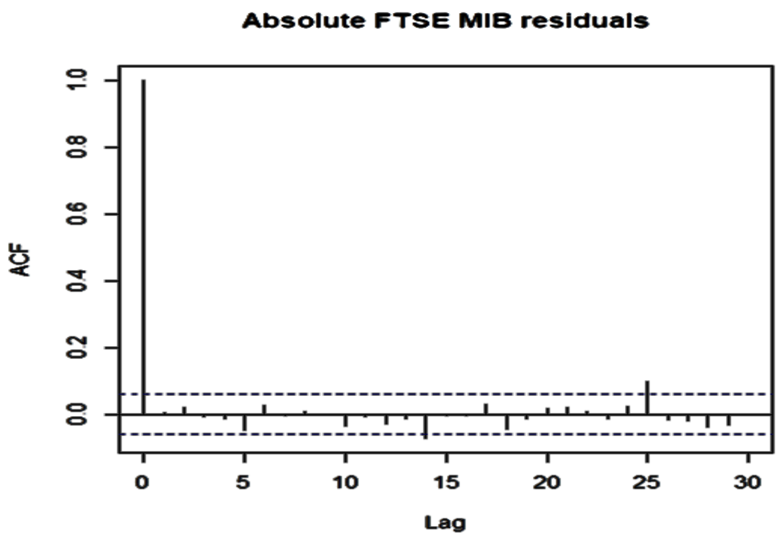
Omega, alpha1 and beta1 are significant and quite similar across all options. However, we are not to indulge in mu, omega, alpha1 or beta1 coefficients any more. For what concerns our analysis, we just wanted to check whether they were similarly estimated, across different marginal options attributed to each index. If all these coefficients were pretty much similar, we could have focused on the other two parameters: degrees of freedom and asymmetry. We wanted to control if both estimates were significant, in order to consider asymmetry and data distribution as a t-Student as two valid possibilities in describing returns behaviour. If we would have found out an insignificant parameter for shape, as an example, we would have thought about quitting t-Student hypothesis for that index. Moreover, we were even much interested in noticing whether shape parameters would have been similar between t-Student and asymmetric t-Student. On the other hand, it was important to control if asymmetric parameter was influenced by normal or t-Student conditional distribution adoption, or would have remained the same. In case we would have noticed differences in degrees of freedom estimation, or asymmetry, we could have thought that those differences were at least partially due to an interconnection between distributional choices: asymmetry or not, for what concerns degrees of freedom, and t-Student or normal, for asymmetry parameter, respectively. This eventual interconnection could have witnessed a capacity of asymmetry to interfere in degrees of freedom estimation, as, in fact, there are only two models that encompass degrees of freedom. They assume a conditional t-Student distribution and, having extremely similar estimated parameters, the unique element that distinguish them is the presence or not of asymmetry. Similarly, there could have been an interconnection between degrees of freedom and asymmetry whether degrees of freedom would have been able to influence skewness. In this case, the same parameter of asymmetry would have changed between two models that present similar parameters, and just a difference in marginal distribution choice. This latter option is completely exhausted by the estimation of degrees of freedom, so the presence or not of degrees of freedom would have influence skewness in the two asymmetric model of ours. According to all what we have said, we can deduce that there seems to have no interconnection between the two: as one parameter estimation does not change, it seems not to be influenced by the presence of the other. As a consequence, we can cross the cases and state that we can assume four different models, according to the presence of both parameters, just one, or even none. Till now, we have said that, for all indices, we are to fall in the case where we have both asymmetry and degrees of freedom, as both estimates are significant. However, in order to estimate Value at Risk, we would like even to try other options, like assuming normality, as an example. We do not want to retreat our empirical deduction, that an asymmetric t-Student should fit all our indices observations well. It is instead for the sake of investigation, as we would like to check whether and in which measure could Value at Risk estimation change, depending on conditional distribution assumption in the model.

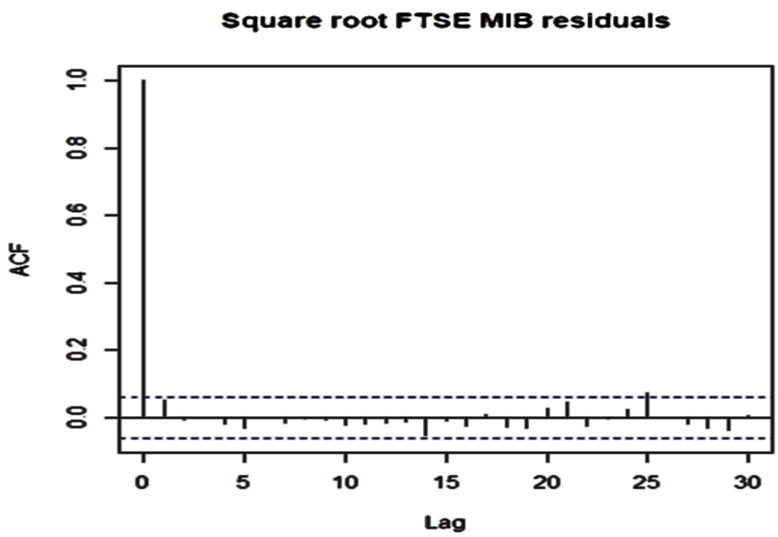
After having taken heteroscedasticity into account, we are to extract, for each index, data representing returns distribution of that index, corrected for heteroscedasticity. According to what we said before, index return is given by two elements: a fixed coefficient that we called α_0 and error term. As we are observing a mean reverting object, α_0 is equal to zero, so we will focus on error term treated for heteroscedasticity, instead. Luckily, R offers an easy code to extract residual series from a GARCH model. We are to perform this code for each conditional distribution option of every index, in order to be able to perform further analysis. Here we expose the four FTSE MIB codes, corresponding to the four options in distribution model. Other indices codes are similarly written.

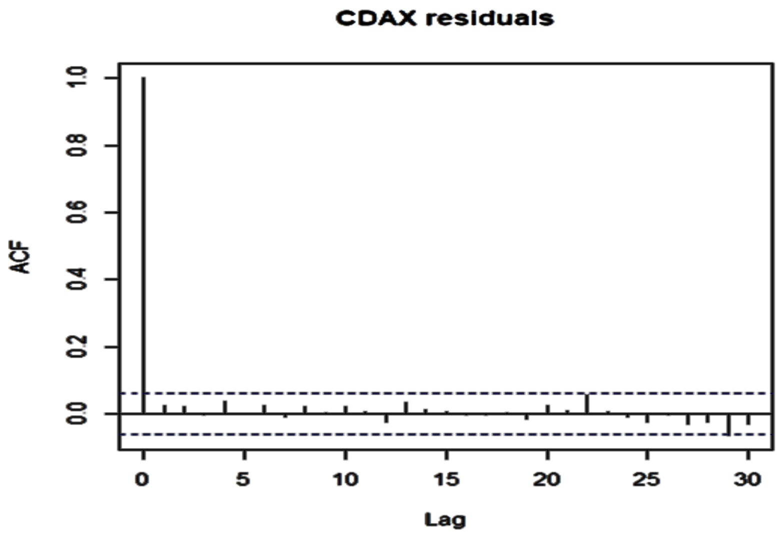
```
FTSEres.std<-residuals(FTSEgarch.std,standardize=T)
FTSEres.sstd<-residuals(FTSEgarch.sstd,standardize=T)
FTSEres.norm<-residuals(FTSEgarch.norm,standardize=T)
FTSEres.snorm<-residuals(FTSEgarch.snorm,standardize=T)
```

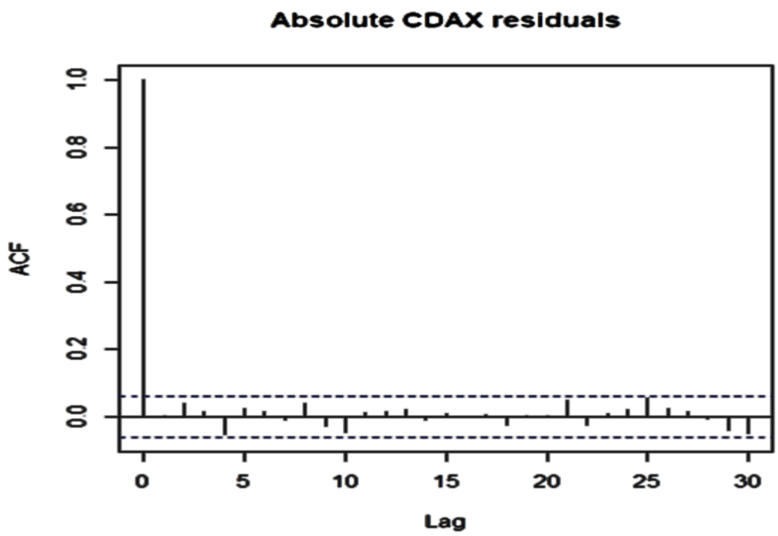
As it is possible to note from codes, residuals distributions have been standardized by the variance of the error term, according to the GARCH(1,1) formula, that we wrote previously. Before proceeding with the analysis, we would like to indulge into two series of graphs. In the first range we reproduce autocorrelograms we previously did for returns, absolute returns and returns at the square root. However, instead of returns, this time we will use standardized residuals, extracted from GARCH model. We want to verify whether there are significant lags but the first. We really expect first autocorrelogram, the one of simple residuals, to not present significant lags but lag zero. However, if we do not detect significant lags after the first into absolute residuals and square root residuals autocorrelograms, we will have the proof that GARCH(1,1) previously employed effectively managed to treat our rough data for heteroscedasticity, and we will be able to continue with our analysis. For rapidity's sake, here we just show autocorrelograms referred to GARCH models with asymmetric t- Student as option, but with different shapes lags won't change.

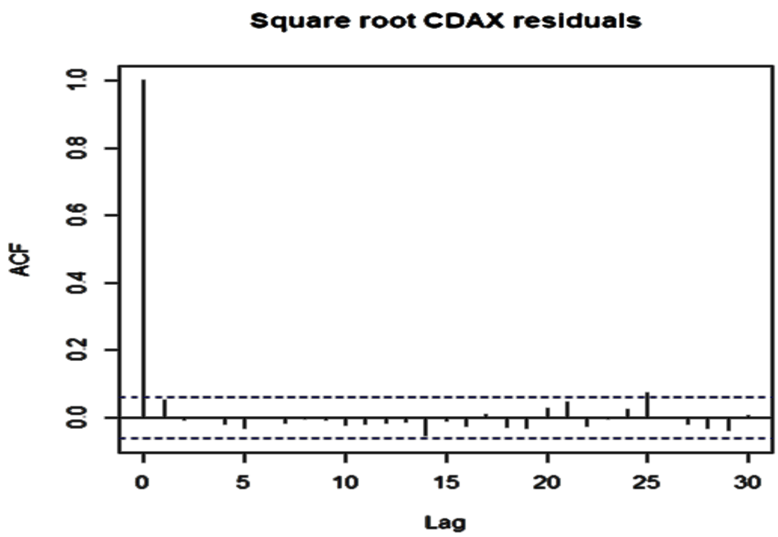


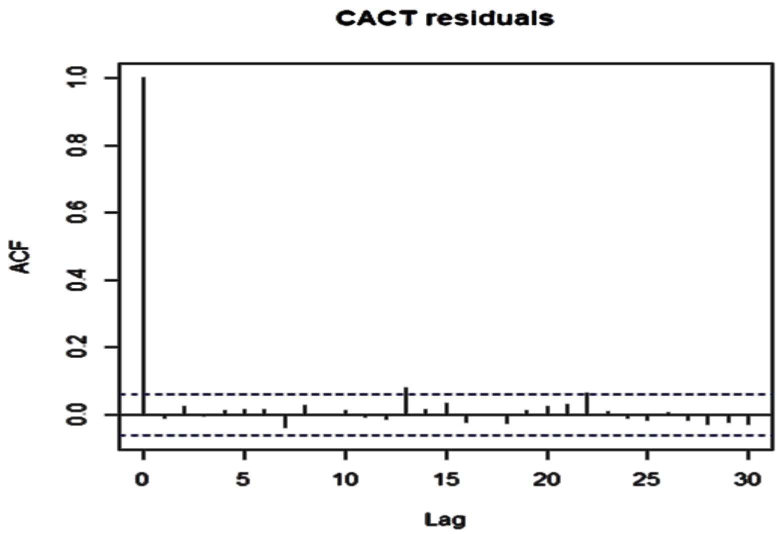


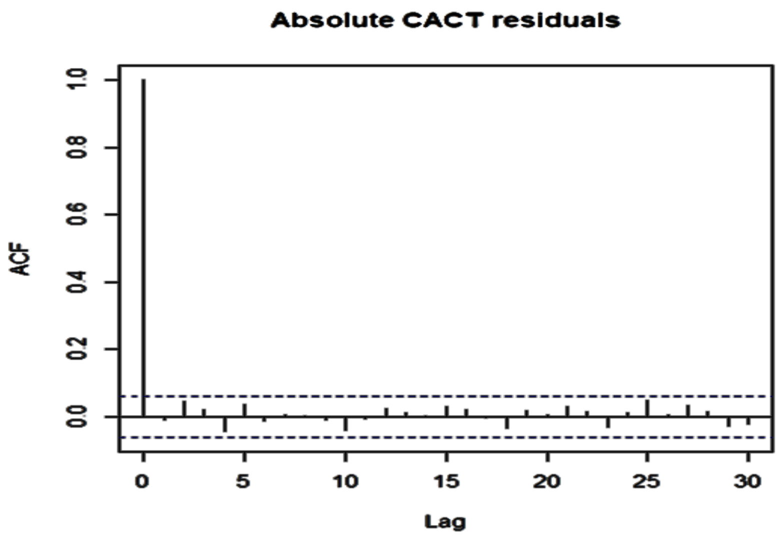

Square root FTSE MIB residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

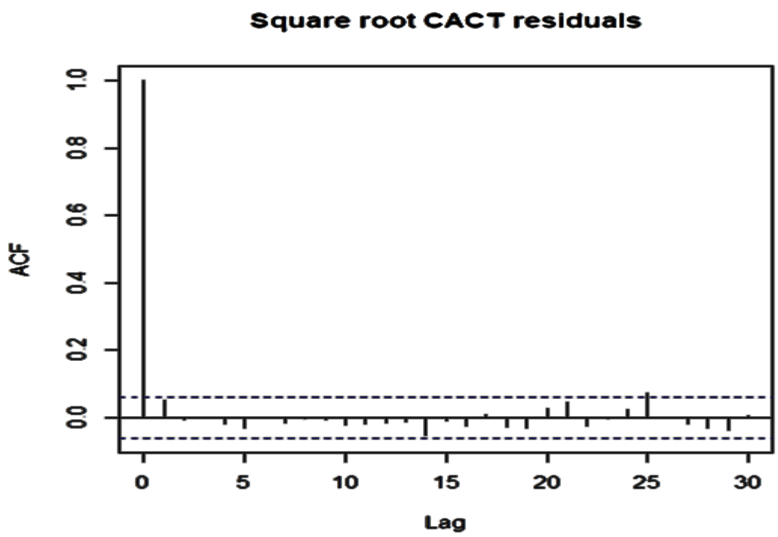

CDAX residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

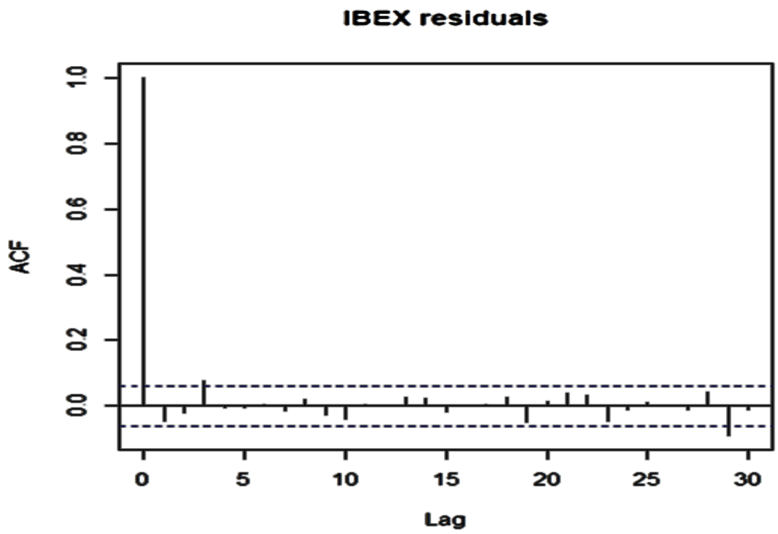

Absolute CDAX residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

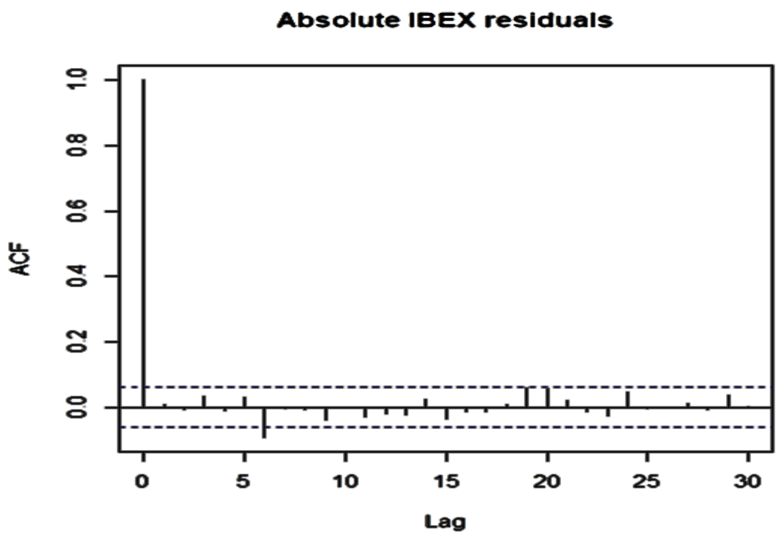

Square root CDAX residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


CACT residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


Absolute CACT residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


Square root CACT residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


IBEX residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag


Absolute IBEX residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
Lag

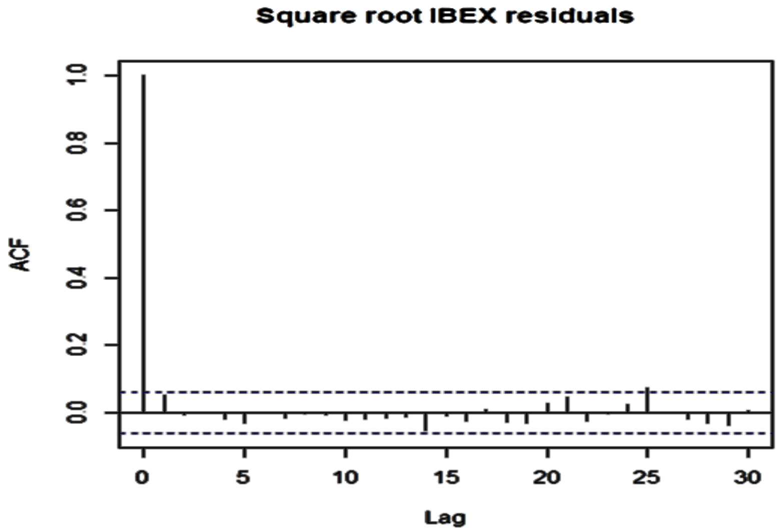


Figure 21

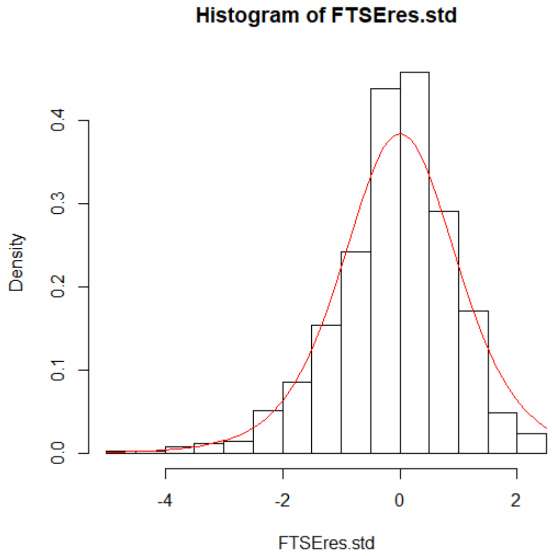
6.3.Graphical representation of GARCH residuals

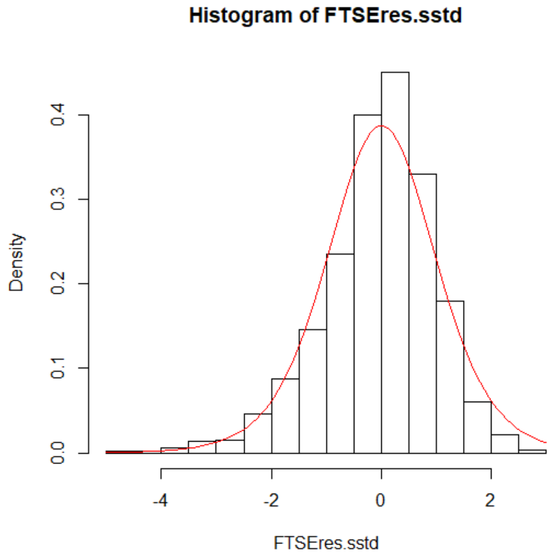
Before passing to copula construction, for each index included in the analysis, we will display four histograms, showing density function of the corresponding standardized residuals. Each index has four graphs, that stay for the different residual distribution assumptions: t-Student, asymmetric t-Student, normal or asymmetric normal. As we have already said before, according to both theory and empirical analysis, an asymmetric t-Student distribution seems the best option to fit every index. However, just for curiosity's sake, we are to follow the same analysis even for all other options. Apart from residuals histograms, we even add in each graph the density function of the corresponding distribution. As an example, the first graph represents FTSE MIB residuals, according to a non-asymmetric Student t-distribution. The curve we added stays for density function of a Student t-distribution, with as many degrees of freedom as GARCH model estimated for FTSE MIB with Student t-option, but not asymmetric. Similarly, in case residuals reflect normality conditional distribution option, the line expresses density function of a normal distribution.

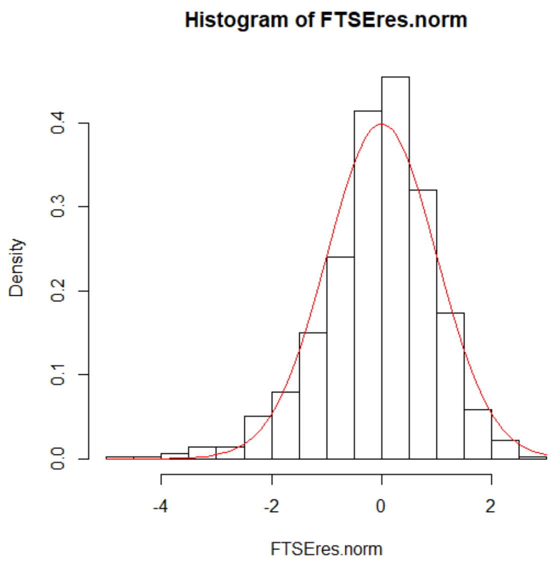
We now expose the codes to create the first picture: the one of FTSE MIB residuals, with t-Student distribution hypothesized:

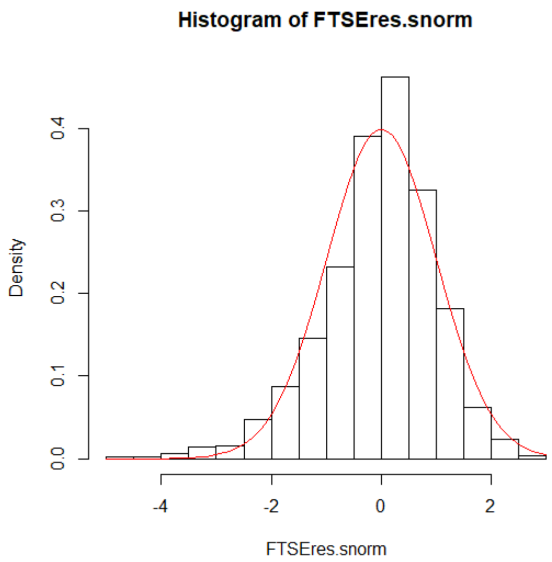
```
hist(FTSEres.std,nclass=20,freq=FALSE)
curve(dt(x,df=6.64),add=T,col="red")
```

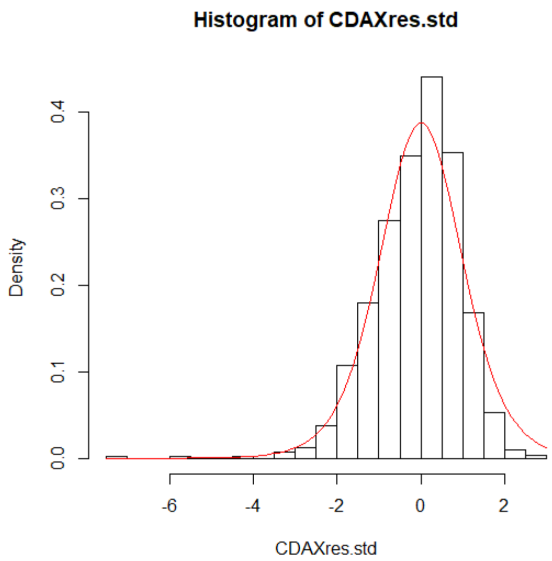
All other graphs are based on this model. Here they are:

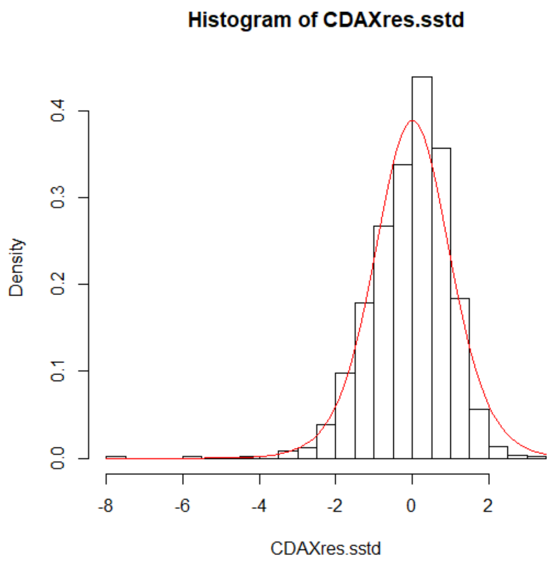

Histogram of FTSEres.std
Density
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
FTSEres.std

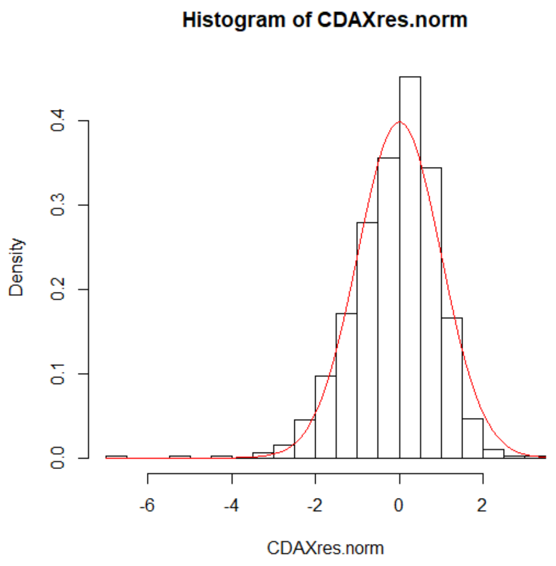

Histogram of FTSEres.sstd
Density
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
FTSEres.sstd

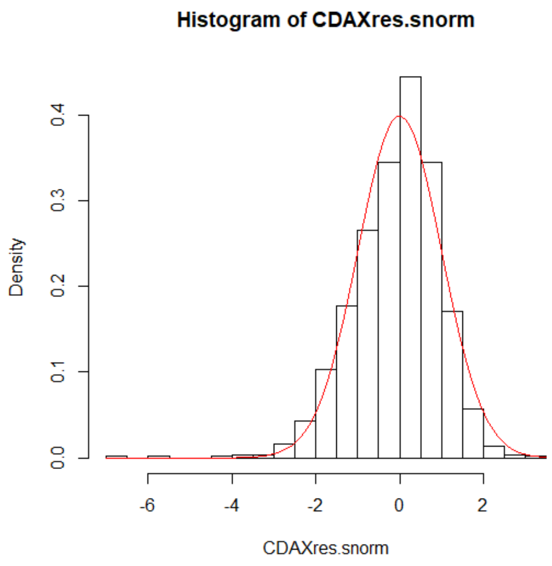

Histogram of FTSEres.norm
Density
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
FTSEres.norm

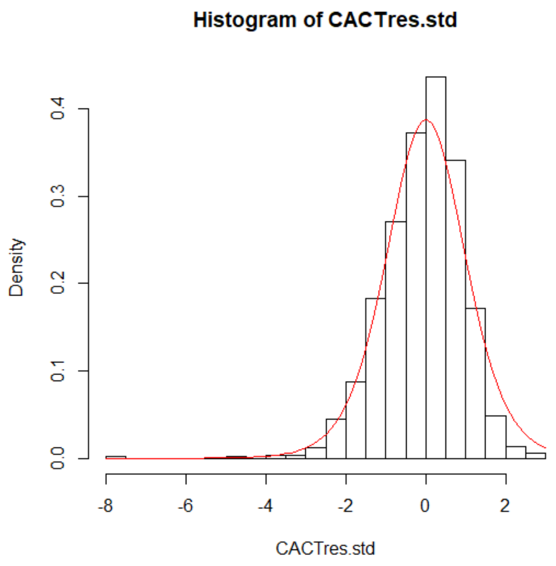

Histogram of FTSEres.snorm
Density
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
FTSEres.snorm

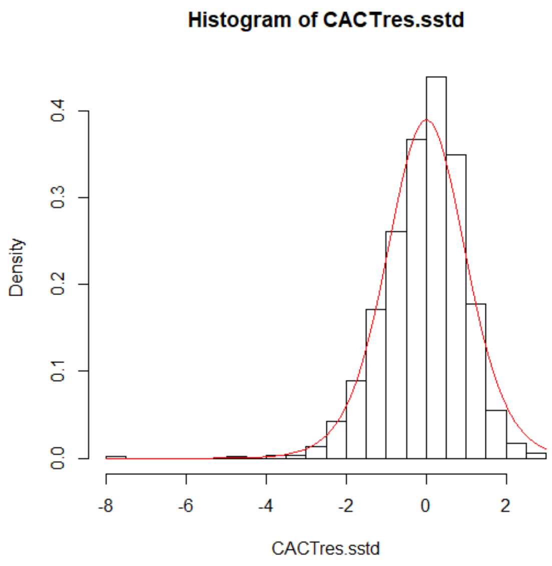

Histogram of CDAXres.std
Density
0.0
0.1
0.2
0.3
0.4
-6
-4
-2
0
2
CDAXres.std

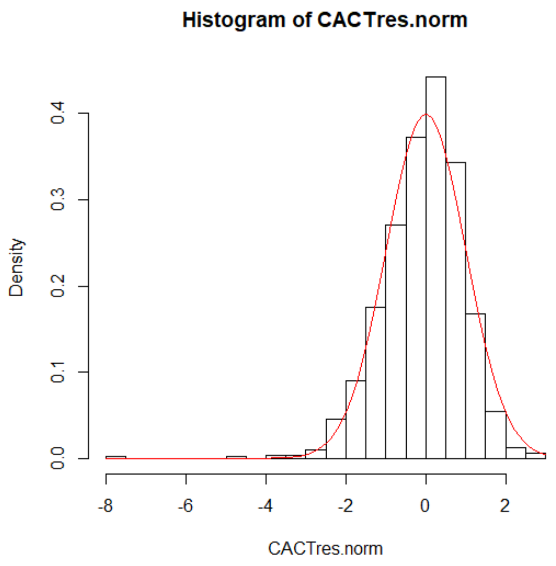

Histogram of CDAXres.sstd
Density
0.0
0.1
0.2
0.3
0.4
-8
-6
-4
-2
0
2
CDAXres.sstd

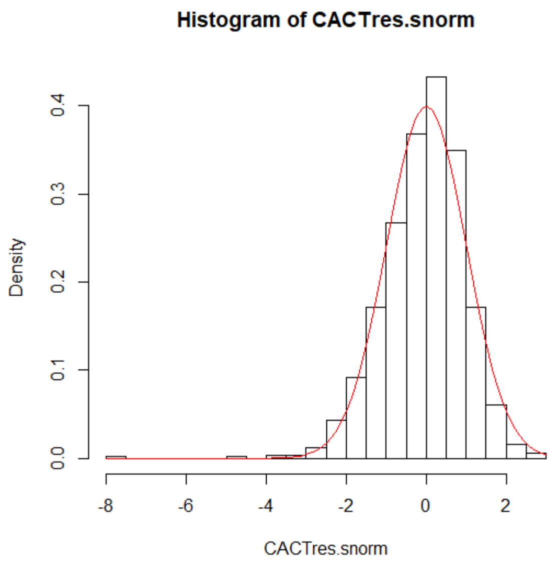

Histogram of CDAXres.norm
Density
0.0
0.1
0.2
0.3
0.4
-6
-4
-2
0
2
CDAXres.norm

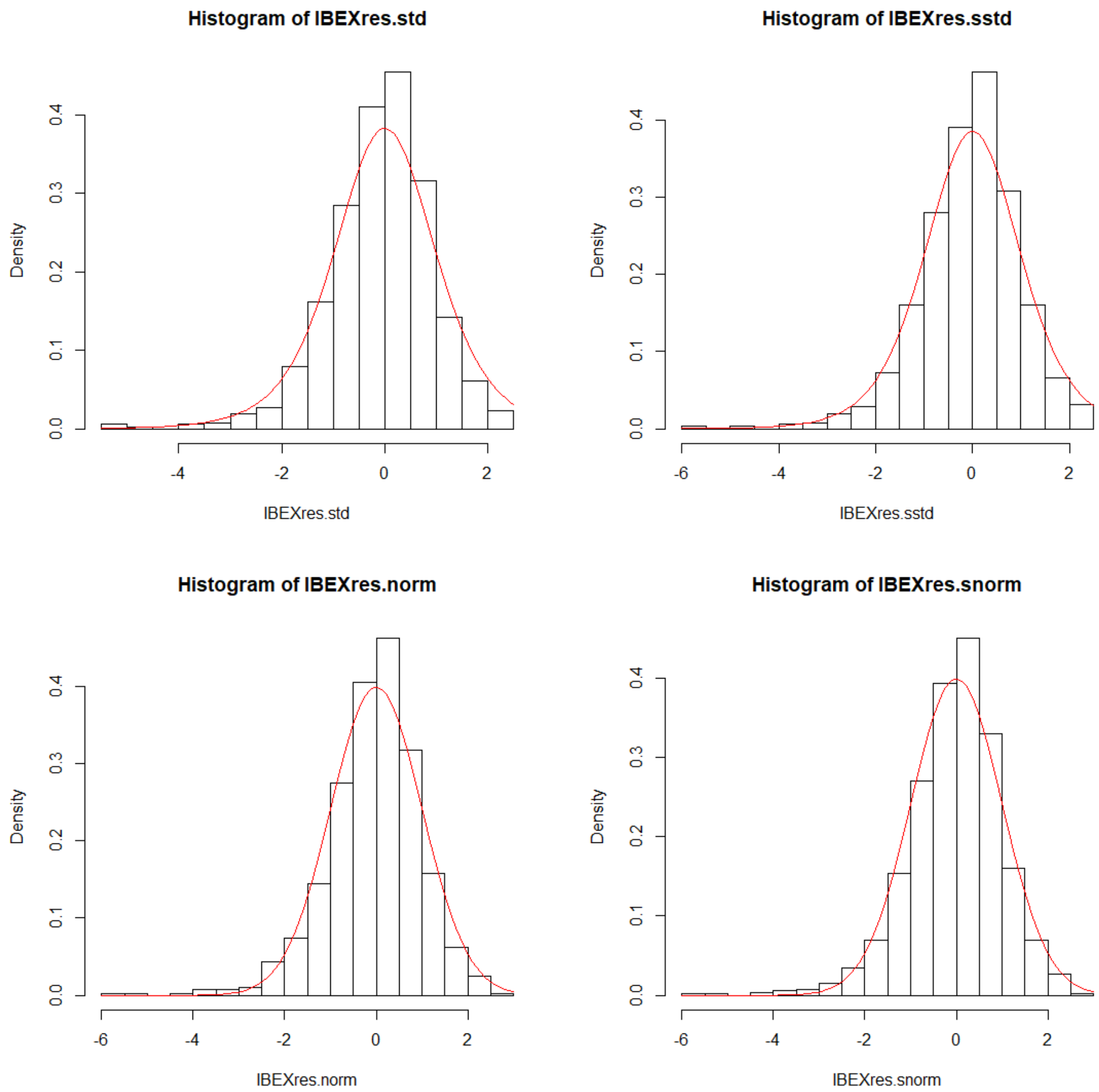

Histogram of CDAXres.snorm
Density
0.0
0.1
0.2
0.3
0.4
-6
-4
-2
0
2
CDAXres.snorm


Histogram of CACTres.std
Density
0.0
0.1
0.2
0.3
0.4
-8
-6
-4
-2
0
2
CACTres.std


Histogram of CACTres.sstd
Density
0.0
0.1
0.2
0.3
0.4
-8
-6
-4
-2
0
2
CACTres.sstd


Histogram of CACTres.norm
Density
0.0
0.1
0.2
0.3
0.4
-8
-6
-4
-2
0
2
CACTres.norm


Histogram of CACTres.snorm
Density
0.0
0.1
0.2
0.3
0.4
-8
-6
-4
-2
0
2
CACTres.snorm



We just want to stress that we generate all graphs by using not approximated degrees of freedom, in order to get more precision. This was only for representation sake: when it comes to copula construction, we will insert the approximated integer number.

6.4.Third step: estimating copula parameters

Now, finally, we are to fit data to copula models. We will fit the four indices to five copulas: two elliptical: Student and Gaussian, and three Archimedean: Frank, Gumbel and Clayton. Then, we will even calculate tail dependence coefficients, and we will check whether our results agree with what was previously stated in the theory.

It is now important to stress one thing: as we said before, both financial theory and empirical analysis seem to agree that a t-Student asymmetric distribution should be able to fit well every index distribution. As a consequence, independently of which copula we choose, we should always assume every conditional distribution has the shape of an asymmetric t-Student. The unique exception is given by normal copula, because this particular case is usually associated to Gaussian conditional distributions. Taking all this into account, we could now choose, for each index, residual distribution that was generated under t-Student asymmetric hypothesis, and start fitting different copulas to these four conditional distributions. After getting copula parameters, we could build a fictional portfolio, composed of our four indices, and, depending on the type of copula we decide to consider, calculating Value at Risk of that portfolio. We could even vary portfolio weights, in order to see whether and in which measure does Value at Risk calculation change. According to the different copula choice, same weights portfolios will probably have different Value at Risk estimation, that we will compare with what we previously exposed into copula theory. However, in order to enhance the deepness of our analysis, we want even to construct copulas on conditional distributions other than the asymmetric Student t-option. After having treated the case that seems to be mostly corroborated by our empirical analysis, we will proceed similarly even with non-asymmetric t-Student, normal and asymmetric normal hypothesis. We will assume all our four indices to follow the same conditional distribution.

As we have anticipated, firstly we are to assume asymmetric Student t-distribution for all indices. We will estimate different copulas parameters and construct differently weighted portfolios. By copulas, we will calculate Value at Risk of these portfolios. We will estimate copula parameters in R. We expose progressively the commands employed, explaining their function.

```
Residuals.sstd<-
cbind(FTSEres.sstd,CDAXres.sstd,CACTres.sstd,IBEXres.sstd)
```

To begin with, we said, we assume an asymmetric t-Student distribution for all indices. Here, we combine our residuals by columns: it will make the work more computationally easy. Now we have to create a generic Student t-copula in four dimensions, where dimension stays for number of distributions:

```
t.cop<-tCopula(dim=4)
```

Mind that this copula has not been fitted to data yet: we just selected copula model.

```
m<-pobs(as.matrix(Residuals.sstd))
```

Here we are asking R to consider our data as a unique matrix, and to compute pseudo observations of our residuals. Now, we can finally combine Student t-copula structure to data: by the following code, we fit residuals matrix to the Student t-copula in four dimensions that we previously created:

```
fitT.sstd<-fitCopula(t.cop,m,method="ml")
```

The last step is extracting copula estimated coefficients:

```
coef(fitT.sstd)
```

rho.1	df
0.5731839	6.3227985

Student t-copula is the unique we are going to treat that has more than one parameter: the first is the rho, and the second is the number of degrees of freedom. The rho indicates the degree of dependence between conditional distributions that compose the copula. It seems obvious, but we want to underline anyway that these degrees of freedom are referred to the copula, while before we estimated degrees of freedom for each conditional distribution: they must be detached.

Now we can avail ourselves of our copula parameters estimation to carve out tail dependence coefficients. We take the two parameters estimated before by the copula built on our data, and we ask R which are tail dependence coefficients of a hypothetical copula with equal parameters:

```
t.cop.sstd<-tCopula(0.5731839,dim=4,df=6.3227985)
TailDep_student.sstd<-lambda(t.cop.sstd)
TailDep_student.sstd
```

lower	upper
0.1997087	0.1997087

It is nice to see that the results agree with what we said before in the theory: being an elliptical copula, Student t-coefficients of tail dependence are identical. Moreover, they are positive, as a Student t-distribution, if compared with a Normal distribution, puts more probability in the tails.

Now that we have disclosed the codes, we can cut it short just by exposing in a table all estimated parameters for our five copula choices. We will even add a second table, dedicated to tail coefficients, and, then, we will spend a bit of time commenting the results. By now, we will assume, in each copula, the same conditional distribution option for all our indices involved: in the first table, copula estimated parameters in the first column are referred to the case in which all series of residuals have been extracted from a GARCH- ARMA model with the same asymmetric t- Student distribution. This condition remains true into first column of the second table, referring to tail dependence coefficients. We will have four columns in each table, constituted by estimated parameters and tail coefficients respectively, according to homogeneous conditional choice. By now, here are estimated parameters and tail coefficients, according to asymmetric t- Student, non- asymmetric t- Student, asymmetric Normal and non-asymmetric Normal, respectively.

Copula parameters

	Asymmetric t- Student	Non- asymmetric t- Student	Asymmetric Normal	Non- asymmetric Normal
Student Copula	0.5731839 (rho) 6.3227985 (df)	0.572762 (rho) 6.336552 (df)	0.5722492 (rho) 6.5119885 (df)	0.571760 (rho) 6.650786 (df)
Normal	0.563937	0.5634775	0.562976	0.5625416
Frank	3.804465	3.798213	3.79577	3.792604
Clayton	0.886848	0.8840733	0.8796065	0.873033
Gumbel	1.510486	1.510046	1.50837	1.507583

Upper and lower tail dependence coefficients

	Asymmetric t- Student	Non- asymmetric t- Student	Asymmetric Normal	Non- asymmetric Normal
Student Copula	0.1997087 0.1997087	0.1990079 0.1990079	0.1930753 0.1930753	0.1884682 0.1884682
Normal Copula	0 0	0 0	0 0	0 0
Frank Copula	0 0	0 0	0 0	0 0
Clayton Copula	0 0.45768	0 0.4565586	0 0.4547445	0 0.4520543
Gumbel Copula	0.4176831 0	0.4174715 0	0.4166641 0	0.4162843 0

We know that the case of Normal copula with asymmetric Student t-residuals distributions is a bit weird, as usually a Normal copula is combined with even normal marginal distributions, and that, assuming different options twists its shape. However, we want to add even this case into the analysis.

It makes not surprise that, from Normal copula, we obtain both lower and upper tail dependence coefficients equal to zero, independently of marginal assumption chosen. Coefficients are identical, as we are treating an elliptical copula, and they are equal to zero, as a Normal copula does not give weight to the tails, apart from the case of perfect correlation. If the estimate is correct, Frank copula should demonstrate both zero lower and upper tail dependence: like a Normal one. As it was previously exposed in theory, Clayton copula gives much weight to joint negative events, as we can infer from positive lower tail dependence coefficient. At the same time, it is not very suitable to express probability of joint positive events, as upper tail dependence is equal to zero. Probably, according to what we anticipated in the theory, this copula would be the best to fit our portfolio, as it gets along with both our empirical results, and with general financial consideration that negative correlation in periods of stress is much stronger than positive correlation in good times. Gumbel copula parameter and Gumbel tail dependence coefficients support general Gumbel shape, perfectly reversed with respect to Clayton copula: it assigns no weight to extreme joint negative events, while it does to positive.

We know that, usually, normal margins are not the most obvious fit for a Student t-copula: Student margins are used, instead. Otherwise, even before we programmed a structure a bit heterodox: a Normal copula with Student t-margins. This is just the opposite case. As we have already said before, we are not only interested in checking whether Value at Risk varies only in a classic contest: we even want to explore what could happen if we insert a twisting case in the analysis, and whether it deviates results. Even being an heterodox case, we can notice that Student t-copula fundamental characteristics are respected: non-zero and identical lower and upper tail dependence.

7.VALUE AT RISK DEFINITION

Now we got all the parameters and we are finally ready to compute Value at Risk. Although, before getting deep into calculations, let's revise what is exactly Value at Risk. T.J. Linsmeier and N.D. Pearson (1996) offer a nice, handy definition: they state that Value at Risk is no more than a single number, that is, a statistical resume of possible portfolio losses. Losses computed for Value at Risk estimation are specified to be due to normal market movements. In normal conditions, losses greater than Value at Risk estimation are, or at least should be, suffered only with a small probability, specified previously.

Getting more specific, given a portfolio, we have to choose a confidence level p. Value at Risk gives a threshold of loss, over a given time horizon. This loss is expected to be exceeded with a probability of only (1-p)% of the time.

Substantially, we can put it even easier: Value at Risk is a quantile in the distribution of profits and losses of a portfolio. It seems to be a very simple indicator as we translate all of the risks of a portfolio into a single number. As a consequence, it is very practical in a boardroom or in an annual report. Because of its practicality, according to O. Guéant, Value at Risk should be one of the most commonly used measures in financial risk management. Unfortunately, O. Guéant even says that easy work ends with Value at Risk definition: theoretical distribution of profits and losses of a portfolio is not observable and must be estimated someway. So, in order to compute Value at Risk we need to get involved in statistic, but not only. As our portfolios are often large and composed of a huge bunch of complicated financial assets, our estimate requires approximation and correct asset pricing. As we cannot select and take into account all risk factors, we need to take note only of the most relevant. Portfolio components need to be matched to these risk factors before being priced. Being a statistical problem, computation of Value at Risk can be carried out using various methods. O. Guéant divides them into three groups:

- Historical simulations
- Parametric methods, also called analytical methods
- Monte Carlo methods

Now, we want to briefly describe each of them, then; we will expose how U. Cherubini, E. Luciano and W. Vecchiato (2004) adopt historical simulation to estimate Value at Risk for a portfolio of two assets. Secondly, we will expose our technique for estimating in R a four assets portfolio Value at Risk. Our procedure will be based on Monte Carlo simulations. Lately, we will put our project into practice, we will show the results and we will comment them.

As we are to start with generalities of historical simulations, we will say that the first step for estimation is choosing a certain number of relevant risk factors, relative to the portfolio. Then, we have to recover data from past behaviour of these risk factors. The objective is drawing a possible evolution of portfolio price and figuring out potential losses that would have been suffered if we had hold that portfolio during the period to which risk factors data are referred. O. Guéant reports that historical simulations are heavily used as they do not require any calibration concerning interdependence structure. To put it easier, it means that we have neither to consider eventual correlation behind variables of interest. The few hypothesis we have to formulate make Value at Risk based on historical simulations quite easy to elaborate.

Lately, things started getting a bit more complex in the mid-90's, with parametric or analytical methods. They are based on strong assumptions about risk factors returns distributions. Notwithstanding being numerous, all parametric methods share a common advantage: Value at Risk can be computed very easily. If we decided to apply a parametric method, in most of the cases we would find us resorting to a Taylor expansion to approximate the portfolio, and then relying on the Greeks of portfolio assets. A nice consequence of this is that we would not need to fully reprice our standard parametric method after every market movement. However, there is even an uncomfortable side: even if Delta approach has been enriched in order to take non-linearities into account, not all linearities can be encompassed by the model.

Anyway, we do not have to care so much about this, as we are to use the latest family of methods: Monte-Carlo simulations. With respect to historical simulations, in Monte-Carlo simulations samples are not bounded to be based on past, recorded realizations of risk factors. Although, we have to estimate distribution parameters for the risk factors and, then, draw scenarios for the joint distribution. The main advantage of this method is that we can tailor a proper distribution for each risk factor. The bad news is that we should have to estimate many parameters, actually all marginal parameters we need to depict every marginal risk factor distribution, and this risks to turn out to be pretty slow.

As we have to choose one of these methods to estimate our Value at Risk, O. Guéant advises us that historical simulations, so, the first method exposed, are mostly non parametric and are therefore able to take into account structure of dependence between different risk factors involved. Otherwise, the adoption of this strategy would mean that we have only to avail ourselves of true historically recorded data. If there is lack of data, we cannot proceed with the analysis. Especially, if we consider we have to estimate an extreme quantile, having few data could seriously bias our estimate. One evolution of this first approach was given by parametric methods: they rely on historical data only with the purpose of fitting some parameters, like standard deviation, for instance. Then, given the parameters, they allow to proceed with the estimation of Value at Risk by both an approximation of the portfolio of interest and eventual distributional assumptions on the risk factors, concerning details that cannot be encompassed by previously estimated parameters. It is here that limitations arise: there can be, O. Guéant states, approximations that may be hardly adapted to extreme risk for certain portfolios. However, we have a main advantage in parametric method: speed of computation.

Otherwise, we are mostly interested in the third family amongst methodologies: Monte-Carlo simulations. Following T.J. Linsmeier and N.D. Pearson (1996), the main difference between Monte-Carlo simulation and historical simulation is that with Monte-Carlo, rather than generating N hypothetical portfolio profits or losses, carrying out the simulation using last N periods observed changes in market factors, we have to choose a statistical distribution believed to fit adequately market factors. Then, we employ a pseudo-random number generator to generate thousands of hypothetical market factors changes. These are then used to elaborate thousands of hypothetical portfolio profits and losses on the current portfolio, and the distribution of these profits and losses. Lastly, we have to estimate the Value at Risk for the portfolio by estimating the corresponding quantile of the resulting distribution of profits and losses. With respect to historical simulations, the great advantage of Monte-Carlo is that the amount of data we can work with is not restricted by historical records. As we are to simulate values in R, we can draw as many trajectories as we want, even one billion, after having calibrated these draws on historical data. In order to make all this discussion more formal, O. Guéant resumes that we have to avail ourselves of past data to fit a distribution f to risk factors. Then we have to generate a large number M of new values for our risk factors. Theoretically, if we have to handle financial assets portfolios, O. Guéant reminds that we could better use risk factors returns, instead of simple risk factors. Anyway, the general idea does not change. Here, O. Guéant calls the new values for the risk factors $(X_{t+1}^1, \dots, X_{t+1}^n)$, that we refer to as:

$$\left(X_{t+1}^{1,m}, \dots, X_{t+1}^{n,m}\right) \quad 1 \le m \le M$$

Then the time consuming step of Monte-Carlo simulations consists of the evaluation of the portfolio for these new values of the risk factors:

$$P_{t,t+1}^m = V\left(X_{t+1}^{1,m}, \dots, X_{t+1}^{n,m}\right) - V(X_t^1, \dots, X_t^n) \quad 1 \le m \le M$$

Once this step has been completed, we can calculate Value at Risk by calculating the quantile of interest in the empirical distribution of profits and losses.

Now that we have exposed the generalities of all three methods, we are to report how U. Cherubini, E. Luciano and W. Vecchiato (2004) apply copulas to historical simulation in order to carry out Value at Risk for a two assets portfolio. We are to explain their application in detail as it offers a starting point on which we work with the aim of calculating Value at Risk for a portfolio of four financial assets. U. Cherubini, E. Luciano and W. Vecchiato resume that, for a given confidence level θ, Value at Risk is the level under which returns will fall only with probability θ. If we denote as Z the portfolio return over a given horizon T, Value at Risk is the threshold such that:

$$Pr(Z \le VaR_Z) = \theta$$

Consider a portfolio of two assets. Let X and Y be their continuous returns, over a common horizon T, and let $\beta \in (0,1)$ be the weight of X. The portfolio return is $Z = \beta X + (1-\beta)Y$, with distribution function:

$$Pr(Z \le z) = Pr(\beta X + (1-\beta)Y \le z)$$

$$= \int_{-\infty}^{+\infty} Pr\left(X \le \frac{1}{\beta}z - \frac{1-\beta}{\beta}y, Y = y\right) f_2(y)dy$$

Now we say that our portfolio can be represented by a copula: single assets distributions are the marginal distributions, while joint copula distribution encompasses every association or dependence relation between the two assets. We apply copula to estimate Value at Risk, given a confidence level θ, for the portfolio. Let x_1 and x_2 be our assets returns, and $\beta \in (0,1)$ the allocation weight, so the portfolio return is given by $z_t = \beta x_{1t} + (1-\beta)x_{2t}$ where, omitting the subscript t,

$$F(x_1, x_2) = Pr(X_1 \le x_1, X_2 \le x_2)$$

$$= Pr\big(F_1(X_1) \le F_1(x_1), F_2(X_2) \le F_2(x_2)\big) = C(F_1(x_1), F_2(x_2))$$

And by derivation we express it in terms of probability density function:

$$f(x_1, x_2) = c\big(F_1(x_1), F_2(x_2)\big) \cdot \prod_{j=1}^{2} f_j(x_j)$$

Hence, the cumulative density function for the portfolio return Z is given by:

$$Ƃ(z) = pr(Z \leq z) = Pr(\beta X_1 + (1-\beta)X_2 \leq z)$$

$$\int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{\frac{1}{\beta}z - \frac{1-\beta}{\beta}x_2} c\big(F_1(x_1), F_2(x_2)\big) f_1(x_1) dx_1 \right\} f_2(x_2) dx_2$$

Where upper integral limit $\frac{1}{\beta}z - \frac{1-\beta}{\beta}x_2$ is obtained by putting $\beta x_1 + (1-\beta)x_2$ in function of x_1. The Value at Risk for the portfolio, at a confidence level $\theta \in (0,1)$ and for a given weight $\beta \in (0,1)$, is the solution z* of the equation $Ƃ(z^*) = \theta$. This result may be extended straight to a portfolio with n-assets, with the condition that all assets weights sum up to 1.

In the beginning, our idea was to exploit this formula to calculate Value at Risk even for our portfolio. We had enough data, as we disposed of weekly observations for more than twenty years for each index. Moreover, we had already estimated conditional distributions parameters, according to different hypothesis. We had even estimated copula parameters, both for elliptical and Archimedean copulas, in order to check if our results would have changed depending on the copula employed. However, problems came out in developing four integrals in a row, one for each asset in our portfolio: this would have consistently slow down our proceedings. Moreover, as it is possible to check in the two assets formulation by U. Cherubini, E. Luciano and W. Vecchiato (2004), we would have had to write explicitly density functions, both for all portfolio assets conditional distributions and for the copula itself. In the end, this would have turned out to be too much computationally intensive, while we wanted to find a much faster and easier way to carve out many Values at Risk, varying conditional assumptions, copula choice and even portfolio weights. In order to deduce an easier expression, we have to consider more carefully U. Cherubini, E. Luciano and W. Vecchiato's (2004) two integrals formulation. We can see that in both cases the lower bound is -∞, while the upper bound changes: it is on the latter we need to focus. We have to start considering the inner integral: its upper bound corresponds to portfolio return formulated in function of x_1, that is, the first asset. Being Value at Risk a quantile, we should find out x_1, as it indicates the level in correspondence of which θ expresses the probability that our portfolio return is smaller or equal than z. In fact, as argument of the first integral, we have marginal distribution of asset x_1 and joint distribution of the copula. In the first integral we have formulated our quantile of interest in function of x_1, and, putting as argument both first asset marginal density function and copula function, it is like we have tried to indicate how much first asset can determine portfolio Value at Risk. By second integral, we should, so, take into account how much second asset can influence quantile determination. This formulation allows to not neglect interdependence between assets, that is substantially the reason for which we use copula instrument. In fact, imagine that, while fitting into the first integral both first asset density function and copula density function, expressing integral upper bound in function of first asset, we cannot completely explain first asset yet, as in joint distribution we have still to express x_2. Only with second integral we can consider even x_2, so only with second integral we can complete joint density function with second asset conditional distribution, and in the end extract quantile z. We would like even to notice that, if we have a portfolio of just two assets, we have to write inner integral upper bound as portfolio return in function of the first asset, but we do not have to express second integral upper bound in function of second asset: we have just to write +∞. This makes things very comfortable and it is substantially due to a system of two equations:

$$\begin{cases} \beta X_1 + (1-\beta) X_2 \le z \\ \beta + (1-\beta) = 1 \end{cases}$$

Second element seems obvious and pertains to portfolio weights, however, it assure us that we do not need to specify additional information in second integral upper bound, as, specifying first asset in first integral upper bound, we do not have to deduce additional information for second integral.

Having a four assets portfolio, applying integrals like U. Cherubini E. Luciano and W. Vecchiato would have been really tricky, so we found out a different method, more computationally practical, that we are to show immediately. We just want to specify one thing before: as it is possible to see, U. Cherubini, E. Luciano and W. Vecchiato use historical simulation: they do not have to simulate data by a random generator on the basis of a previously fitted distribution of conditional returns. They simple resort to historical, recorded data. We, instead, want to apply Monte-Carlo simulation to increase precision. We have already estimated degrees of freedom of each Student t-asset distribution, and we do not need other parameters. Moreover, we have even estimated copula parameters for both Archimedean and elliptical copulas. As a consequence, now we choose a copula. We take as a starting point residual distributions, on which we estimated all the parameters of the copula. Then, by a random generator in R, we estimate a really high number of fictional joint distributions, that have all to share both marginal and copula parameters estimated previously. We build our portfolio selecting the weights of the assets, paying attention that all weights must sum up to 1. In addition to this, it is extremely important to underline that, when we indicate portfolio assets, we are not referring to recorded historical returns, but to random generated assets returns, that have been created by R when we have scattered a series of random generated copulas with fixed number of marginal distributions and fitted parameters for each random distribution. After having settled the weights of our portfolio, we are to calculate its Value at Risk just by extracting the corresponding quantile. Having used a copula, our result takes automatically into account all possible interdependences between assets.

Getting into practice, we will start by estimating Value at Risk for a portfolio represented by a Student t-copula with Student t-assets returns distributions. Later, we will even calculate Value at Risk by a Normal copula with Student t-margins, and by a Student t-copula with normal margins. Maybe these latter cases could appear a bit bizarre: usually, Normal copula is associated with Gaussian marginal distributions, and vice versa. However, we want to apply all five copulas previously exposed in the analysis. The objective is checking if Value at Risk changes and how, according to both copula and margins choice. We will even alter the weights of the portfolio, to see if estimate changes. Lastly, we will control if all our results agree with what was stated in theory before, and we will try to indicate which could be the better copula to take portfolio risk into account. Now, we are to carefully depict all passages with Student t-copula, and later we will expose a table with all Value at Risk estimates.

8.CORRELATIONS

Before proceeding with copula fitting process, we waste a bit of time checking the correlation between the variables. According to what we have exposed previously, we are not to give linear correlation matrix, as it is not the optimal measure of association. We will show rank correlation instead: Kendall's tau and Spearman's rho coefficients. Here we expose the codes to associate residuals by columns and carve out matrix of Kendall's and Spearman's coefficients. We want to notice that, in the first table, we refer to all our four indices residuals obtained by a GARCH- ARMA model with asymmetric t- Student marginal shape option. Later, we will draw similar matrix even for the other marginal distribution options. We want to be sure the magnitude of association does not change, depending whether we choose to adopt one distribution or the other:

```
Residuals.sstd<-cbind(FTSEres.sstd,CDAXres.sstd,CACTres.sstd,
IBEXres.sstd)
cor(Residuals.sstd,method="kendall")
```

Asymmetric t-Student residuals- Kendall

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.2077630	0.2375144	0.3500005
CDAXres	0.2077630	1.0000000	0.7437551	0.4221804
CACTres	0.2375144	0.7437551	1.0000000	0.4659060
IBEXres	0.3500005	0.4221804	0.4659060	1.0000000

Asymmetric t-Student residuals- Spearman

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.3071783	0.3498962	0.5040315
CDAXres	0.3071783	1.0000000	0.9116999	0.5926453
CACTres	0.3498962	0.9116999	1.0000000	0.6436635
IBEXres	0.5040315	0.5926453	0.6436635	1.0000000

Non- asymmetric t-Student residuals- Kendall

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.2078809	0.2375334	0.3494604
CDAXres	0.2078809	1.0000000	0.7432378	0.4211572
CACTres	0.2375334	0.7432378	1.0000000	0.4652023
IBEXres	0.3494604	0.4211572	0.4652023	1.0000000

Non- asymmetric t-Student residuals- Spearman

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.3072186	0.3497704	0.5029552
CDAXres	0.3072186	1.0000000	0.9114143	0.5916194
CACTres	0.3497704	0.9114143	1.0000000	0.6429199
IBEXres	0.5029552	0.5916194	0.6429199	1.0000000

Asymmetric Normal residuals- Kendall

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.2064204	0.2370846	0.3502477
CDAXres	0.2064204	1.0000000	0.7402101	0.4210850
CACTres	0.2370846	0.7402101	1.0000000	0.4651376
IBEXres	0.3502477	0.4210850	0.4651376	1.0000000

Asymmetric Normal residuals- Spearman

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.3051137	0.3494250	0.5050080
CDAXres	0.3051137	1.0000000	0.9096571	0.5919211
CACTres	0.3494250	0.9096571	1.0000000	0.6430488
IBEXres	0.5050080	0.5919211	0.6430488	1.0000000

Non- asymmetric Normal residuals- Kendall

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.2064889	0.2372633	0.3501488
CDAXres	0.2064889	1.0000000	0.7413626	0.4212371
CACTres	0.2372633	0.7413626	1.0000000	0.4657880
IBEXres	0.3501488	0.4212371	0.4657880	1.0000000

Non- asymmetric Normal residuals- Spearman

	FTSEres	CDAXres	CACTres	IBEXres
FTSEres	1.0000000	0.3054580	0.3496436	0.5048907
CDAXres	0.3054580	1.0000000	0.9102985	0.5920506
CACTres	0.3496436	0.9102985	1.0000000	0.6435536
IBEXres	0.5048907	0.5920506	0.6435536	1.0000000

It is nice to see that choice of marginal distribution shape does not alter association measures between indices. The adoption of Kendall’s tau rather than Spearman’s rho has an influence on the magnitude of the association, however, it has not on the order of the couples, if we dispose them from less positively related to more strongly positively related. The logic behind this is not just getting an idea whether two variables are strongly correlated or not. We will recover this table when we will have to calculate Value at Risk for portfolios constituted of our four indices. Theoretically, if we assigned strong weight to two assets that are even relatively strongly correlated, like CACT and CDAX, we would change for worse Value at Risk estimation, as we are making our portfolio riskier.

9.VALUE AT RISK ESTIMATION

As now we can finally pass to Value at Risk estimation, we will start by exposing and commenting every line of code we are to use, in order to clarify our proceedings. We will refer to a Student t-copula with asymmetric Student t-conditional distributions. After this explicative example, we will resume all Value at Risk estimates in a series of tables, comprehensive of every possibility given by copula choice and marginal shape option. We will even adopt different portfolio compositions, checking in which measure does Value at Risk vary, depending whether we focus more or less on relatively strongly positively correlated assets.

Anyway, for all cases we need to use the following command, whose point is generating an object, called “r”, that is set equal to the number of random copulas we want to generate. This is the first step of Monte Carlo simulation: when we are to generate an extremely high number of objects, on which to deduce our estimate. As we are lucky and we dispose of a powerful software, we settle r equal to 100000, as R can perfectly handle 100000 randomly generated copulas, whose parameters have been previously estimated from real series of residuals.

```
r<-100000
```

At this point, before applying random generation and Monte Carlo simulation, we need to fit the copula. We do not want to start immediately with something heterodox, like a Normal copula with Student t-conditional distributions: we will choose a Student t-copula with asymmetric Student t-margins. Maybe this will complicate a bit our codes, as, while for a Gaussian copula we just need a unique parameter of joint dependence, for Student t-copula we need two: degrees of freedom of the copula and rho, as previously said. In the following codes, we briefly resume how to estimate copula parameters from our matrix of residuals: we generate the structure of a Student t-copula in four dimensions, then we fit to it the pseudo matrix of residuals.

There is one thing we would like the reader to mind: as we exposed before in the theory, we are to use Inference for Margins Method, that is more practical than Maximum Likelihood Estimation, and releases efficient results. When we ask the software to apply "ml" in estimating copula parameters, this "ml" is not to be confused with Maximum Likelihood Estimation. True maximum likelihood method would release both estimates for copula parameters and for conditional distribution parameters in one single step. Here, instead, the code that we are to use allows us to estimate separately conditional distributions parameters and copula parameters, as in the Inference for Margins method. By conditional distribution parameters we mean, for example, degrees of freedom of FTSE MIB, CDAX and other indices distributions, in the case we assume for GARCH-ARMA model a Student t-distribution option. In the case we have assumed a Normal conditional distribution instead, we do not have to insert parameters: later we will have just to write "norm" for indicating marginal shape. As a conclusion, here we are with the first codes:

```
t.cop<-tCopula(dim=4)
m<-pobs(as.matrix(Residuals.sstd))
fitT.sstd<-fitCopula(t.cop,m,method="ml")
coef(fit.sstd)
```

Where, by last line, we ask the software to release estimated copula parameters. Now, given both our conditional distributions parameters and our copula parameters, we can construct a fictional copula of the same type, with the same number of dimensions, with both previously estimated copula parameters, like rho and degrees of freedom of the copula, and marginal parameters, like degrees of freedom of each marginal distribution, that was assumed to be a t-Student. By writing "t" for each conditional distribution, we even specify that we assume every index series of residuals to be well described by a t-Student. When we will not have Student t-margins, but Normal margins, we will have to insert the mean and the standard deviation as parameters. Moreover, it is nice to notice that we are not bound to write always the same type of distribution for all margins: we could even ask for a copula with two t-Student and two Normal conditional distributions, as an example.

```
t.cop_sstd<-mvdc(tCopula(param=0.5731839,df=6,dim=4),
```

```
margins=c("t","t","t","t"),
paramMargins=list(list(df=8),list(df=10),list(df=11),list(df=7)))
```

We just want to specify one thing: last copula has been drawn on the basis of previously estimated parameters, however, it is completely disentangled from real distributions of indices' residuals: it is fictional.

At this point, we have to proceed with simulation: according to the copula structure we have just specified, we ask R to generate 100000 random copulas with the same characteristics:

```
Sim_student_sstd<-rMvdc(r,t.cop_sstd)
```

We have to specify three confidence levels (1-α), in correspondence of which calculate Value at Risk: returns will fall only with probability (1-α). We generate a vector alpha with our chosen levels. Levels usually taken corresponds to 1%, 5% and 10%, so we are to follow this tradition, however, even lower or higher references could be adopted.

```
alfa<-c(0.01,0.05,0.1)
```

Next we have to detach our marginal distributions from the copula and give them a name. This will be useful when we will have to build our portfolios, because the conditional distributions will be our assets. In order to better understand next codes, remember that we have 100000 copula simulations, and imagine our assets disposed into a matrix. Having 100000 simulations of copula built on our four assets, this means that we have automatically even 100000 simulations for each of our conditional asset distribution. Think assets are disposed into a matrix by columns: with next codes we are just calling column by column, naming each of them as the corresponding asset.

```
FTSE_student.cop_sstd<-(Sim_student_sstd[,1])
CDAX_student.cop_sstd<-(Sim_student_sstd[,2])
CACT_student.cop_sstd<-(Sim_student_sstd[,3])
IBEX_student.cop_sstd<-(Sim_student_sstd[,4])
```

Now we will create a portfolio with equal weights and with all the assets we dispose of:

```
portfolio_N1<-
0.25*FTSE_student.cop_sstd+0.25*CDAX_student.cop_sstd+
0.25*CACT_student.cop_sstd+0.25*IBEX_student.cop_sstd
```

As Value at Risk is substantially a quantile of the portfolio, in order to get its estimate we have finally to extract the quantiles written in vector alpha from our portfolio with equal weights:

```
quantile(portfolio_N1,alfa)
```

1%	5%	10%
-2.361328	-1.515750	-1.139032

Here we have quantiles of the easiest portfolio to compose : we just gave all assets the same weight. Although, before dedicate ourselves to other copulas than the Student with asymmetric Student t-margins, we would like to linger a bit on portfolio composition. In particular, we wonder if, changing assets weights, we would obtain strong differences in Value at Risk.

9.1.Portfolio variance

We hypothesize four differently weighted portfolios, more or less unbalanced, and we will calculate their variances. We choose portfolios with minimum, medium and high variance and, then, we estimate their Value at Risk. Theoretically, Value at Risk should be higher the stronger is the variance in the portfolio: as variance substantially expresses risk, and Value at Risk is a way to measure risk. Anyway, we cannot just say that the portfolio with identical weights is the most balanced, as we have even to take into account association measures. As we explained previously when we exposed association measures, we are not to refer to linear correlation coefficient, as it, by its name, takes into account only linear correlation. We are to use rank association instead: Kendall's tau or Spearman's rho. From association matrices we can check that they are quite similar, Spearman's rho tends just to be constantly a bit higher than Kendall's tau: in next calculations, we will resort to Spearman's rho. The general formula to get variance for a portfolio of four assets is:

$$\sigma_P^2 = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + x_4^2\sigma_4^2 + 2x_1x_2\sigma_1\sigma_2\tau_{1,2} + 2x_1x_3\sigma_1\sigma_3\tau_{1,3} + 2x_2x_3\sigma_2\sigma_3\tau_{2,3} + 2x_1x_4\sigma_1\sigma_4\tau_{1,4} + 2x_2x_4\sigma_2\sigma_4\tau_{2,4} + 2x_3x_4\sigma_3\sigma_4\tau_{3,4}$$

With respect to traditional formulation, we are to change linear correlation coefficient with Spearman's rho coefficient. This is just a brief test, in order to choose different portfolio weights for our Value at Risk experiment. Now we expose the code used to calculate variance of the same weights portfolio, where later we will just expose variances obtained from other more or less balanced portfolios.

```
port_variance<-
(var(FTSEres.sstd)*0.25^2)+(var(CDAXres.sstd)*0.25^2)+
+(var(CACTres.sstd)*0.25^2)+(var(IBEXres.sstd)*0.25^2)+
+(2*0.25*0.25*sd(FTSEres.sstd)*sd(CDAXres.sstd)*0.3071783)+
+(2*0.25*0.25*sd(FTSEres.sstd)*sd(CACTres.sstd)*0.3498962)+
+(2*0.25*0.25*sd(CDAXres.sstd)*sd(CACTres.sstd)*0.9116999)+
+(2*0.25*0.25*sd(FTSEres.sstd)*sd(IBEXres.sstd)*0.5040315)+
```

```
+(2*0.25*0.25*sd(CDAXres.sstd)*sd(IBEXres.sstd)*0.5926453)+
+(2*0.25*0.25*sd(CACTres.sstd)*sd(IBEXres.sstd)*0.6436635)
```

And we get a variance equal to 0.6932016.

Now, we will try to construct a portfolio with lower variance, taking into account that the couples FTSE MIB and CDAX, and FTSE MIB and CACT seem to be not strongly positively correlated; while couples CDAX and CACT, and CACT and IBEX seem to be more strongly positively correlated. When we want to reduce variance, we will try to avoid assigning high weights both to CDAX and CACT, or both to CACT and IBEX. Similarly, in order to augment variance, we will try to give much weight both to CDAX and CACT, or both to CACT and IBEX.

In a few pages, we are to calculate Value at Risk of the portfolio with identical weights. In order to calculate another Value at risk, referred to a differently weighted portfolio, with the same assets, we try to build a new portfolio, less risky. Similarly, we will create even a riskier one, playing with the weights. We start with the less risky: as we said that CDAX and CACT share strong positive dependence, we are to assign them strongly different weights: 45% and 5% respectively. Similarly, as we observed that CACT and IBEX are positively entangled, we are to invest 20% of our portfolio in IBEX, only 5% in CACT and what remains in FTSE MIB. As a consequence, the two assets that are more strongly present in the portfolio, so that have the highest weight, are only weakly dependent. We obtain a variance equal to 0.6786376, that is smaller than the one calculated with equal weights portfolio.

In order to create a high variance portfolio, we reason in the same way: we assign much weight both to CACT and to CDAX, as they are positively entangled. Theoretically, we could have done the same with the couple CACT and IBEX, but Spearman's coefficient is higher for CACT and CDAX, so we expect the effect to be stronger. We assign 5% to FTSE MIB, 45% to CACT, 30% to CDAX and 20% to IBEX. We can use same codes as before, after having changed weights. We find out that variance is 0.8423955, that is the highest variance till now.

Lastly, we just want to create a fourth portfolio, with intermediate weights, but different from the equal weights portfolio. We settle these random weights like this: 25% FTSE MIB, 25% CDAX, 15% CACT and 35% IBEX. Its variance is 0.6913021.

Before proceeding with Value at Risk estimation, we want just to clarify a couple of things. We were looking for four different portfolio compositions in order to check whether and how much Value at Risk could be influenced by a more or less balanced choice of assets.

However, we are not interested in determining the lowest or highest variance portfolio: in that case we would have thought about a more scientific way to do it. We just wanted to have four portfolios from which we expected relatively different Value at Risk estimations.

9.2.Our results

Now that we have even found out assets weights balance, we can finally pass to estimate Value at Risk of all portfolios, assuming different conditional distributions for assets and different copulas entangling them. We will show some tables, in order to demonstrate all possible Value at Risk estimates. For each portfolio composition, we will present three tables, one for each confidence level at which we have to calculate Value at Risk: confidence levels that we previously exposed in vector alpha. In each table, we will gather Value at Risk estimates, with that portfolio composition and at that confidence level, with all possible different combinations of copula and conditional distributions of assets choice. We just have to add a small precision: almost all codes are similar to the one we previously exposed, with a Student t-copula and asymmetric Student t-margins. The unique difference is given by the case in which we have to specify Normal margins, instead of Student t-margins. Here, instead of having to insert degrees of freedom of each asset distribution, in order to define Normal conditional distribution we have to indicate mean and standard deviation of the corresponding series of asset residuals. Moreover, when now we have to specify which type of marginal distribution do we want for our assets, we do not have the option to include asymmetry. However, this constitutes no problem. Asymmetry has been already taken into consideration when we have calculated copula parameters: we have fitted copula structure to the matrix of asymmetric residuals. In the following we will parade all the tables, and then we will comment the results.

25% FTSE MIB+ 25% CDAX+ 25% CACT+ 25% IBEX

Alpha 1%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-2.273783	-2.350577	-1.987967	-1.958898
Student Copula	-2.361328	-2.438872	-2.052793	-2.021525
Frank Copula	-1.915491	-1.962420	-1.702977	-1.690158
Clayton Copula	-2.553492	-2.638603	-2.203706	-2.172285
Gumbel Copula	-1.958684	-2.03379	-1.735223	-1.717676

25% FTSE MIB+ 25% CDAX+ 25% CACT+ 25% IBEX
Alpha 5%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.517775	-1.548442	-1.428042	-1.411164
Student Copula	-1.515750	-1.546642	-1.439146	-1.422116
Frank Copula	-1.419695	-1.453029	-1.342340	-1.326182
Clayton Copula	-1.579212	-1.608470	-1.486565	-1.475454
Gumbel Copula	-1.359511	-1.39206	-1.280890	-1.272408

25% FTSE MIB+ 25% CDAX+ 25% CACT+ 25% IBEX
Alpha 10%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.153520	-1.171162	-1.125645	-1.115217
Student Copula	-1.139032	-1.156203	-1.119314	-1.109227
Frank Copula	-1.151975	-1.176161	-1.120661	-1.108886
Clayton Copula	-1.135956	-1.154255	-1.111016	-1.104799
Gumbel Copula	-1.065898	-1.08710	-1.029268	-1.031076

30% FTSE MIB+ 45% CDAX+ 5% CACT+ 20% IBEX
Alpha 1%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-2.345511	-2.410581	-2.053015	-2.018036
Student Copula	-2.427182	-2.498765	-2.118691	-2.082708
Frank Copula	-2.052103	-2.097778	-1.829550	-1.799066
Clayton Copula	-2.598529	-2.666584	-2.238795	-2.210347
Gumbel Copula	-2.086616	-2.148906	-1.850365	-1.821784

30% FTSE MIB+ 45% CDAX+ 5% CACT+ 20% IBEX
Alpha 5%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.559266	-1.587175	-1.475301	-1.454653
Student Copula	-1.549433	-1.575947	-1.479572	-1.458572
Frank Copula	-1.485832	-1.513011	-1.406991	-1.388428
Clayton Copula	-1.612940	-1.635563	-1.531010	-1.513050
Gumbel Copula	-1.434196	-1.453878	-1.353239	-1.338759

30% FTSE MIB+ 45% CDAX+ 5% CACT+ 20% IBEX
Alpha 10%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.181487	-1.197403	-1.161426	-1.147828
Student Copula	-1.168136	-1.183359	-1.159004	-1.145070
Frank Copula	-1.186059	-1.196856	-1.162928	-1.141436
Clayton Copula	-1.166777	-1.181943	-1.149227	-1.138901
Gumbel Copula	-1.114176	-1.129925	-1.082104	-1.077463

5% FTSE MIB+ 30% CDAX+ 45% CACT+ 20% IBEX
Alpha 1%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-2.299105	-2.374271	-2.054346	-2.022287
Student Copula	-2.385680	-2.465041	-2.106974	-2.073399
Frank Copula	-2.002010	-2.068352	-1.826544	-1.796237
Clayton Copula	-2.534295	-2.634373	-2.240844	-2.212952
Gumbel Copula	-2.031732	-2.112251	-1.839434	-1.823034

5% FTSE MIB+ 30% CDAX+ 45% CACT+ 20% IBEX
Alpha 5%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.539699	-1.567895	-1.472185	-1.452831
Student Copula	-1.535116	-1.564994	-1.481094	-1.461173
Frank Copula	-1.464784	-1.496408	-1.405392	-1.386017
Clayton Copula	-1.601083	-1.633087	-1.525380	-1.510051
Gumbel Copula	-1.408070	-1.441815	-1.350067	-1.339920

5% FTSE MIB+ 30% CDAX+ 45% CACT+ 20% IBEX
Alpha 10%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.171712	-1.190289	-1.157130	-1.145709
Student Copula	-1.158878	-1.176516	-1.154128	-1.142176
Frank Copula	-1.175686	-1.194706	-1.155186	-1.141895
Clayton Copula	-1.157674	-1.177220	-1.157611	-1.136852
Gumbel Copula	-1.093044	-1.120841	-1.080515	-1.074959

25% FTSE MIB+ 25% CDAX+ 15% CACT+ 35% IBEX
Alpha 1%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-2.315474	-2.400030	-1.998944	-1.970257
Student Copula	-2.405955	-2.490284	-2.062076	-2.032489
Frank Copula	-1.983487	-2.035792	-1.731719	-1.719383
Clayton Copula	-2.585115	-2.678083	-2.203174	-2.178765
Gumbel Copula	-2.018562	-2.104316	-1.761982	-1.733227

25% FTSE MIB+ 25% CDAX+ 15% CACT+ 35% IBEX
Alpha 5%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.535986	-1.570112	-1.436407	-1.419988
Student Copula	-1.535072	-1.566793	-1.440916	-1.424779
Frank Copula	-1.444082	-1.476504	-1.354960	-1.336737
Clayton Copula	-1.595229	-1.626927	-1.493270	-1.479839
Gumbel Copula	-1.390298	-1.421743	-1.292169	-1.286342

25% FTSE MIB+ 25% CDAX+ 15% CACT+ 35% IBEX
Alpha 10%

	Asymmetric t-Student margins	Non- asymmetric t-Student margins	Asymmetric Normal margins	Non- asymmetric Normal margins
Normal Copula	-1.167411	-1.185542	-1.130209	-1.120319
Student Copula	-1.151383	-1.169316	-1.124270	-1.114445
Frank Copula	-1.165905	-1.185639	-1.126637	-1.113432
Clayton Copula	-1.147309	-1.166420	-1.118257	-1.114055
Gumbel Copula	-1.087356	-1.106299	-1.039268	-1.039306

Now that we have elaborated all possible Value at Risk estimates, we can spend a couple of words commenting our results. In the first three tables, we refer always to the first portfolio, the one composed of equal weighted assets. In each table we compute Value at Risk estimates for each copula considered into the analysis. We have five copulas, two elliptical and three Archimedean. For each possible copula choice, we compute Value at Risk with four different options: whether all our conditional distributions are shaped like a t-Student or a Normal, and whether do they encompass asymmetry option or not. We get a total of twenty estimates in each table. The first table is referred to Value at Risk estimates of the first portfolio, adopting a confidence level of 1%. In the second table, referring the same first portfolio, we have a confidence level of 5%, and in the third table of 10%. Then, we have shown a similar organization for all other three portfolios we proposed, adopting for each all possible choices of copulas and of conditional distributions type. For each portfolio we display three tables, given by the three confidence levels we decided to treat. We start from the first table, of the first portfolio at a confidence level of 1%. We can notice that, given the same copula, conditional distributions choice has an influence on portfolio Value at Risk. Generally, we can notice that Value at Risk computed by Student t-marginal option tends to be generally higher than the one computed with Normal option, independently of asymmetry. This can be substantially due to longer tails that Student t-distribution has, with respect to Normal distribution. We do not find a substantial, constant difference between Student t-margins with asymmetry option or not. However, we can notice that asymmetric normal conditional distributions, if compared with non- asymmetric normal conditional distributions, given the same copula, tend to give lower Value at Risk estimates. This is due to the presence of negative asymmetry, that drags the distribution to the left, showing a higher density function in the part of graph that corresponds to poor portfolio results. It is possible to check that this situation, that, resuming, copulas with Student t-margins release lower Value at Risk estimates than those with Normal margins, and that asymmetry into normal margins has an influence as well, is present in all tables, independently of the type of copula, portfolio or confidence level we choose. After having taken conditional distributions into consideration, we can pass to copula choice. Here, it is necessary to distinguish between confidence levels, as Value at Risk estimates with low confidence levels, like 1% or 5%, release results that confirm what we said previously into the theory, whereas estimates of every portfolio, at confidence level of 10%, seem to be more blurred. Given conditional distributions choice, we can verify that, at a confidence level of 1%, the lowest Value at Risk is always the one estimated with a Clayton copula. This gets along with what was written into the theory and with what we previously observed, that Clayton copula is the one that, among all elliptical and Archimedean copulas that we analysed, puts more weights on the negative tail, that means that it gives more probability to joint negative events. Having calculated a low quantile, it was to be expected that this copula estimate is dragged down by the higher probability that its distribution puts on the left part of the density function. After Clayton copula, we can check that immediately lower Value at Risk estimate is the one of Student t-copula with Student t-conditional distributions, independently whether they are asymmetric or not. Mind that even for Clayton copula we referred to the case with Student t-margins. After Student t-copula, immediately lower Value at Risk is the one of Gumbel copula, with Student t-margins as well. The fact that an elliptical copula, as the t-Student is, releases lower, so riskier, estimates than an Archimedean copula, should be explained by the particular shape of a Gumbel copula. In fact, while an elliptical copula attributes the same weights both to positive and negative joint extreme events, with respect to Normal copula, Student t-copula gives a positive probability to the extremes. By contrast, a Gumbel copula is more optimistic, and attributes a positive weight only to joint extreme positive events, while it has zero negative tail dependence. Probably, if we calculated Value at Risk at a confidence level of 95%, this would appear in a Gumbel copula estimate higher than Clayton or Student. However, here we are in lowest quantiles, and the special characteristics of Gumbel copula do not appear. Higher than Gumbel, there is Value at Risk estimate of Frank copula that, as we said before, gives no additional weight neither to lower nor upper tail dependence. Last and highest value is the one given by Normal copula, with normal conditional distributions, independently whether with asymmetry or not. This is due to the conjuncture that both the copula and the margins give no weight to extreme events. Now, we have referred to the most general cases, Normal copula with Normal margins and Archimedean copula with Student margins, to illustrate results we got. However, if we focus on all possible combinations, like Archimedean copula or Student copulas with Normal margins, we can get an entire range of possible intermediate cases and estimates. If we refer to a confidence level of 5%, we can check that Clayton copula with Student t-conditional distributions remains the one that release lowest Value at Risk estimates. Immediately after we can find Student t-copula with Student t-conditional distributions. Then, however, Gumbel and Frank copula are reversed, and we can see that Frank copula tends to release lower Value at Risk results than Gumbel copula. Last, as before, is Normal copula with Normal margins. Even in case of a confidence level of 5%, we can verify that all substantial traditional copulas assumptions are respected: Clayton copula, giving more weight to negative extreme, offers lower Value at Risk estimates. Immediately after there is Student t-copula that, while being an elliptical copula, attributes anyway a positive probability on both extremes. Higher estimates are to be referred to copulas that do not encompass joint negative extreme events and, so, whose density functions are not dragged forwards the left. Last, as usual, there is always traditional Normal copula with Normal margins, whose density function is mostly concentrated around the media. Among this reference case, there is a long range of intermediate levels, given by various couples between copula and margins. However, we wanted essentially to verify that the footholds follow the theory. However, when we refer to portfolio estimate of Value at Risk at a confidence level of 10%, all these considerations cannot be made. As we are not more into the negative extreme of the density function, copula characteristics do not appear strongly as before. Value at Risk estimate for a Clayton copula with Student t-conditional distributions is now higher than the one of Frank and Student copula, while Normal copula with Normal margins remains anyway the last. These general traits we have displayed for the first portfolio, the one with equal weights, remain substantially true for all portfolios we analyse. However, we cannot individuate a difference into Value at Risk estimates, given higher or lower variance of the portfolio. Probably, this depends to the fact that portfolios variances are not so far between then. If we check again Spearman's rho matrices, we can verify that many couples of assets have a relatively strong positive dependence. This means that, given the condition that we want to create a portfolio with all our four assets, we cannot reduce variance under a certain threshold. If we want to include all assets, there will be always a couple that will enhance variance. Theoretically, we could have tried to elaborate much lower or much higher variance portfolio by including only three, or even two assets in our portfolio. However, this would have been far from the scope of our analysis, as we aimed at analysing copula as a dependence instrument and elaborate Value at Risk of a portfolio represented by a copula with more than two conditional distributions. Portfolio optimization was not among the aims of our analysis.

APPENDIX A: ARMA(4,0)- GARCH(1,1) TEST

To begin with, we expose the codes with which we create the GARCH(1,1)- ARMA(4,0) model for FTSE MIB, we fit it to our time series of returns and we ask for the estimated parameters:

```
test<-
ugarchspec(variance.model=list(model="sGARCH",garchOrder=c(1,1)),
mean.model=list(armaOrder=c(4,0)))
FTSE_test<-ugarchfit(test,FTSE_rts)
FTSE_test
```

	Estimate	Std. Error	t value	Pr(>|t|)
Mu	0.001239	0.000784	1.5792	0.114300
ar1	-0.036978	0.033482	-1.1044	0.269410
ar2	0.034072	0.033169	1.0272	0.304326
ar3	0.033697	0.033097	1.0181	0.308615
ar4	-0.051910	0.032258	-1.6092	0.107566
omega	0.000018	0.000008	2.2330	0.025548
alpha1	0.132290	0.025017	5.2880	0.000000
beta1	0.859380	0.025112	34.2213	0.000000

In the first command, we set the model: by "sGARCH" we ask the software to apply the standard GARCH model, of which we specify the order (1,1). Then we ask for an ARMA, where the autoregressive part is of order 4, while the moving average is 0, as previously specified. Lastly, we would have to choose a conditional distribution to fit our data from the followings, offered by R: "norm", "std", "snorm" and "sstd". The first two are the classic ones: normal distribution and Student t-distribution respectively. "Snorm" stays for an asymmetric normal distribution, while "sstd" indicates, logically, an asymmetric Student t-distribution. As we do not want that the choice of conditional distribution has some influence on eventual autocorrelation in returns, we are not to indicate distribution model option in this test. We will specify it later for GARCH(1,1)- ARMA(0,0). As it is possible to check from the second command, we fit FTSE MIB data to GARCH ARMA model, and then we ask the software to show estimated parameters, that are displayed in the table. Mu indicates the expected value of time series returns, that corresponds to zero with no surprise. Ar1, ar2, ar3 and ar4 are the estimated coefficients of autoregression in returns. Ar1 says how much return in time t depends on return in time t-1, and so forth. As it is possible to check in the last column, all p-value are higher than 0,05. As a conclusion, we do not refuse null hypothesis for all ARMA(4,0) parameters. As null hypothesis stands for parameter not being significant, we can deduce that, effectively, autocorrelograms of returns have been influenced by autocorrelograms of absolute returns and of returns at square root, as well as in Ljung- Box tests. We can conclude that there is no autocorrelation in returns: they are independently distributed.

Now we perform the same test even for IBEX, proving that even IBEX returns are independently distributed:

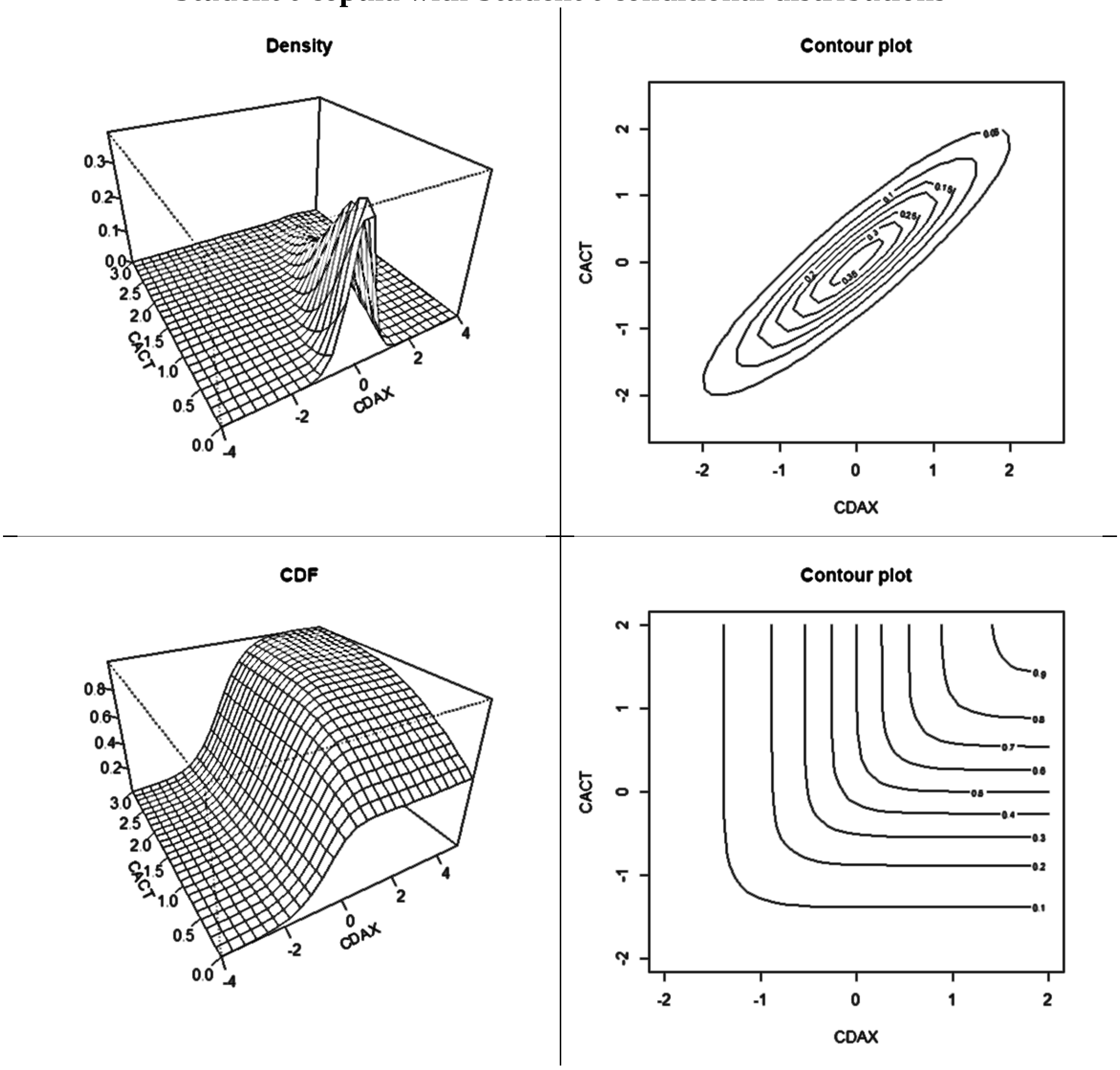
```
IBEX_test<-ugarchfit(test,IBEX_rts)
IBEX_test
```

	Estimate	Std. Error	t value	Pr(>|t|)
mu	0.001529	0.000848	1.80198	0.071548
ar1	-0.060252	0.034600	-1.74139	0.081616
ar2	-0.032111	0.034077	-0.94231	0.346036
ar3	0.045248	0.033914	1.33422	0.182130
ar4	-0.005970	0.032630	-0.18297	0.854819
omega	0.000070	0.000030	2.33026	0.019793
alpha1	0.140299	0.035791	3.91990	0.000089
beta1	0.801911	0.054903	14.60596	0.000000

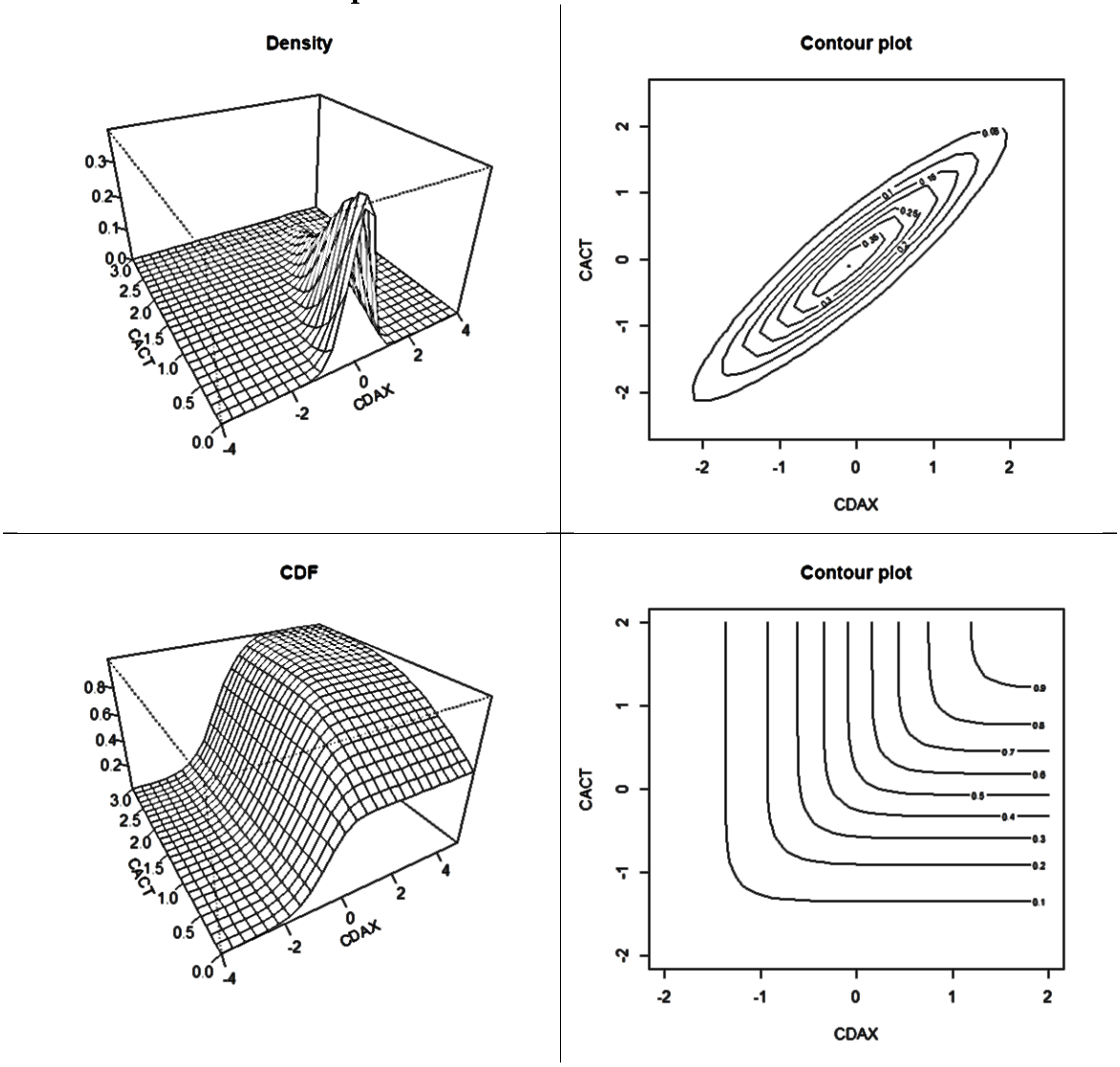
APPENDIX B: GRAPHICAL REPRESENTATIONS

Now that we have estimated Value at Risk, we can spend last pages of this work by enjoying some graphical representations. While before, into the theory, we displayed copulas represented by other authors, now we are to choose two of our assets, like CDAX and CACT, as they are quite strongly entangled and so we will be able to see joint dependence more clearly. We will take only two assets out of four as the software cannot handle graphically more. We will construct copulas with these two assets, displaying for each density function, cumulative density function and level curves. We will refer to all five copulas we coped with in our analysis. For each copula we will realise both the case with Student t-conditional distributions and the one for Normal margins. This will be done even for the heterodox case: Student t-copula with Normal margins and vice versa, just for analysis and curiosity's sake. Every time we will describe which copula we are to treat in the title. Finally, we are to start:

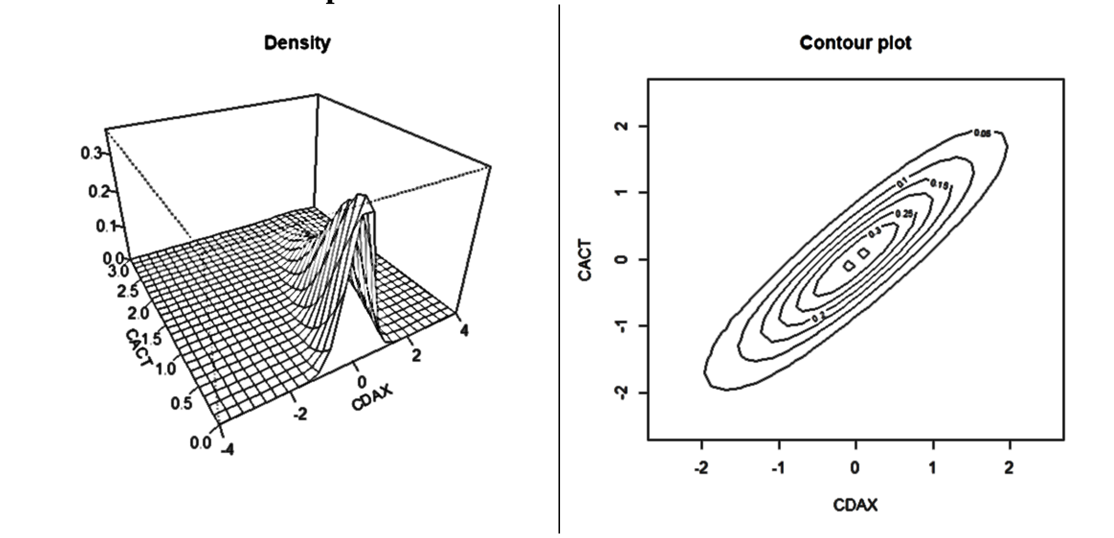
Student t-copula with Student t-conditional distributions



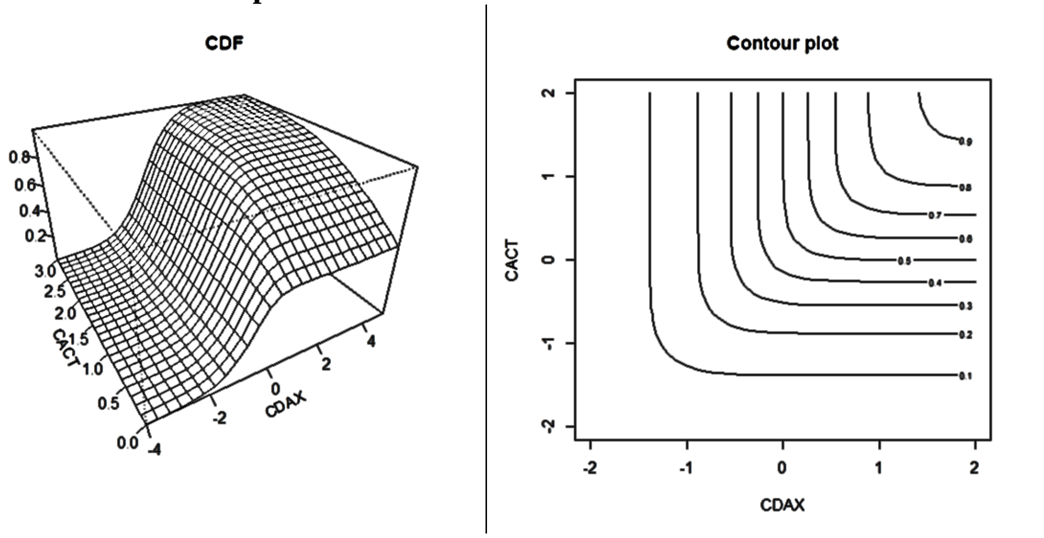
Student t-copula with Normal conditional distributions



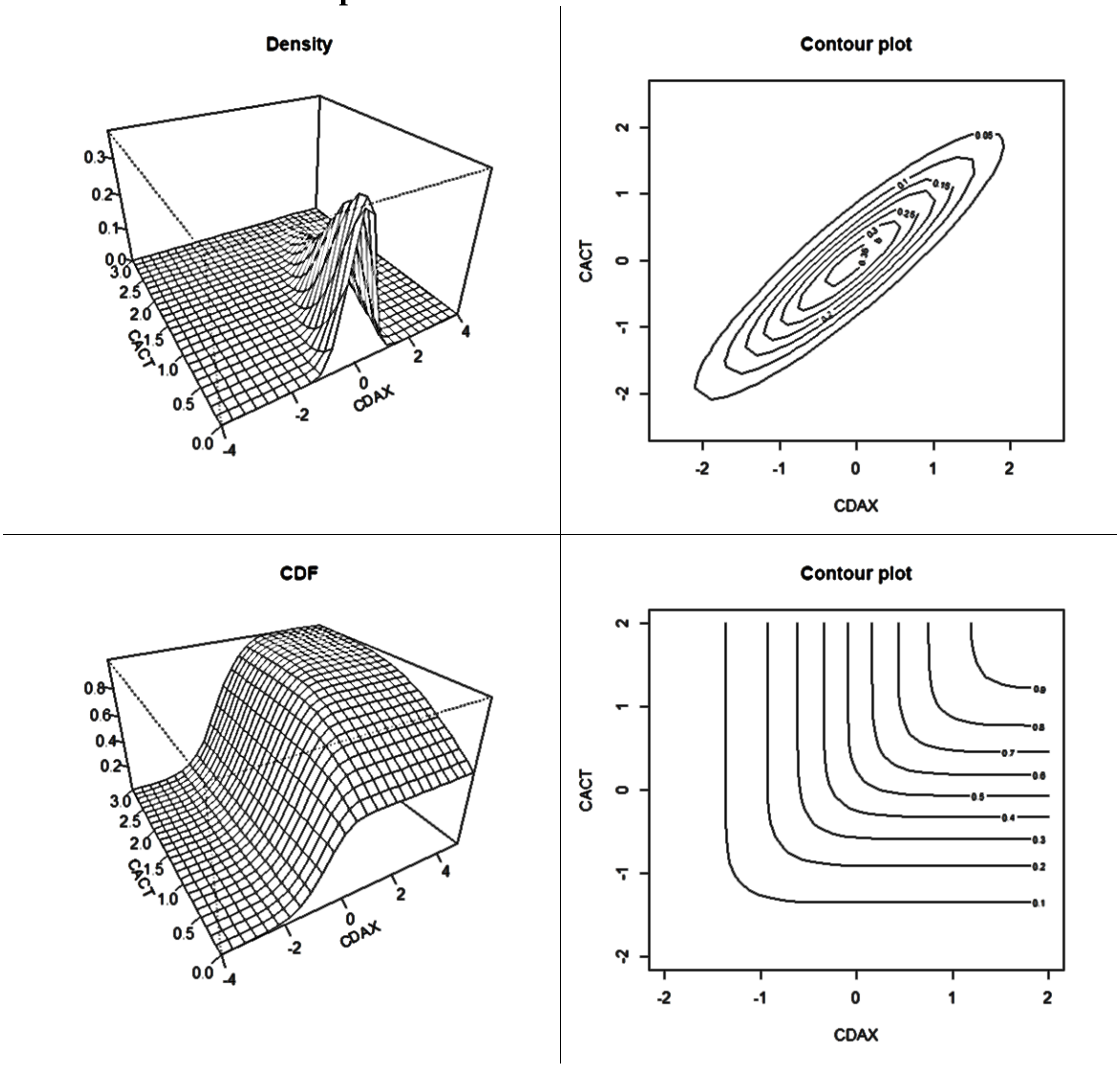
Normal copula with Student t-conditional distributions



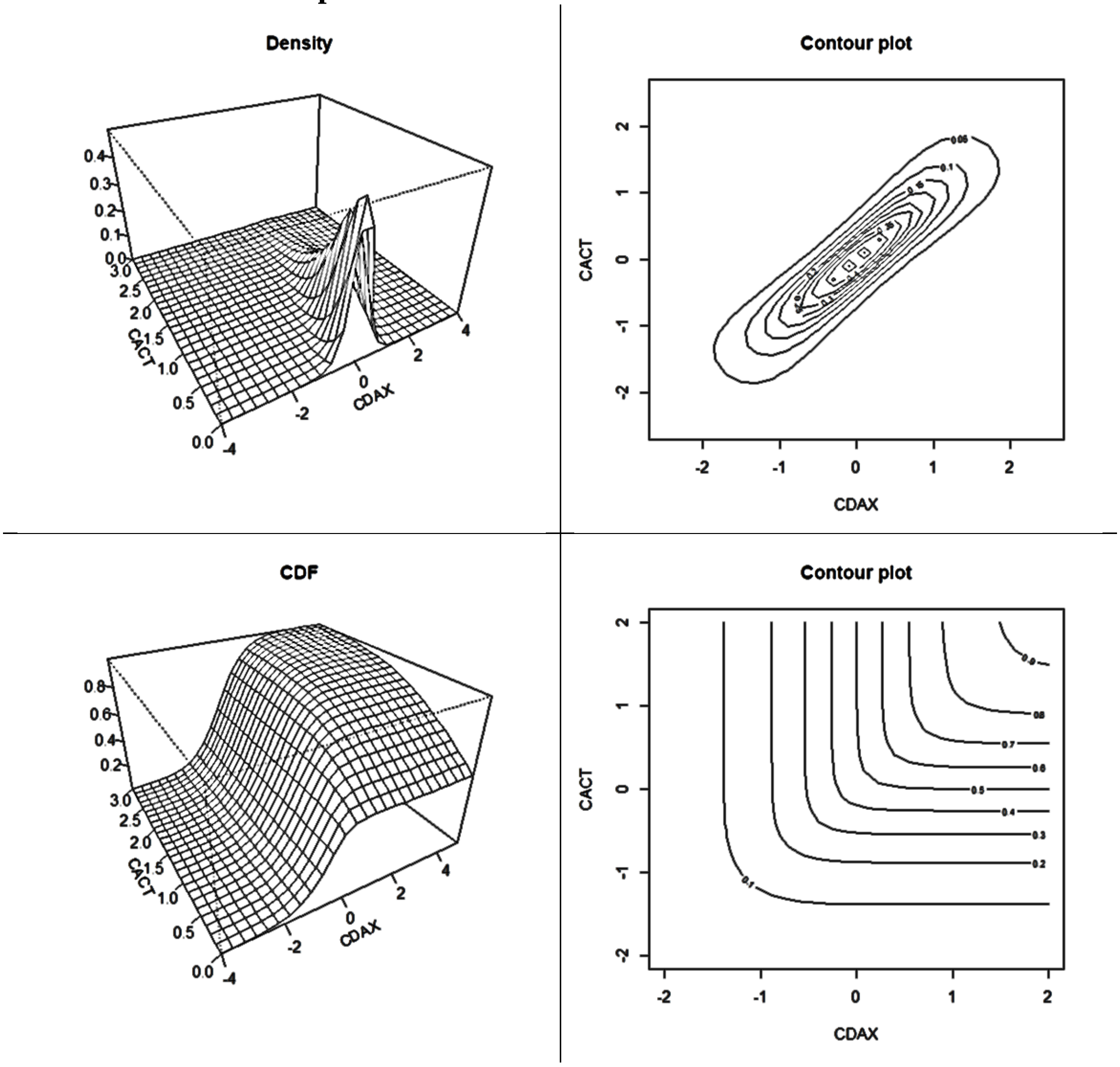
Normal copula with Student t-conditional distributions



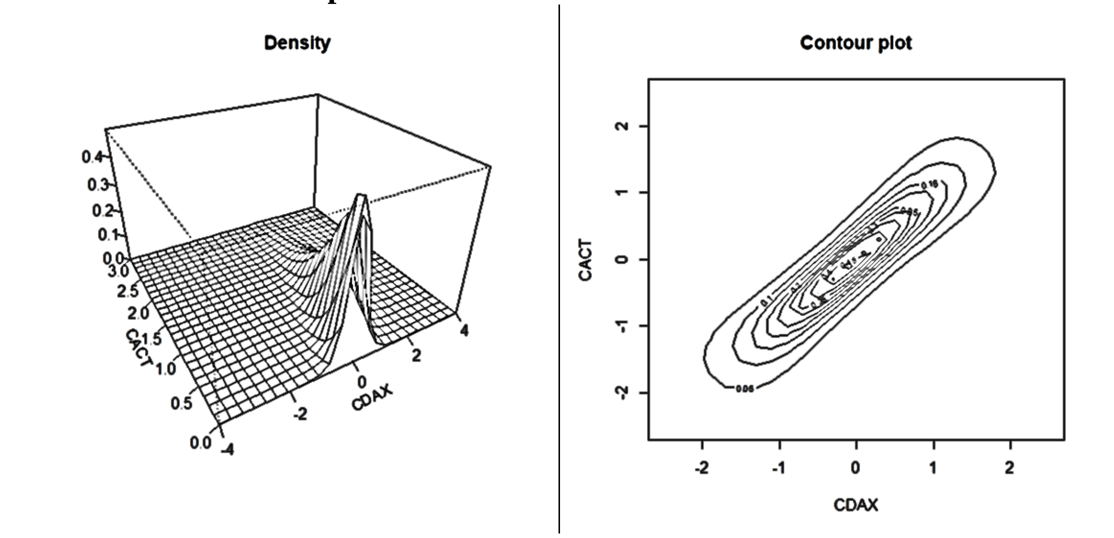
Normal copula with Normal conditional distributions



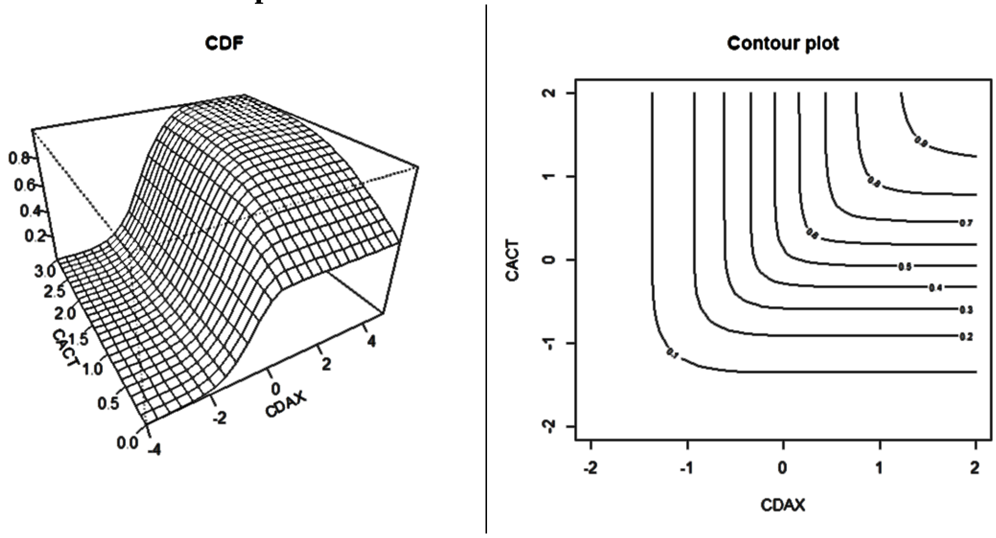
Frank copula with Student t-conditional distributions



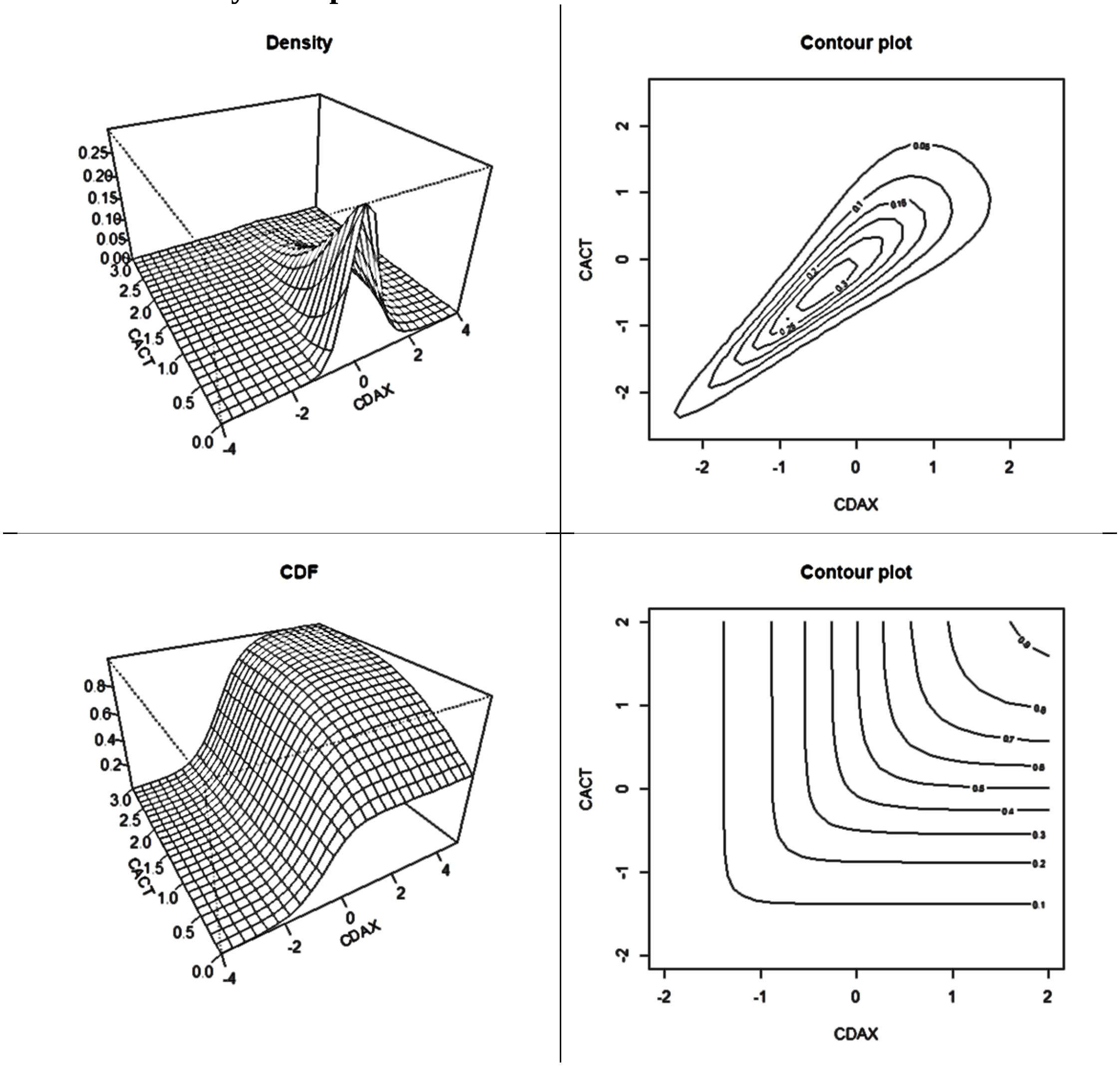
Frank copula with Normal conditional distributions



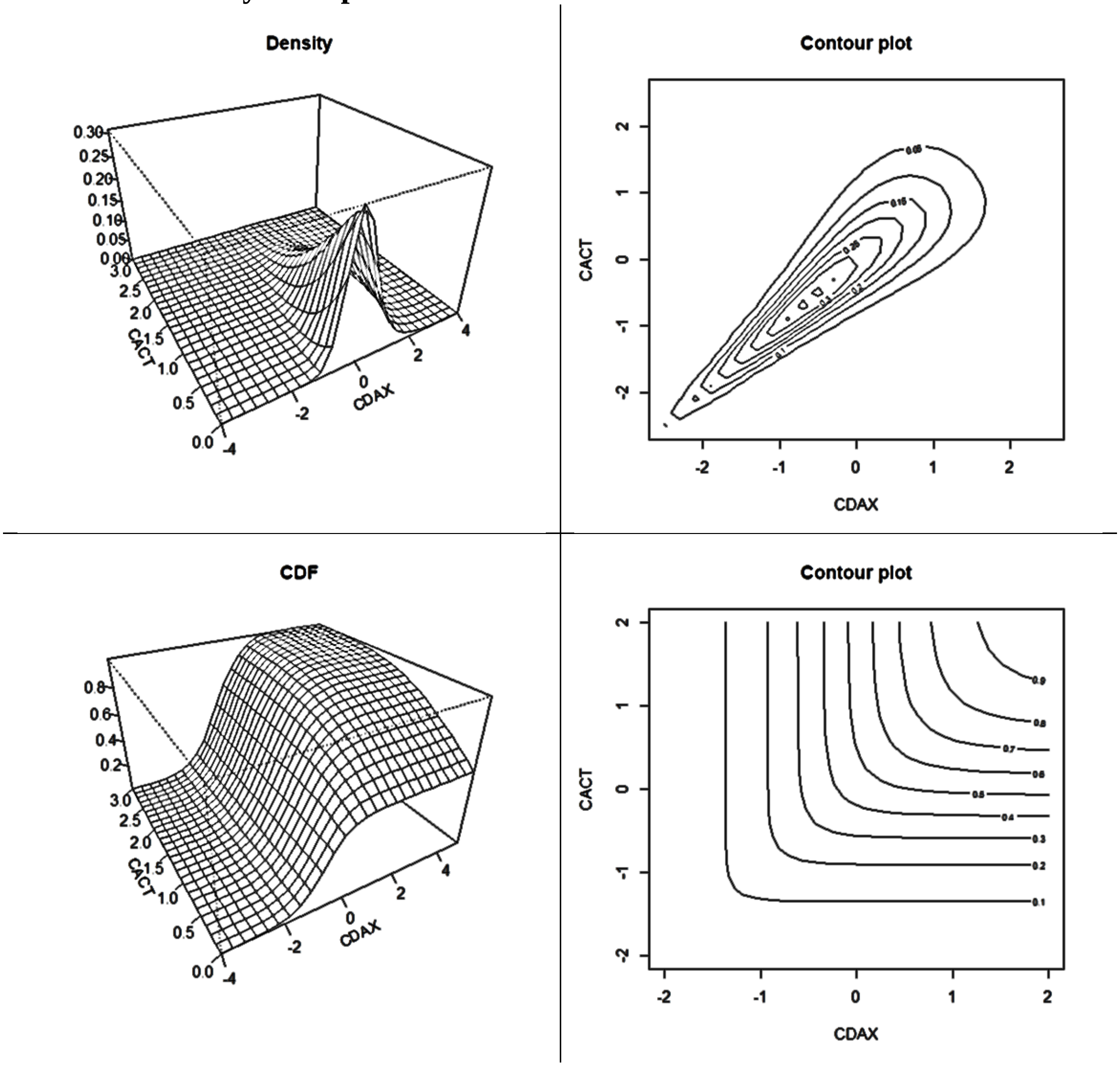
Frank copula with Normal conditional distributions



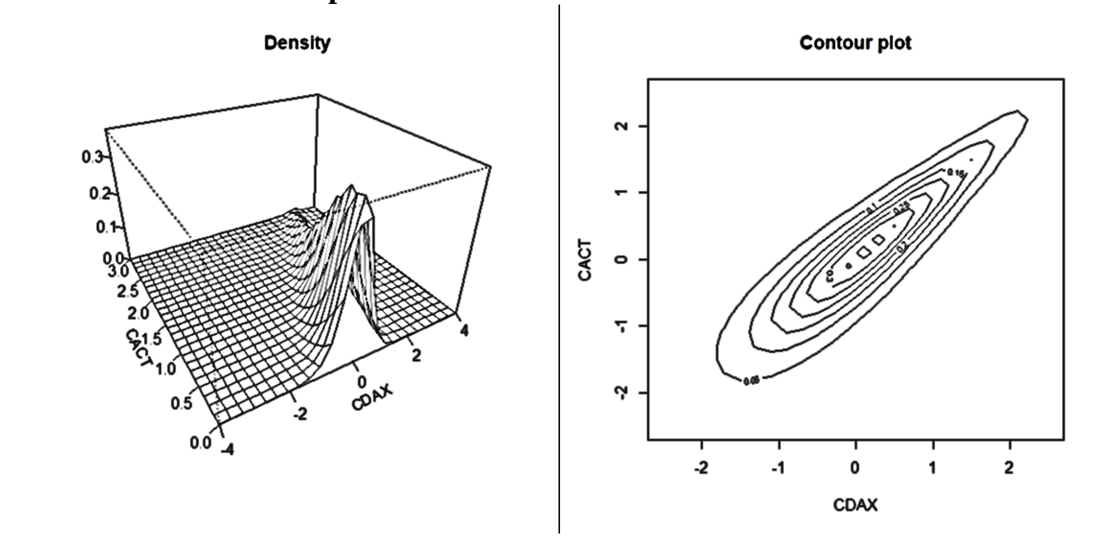
Clayton copula with Student t-conditional distributions



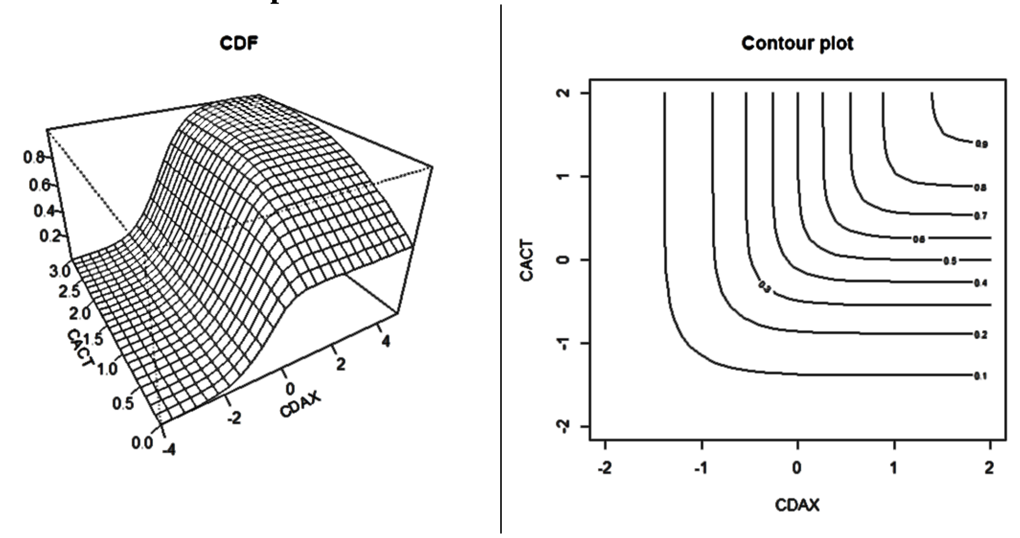
Clayton copula with Normal conditional distributions



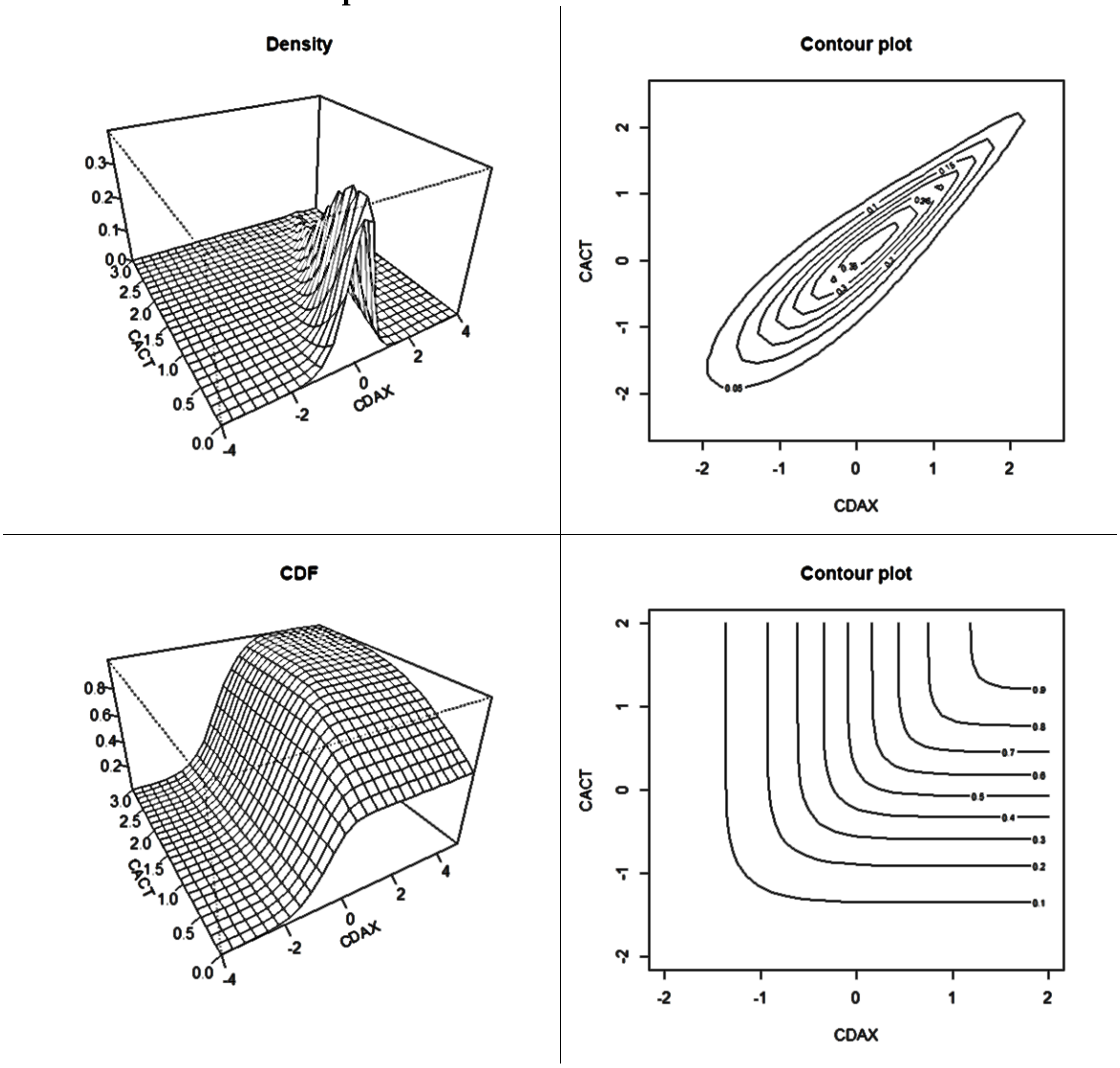
Gumbel copula with Student t-conditional distributions



Gumbel copula with Student t-conditional distributions



Gumbel copula with Normal conditional distributions



APPENDIX C: R-CODE

```
# Install and charge package for reading excel file:
install.packages("xlsx")
library(xlsx)

# Import data:
setwd("C:/Users/HP/Desktop/Tesi/Work in progress")
FTSE_p<-
read.xlsx(file="Indici-dati settimanali.xlsx",sheetName="FTSE MIB")
# As we use same commands for all indices, we are to present only
the case of FTSE MIB, for practicality purpose.

# Convert data frames into time series:
FTSE_pts<-ts(FTSE_p[,2])
plot(FTSE_pts,main="FTSE MIB")

# Carve out logarithmic returns and plot the graph of returns time
series:
LogFTSE_p<-log(FTSE_p[,2])
FTSE_r<-diff(LogFTSE_p,1)
FTSE_rts<-ts(FTSE_r)
plot(FTSE_rts,main="FTSE MIB")

# Some descriptive statistics:
mean(FTSE_r)
median(FTSE_r)
sd(FTSE_r)
var(FTSE_r)
install.packages("moments")
library(moments)
skewness(FTSE_r)
kurtosis(FTSE_r)
```

```
# Jarque-Bera Normality test:
install.packages("tseries")
library(tseries)
jarque.bera.test(FTSE_rts) # p-value < 2.2e-16
--------------------------------------------------------------------
# Check auto-correlation in returns, in absolute value returns and
in returns at square root:
absFTSE_rts<-abs(FTSE_rts)
expFTSE_rts<-FTSE_rts^2
acf(FTSE_rts,main="FTSE MIB returns")
acf(absFTSE_rts,main="Absolute FTSE MIB returns")
acf(expFTSE_rts,main="Square root FTSE MIB returns")

# Ljung-Box auto-correlation test:
install.packages("stats")
library(stats)
Box.test(FTSE_rts,type="Ljung-Box")   # p-value = 0.03607
Box.test(absFTSE_rts,type="Ljung-Box")# p-value = 2.032e-14
Box.test(expFTSE_rts,type="Ljung-Box")# p-value < 2.2e-16
--------------------------------------------------------------------
# Elaborate a ARMA(4,0)- GARCH(1,1) model to test auto-regression in
FTSE MIB:
install.packages("rugarch")
library(rugarch)
test<-
ugarchspec(variance.model=list(model="sGARCH",garchOrder=c(1,1)),
mean.model=list(armaOrder=c(4,0)))
FTSE_test<-ugarchfit(test,FTSE_rts)
FTSE_test
# We execute the same test even for IBEX, but not for CDAX and CACT.
--------------------------------------------------------------------
# Create univariate GARCH specification objects:
gspec.ru.sstd<-
ugarchspec(variance.model=list(model="sGARCH",garchOrder=c(1,1)),
```

```
mean.model=list(armaOrder=c(0,0)),distribution.model="sstd")

FTSEgarch.sstd<-ugarchfit(gspec.ru.sstd,FTSE_rts)
FTSEres.sstd<-residuals(FTSEgarch.sstd,standardize=T)

hist(FTSEres.sstd,nclass=20,freq=FALSE)
curve(dt(x,df=8.43),add=T,col="red")
--------------------------------------------------------------------
absFTSEres.sstd<-abs(FTSEres.sstd)
expFTSEres.sstd<-FTSEres.sstd^2

acf(FTSEres.sstd,main="FTSE MIB residuals")
acf(absFTSEres.sstd,main="Absolute FTSE MIB residuals")
acf(expFTSEres.sstd,main="Square root FTSE MIB residuals")
--------------------------------------------------------------------
# Fit the data set to different copula models and estimate copula parameters.
Here we expose the case where all conditional distributions are supposed to be asymmetric Student t-distributions: code for other conditional distributions hypothesis is similar.
install.packages("copula")
library(copula)
Residuals.sstd<-
cbind(FTSEres.sstd,CDAXres.sstd,CACTres.sstd,IBEXres.sstd)

t.cop<-tCopula(dim=4)
m<-pobs(as.matrix(Residuals.sstd))
fitT.sstd<-fitCopula(t.cop,m,method="ml")
coef(fitT.sstd)
t.cop.sstd<-tCopula(0.5731839,dim=4,df=6.3227985)
TailDep_student.sstd<-lambda(t.cop.sstd)
TailDep_student.sstd

normal.cop<-normalCopula(dim=4)
```

```
m<-pobs(as.matrix(Residuals.sstd))
fitN.sstd<-fitCopula(normal.cop,m,method="ml")
coef(fitN.sstd)
normal.cop.sstd<-normalCopula(0.563937,dim=4)
TailDep_normal.sstd<-lambda(normal.cop.sstd)
TailDep_normal.sstd

frank.cop<-frankCopula(dim=4)
m<-pobs(as.matrix(Residuals.sstd))
fitF.sstd<-fitCopula(frank.cop,m,method="ml")
coef(fitF.sstd)
frank.cop.sstd<-frankCopula(3.804465,dim=4)
TailDep_frank.sstd<-lambda(frank.cop.sstd)
TailDep_frank.sstd

clayton.cop<-claytonCopula(dim=4)
m<-pobs(as.matrix(Residuals.sstd))
fitC.sstd<-fitCopula(clayton.cop,m,method="ml")
coef(fitC.sstd)
clayton.cop.sstd<-claytonCopula(0.886848,dim=4)
TailDep_clayton.sstd<-lambda(clayton.cop.sstd)
TailDep_clayton.sstd

gumbel.cop<-gumbelCopula(dim=4)
m<-pobs(as.matrix(Residuals.sstd))
fitG.sstd<-fitCopula(gumbel.cop,m,method="ml")
coef(fitG.sstd)
gumbel.cop.sstd<-gumbelCopula(1.510486,dim=4)
TailDep_gumbel.sstd<-lambda(gumbel.cop.sstd)
TailDep_gumbel.sstd
--------------------------------------------------------------------
Residuals.sstd<-
cbind(FTSEres.sstd,CDAXres.sstd,CACTres.sstd,IBEXres.sstd)
cor(Residuals.sstd,method="kendall")
```

```
cor(Residuals.sstd,method="spearman")
-------------------------------------------------------------------
# Calculate portfolio variance by Spearman's rho:
port_variance<-
(var(FTSEres.sstd)*0.25^2)+(var(CDAXres.sstd)*0.25^2)+
(var(CACTres.sstd)*0.25^2)+(var(IBEXres.sstd)*0.25^2)+
+(2*0.25*0.25*sd(FTSEres.sstd)*sd(CDAXres.sstd)*0.3071783)+
+(2*0.25*0.25*sd(FTSEres.sstd)*sd(CACTres.sstd)*0.3498962)+
+(2*0.25*0.25*sd(CDAXres.sstd)*sd(CACTres.sstd)*0.9116999)+
+(2*0.25*0.25*sd(FTSEres.sstd)*sd(IBEXres.sstd)*0.5040315)+
+(2*0.25*0.25*sd(CDAXres.sstd)*sd(IBEXres.sstd)*0.5926453)+
+(2*0.25*0.25*sd(CACTres.sstd)*sd(IBEXres.sstd)*0.6436635)
port_variance         # 0.6932016
# We only expose portfolio variance for the first portfolio: for the others it is just necessary to change coefficients.
-------------------------------------------------------------------
# Calculate Value at Risk with five different copulas.
# For practicality purpose, we only exhibit asymmetric Student t-conditional distributions case. Cases with other conditional distributions are similarly coded. The unique difference is, when writing conditional distributions parameters in copula structure, that a Student t-conditional distribution requires number of degrees of freedom, while, in case of Normality, we have to insert mean and standard deviation.

# Normal copula, with asymmetric Student t-marginal distributions:
Residuals.sstd<-
cbind(FTSEres.sstd,CDAXres.sstd,CACTres.sstd,IBEXres.sstd)
r<-100000
set.seed(123)
normal.cop_sstd<-
mvdc(normalCopula(param=0.563937,dim=4),margins=c("t","t","t","t"),
paramMargins=list(list(df=8),list(10),list(df=11),list(df=7)))
```

```
Sim_normal_sstd<-rMvdc(r,normal.cop_sstd)
alfa<-c(0.01,0.05,0.10)

FTSE_normal.cop_sstd<-(Sim_normal_sstd[,1])
CDAX_normal.cop_sstd<-(Sim_normal_sstd[,2])
CACT_normal.cop_sstd<-(Sim_normal_sstd[,3])
IBEX_normal.cop_sstd<-(Sim_normal_sstd[,4])

portfolio_N1<-
0.25*FTSE_normal.cop_sstd+0.25*CDAX_normal.cop_sstd+
0.25*CACT_normal.cop_sstd+0.25*IBEX_normal.cop_sstd
quantile(portfolio_N1,alfa)
# From now on, we show the code only for the first portfolio: for
other portfolios it is just necessary to change weights.
----------------------------------------------------------------------
# Student t-copula, with asymmetric Student t-conditional
distributions:
t.cop_sstd<-
mvdc(tCopula(param=0.5731839,df=6,dim=4),margins=c("t","t","t","t"),
paramMargins=list(list(df=8),list(10),list(df=11),list(df=7)))

Sim_student_sstd<-rMvdc(r,t.cop_sstd)
alfa<-c(0.01,0.05,0.10)

FTSE_student.cop_sstd<-(Sim_student_sstd[,1])
CDAX_student.cop_sstd<-(Sim_student_sstd[,2])
CACT_student.cop_sstd<-(Sim_student_sstd[,3])
IBEX_student.cop_sstd<-(Sim_student_sstd[,4])

portfolio_N1<-
0.25*FTSE_student.cop_sstd+0.25*CDAX_student.cop_sstd+
+0.25*CACT_student.cop_sstd+0.25*IBEX_student.cop_sstd
quantile(portfolio_N1,alfa)
# Frank copula, with asymmetric Student t-conditional distributions:
```

```
frank.cop_sstd<-
mvdc(frankCopula(param=3.804465,dim=4),margins=c("t","t","t","t"),
paramMargins=list(list(df=8),list(10),list(df=11),list(df=7)))

Sim_frank_sstd<-rMvdc(r,frank.cop_sstd)
alfa<-c(0.01,0.05,0.10)

FTSE_frank.cop_sstd<-(Sim_frank_sstd[,1])
CDAX_frank.cop_sstd<-(Sim_frank_sstd[,2])
CACT_frank.cop_sstd<-(Sim_frank_sstd[,3])
IBEX_frank.cop_sstd<-(Sim_frank_sstd[,4])

portfolio_N1<-
0.25*FTSE_frank.cop_sstd+0.25*CDAX_frank.cop_sstd+
0.25*CACT_frank.cop_sstd+0.25*IBEX_frank.cop_sstd
quantile(portfolio_N1,alfa)
--------------------------------------------------------------------
# Clayton copula, with asymmetric Student t-conditional
distributions:
clayton.cop_sstd<-
mvdc(claytonCopula(param=0.886848,dim=4),margins=c("t","t","t","t"),
paramMargins=list(list(df=8),list(10),list(df=11),list(df=7)))

Sim_clayton_sstd<-rMvdc(r,clayton.cop_sstd)
alfa<-c(0.01,0.05,0.10)

FTSE_clayton.cop_sstd<-(Sim_clayton_sstd[,1])
CDAX_clayton.cop_sstd<-(Sim_clayton_sstd[,2])
CACT_clayton.cop_sstd<-(Sim_clayton_sstd[,3])
IBEX_clayton.cop_sstd<-(Sim_clayton_sstd[,4])

portfolio_N1<-
0.25*FTSE_clayton.cop_sstd+0.25*CDAX_clayton.cop_sstd+
+0.25*CACT_clayton.cop_sstd+0.25*IBEX_clayton.cop_sstd
```

```
quantile(portfolio_N1,alfa)
--------------------------------------------------------------------
# Gumbel copula, with asymmetric Student t-conditional
distributions:
gumbel.cop_sstd<-
mvdc(gumbelCopula(param=1.510486,dim=4),margins=c("t","t","t","t"),
paramMargins=list(list(df=8),list(10),list(df=11),list(df=7)))

Sim_gumbel_sstd<-rMvdc(r,gumbel.cop_sstd)
alfa<-c(0.01,0.05,0.10)

FTSE_gumbel.cop_sstd<-(Sim_gumbel_sstd[,1])
CDAX_gumbel.cop_sstd<-(Sim_gumbel_sstd[,2])
CACT_gumbel.cop_sstd<-(Sim_gumbel_sstd[,3])
IBEX_gumbel.cop_sstd<-(Sim_gumbel_sstd[,4])

portfolio_N1<-
0.25*FTSE_gumbel.cop_sstd+0.25*CDAX_gumbel.cop_sstd+
+0.25*CACT_gumbel.cop_sstd+0.25*IBEX_gumbel.cop_sstd
quantile(portfolio_N1,alfa)
--------------------------------------------------------------------
# Graphical representations of copulas built on CDAX and CACT:
my_data1<-cbind(CDAXres.std,CACTres.std)
var_a<-pobs(my_data1)[,1]
var_b<-pobs(my_data1)[,2]

# By fitting our data to bivariate copula structure, we estimate
different copula parameters. We expose the case in which conditional
distributions are t-Student: for Normal conditional distributions
only marginal parameters change.

t.cop1<-tCopula(dim=2)
k<-pobs(as.matrix(my_data1))
fitT.1<-fitCopula(t.cop1,k,method="ml")
```

```
coef(fitT.1)
tau(tCopula(param=0.9134227,df=12))   # 0.7331413

normal.cop1<-normalCopula(dim=2)
k<-pobs(as.matrix(my_data1))
fitN.1<-fitCopula(normal.cop1,k,method="ml")
coef(fitN.1)
tau(normalCopula(param=0.9120054))    # 0.730933

frank.cop1<-frankCopula(dim=2)
k<-pobs(as.matrix(my_data1))
fitF.1<-fitCopula(frank.cop1,k,method="ml")
coef(fitF.1)
tau(frankCopula(param=13.5122))       # 0.7400084

clayton.cop1<-claytonCopula(dim=2)
k<-pobs(as.matrix(my_data1))
fitC.1<-fitCopula(clayton.cop1,k,method="ml")
coef(fitC.1)
tau(claytonCopula(param=3.527801))    # 0.6381925

gumbel.cop1<-gumbelCopula(dim=2)
k<-pobs(as.matrix(my_data1))
fitG.1<-fitCopula(gumbel.cop1,k,method="ml")
coef(fitG.1)
tau(gumbelCopula(param=3.374522))     # 0.7036617
--------------------------------------------------------------------
T_dist1<-
mvdc(tCopula(param=0.9134227,dim=2,df=12),margins=c("t","t"),
paramMargins=list(list(df=9),list(df=9)))
v<-rMvdc(5000,T_dist1)
pdf_mvd<-dMvdc(v,T_dist1)
cdf_mvd<-pMvdc(v,T_dist1)
install.packages("scatterplot3d")
```

```
library(scatterplot3d)

persp(T_dist1,dMvdc,xlim=c(-
4,4),ylim=c(0,3),main="Density",xlab="CDAX", ylab="CACT",zlab=" ")
contour(T_dist1,dMvdc,xlim=c(-2.5,2.5),ylim=c(-
2.5,2.5),main="Contour plot",
xlab="CDAX",ylab="CACT")
persp(T_dist1,pMvdc,xlim=c(-4,5),ylim=c(0,3),main="CDF",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(T_dist1,pMvdc,xlim=c(-2,2),ylim=c(-2,2),main="Contour plot",
xlab="CDAX",ylab="CACT")
--------------------------------------------------------------------
N_dist1<-
mvdc(normalCopula(param=0.9120054,dim=2),margins=c("t","t"),
paramMargins=list(list(df=9),list(df=9)))
v<-rMvdc(5000,N_dist1)
pdf_mvd<-dMvdc(v,N_dist1)
cdf_mvd<-pMvdc(v,N_dist1)

persp(N_dist1,dMvdc,xlim=c(-4,4),ylim=c(0,3),main="Density",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(N_dist1,dMvdc,xlim=c(-2.5,2.5),ylim=c(-
2.5,2.5),main="Contour plot",
xlab="CDAX",ylab="CACT")
persp(N_dist1,pMvdc,xlim=c(-4,5),ylim=c(0,3),main="CDF",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(N_dist1,pMvdc,xlim=c(-2,2),ylim=c(-2,2),main="Contour  plot",
xlab="CDAX",ylab="CACT")
--------------------------------------------------------------------
F_dist1<-mvdc(frankCopula(param=13.5122,dim=2),margins=c("t","t"),
paramMargins=list(list(df=9),list(df=9)))
v<-rMvdc(5000,F_dist1)
pdf_mvd<-dMvdc(v,F_dist1)
cdf_mvd<-pMvdc(v,F_dist1)
```

```
persp(F_dist1,dMvdc,xlim=c(-4,4),ylim=c(0,3),main="Density",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(F_dist1,dMvdc,xlim=c(-2.5,2.5),ylim=c(-
2.5,2.5),main="Contour plot",
xlab="CDAX",ylab="CACT")
persp(F_dist1,pMvdc,xlim=c(-4,5),ylim=c(0,3),main="CDF",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(F_dist1,pMvdc,xlim=c(-2,2),ylim=c(-2,2),main="Contour plot",
xlab="CDAX",ylab="CACT")
--------------------------------------------------------------------
C_dist1<-
mvdc(claytonCopula(param=3.527801,dim=2),margins=c("t","t"),
paramMargins=list(list(df=9),list(df=9)))
v<-rMvdc(5000,C_dist1)
pdf_mvd<-dMvdc(v,C_dist1)
cdf_mvd<-pMvdc(v,C_dist1)

persp(C_dist1,dMvdc,xlim=c(-4,4),ylim=c(0,3),main="Density",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(C_dist1,dMvdc,xlim=c(-2.5,2.5),ylim=c(-
2.5,2.5),main="Contour plot",
xlab="CDAX",ylab="CACT")
persp(C_dist1,pMvdc,xlim=c(-4,5),ylim=c(0,3),main="CDF",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(C_dist1,pMvdc,xlim=c(-2,2),ylim=c(-2,2),main="Contour plot",
xlab="CDAX",ylab="CACT")
--------------------------------------------------------------------
G_dist1<-mvdc(gumbelCopula(param=3.374522,dim=2),margins=c("t","t"),
paramMargins=list(list(df=9),list(df=9)))
v<-rMvdc(5000,G_dist1)
pdf_mvd<-dMvdc(v,G_dist1)
cdf_mvd<-pMvdc(v,G_dist1)
```

```
persp(G_dist1,dMvdc,xlim=c(-4,4),ylim=c(0,3),main="Density",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(G_dist1,dMvdc,xlim=c(-2.5,2.5),ylim=c(-
2.5,2.5),main="Contour plot",
xlab="CDAX",ylab="CACT")
persp(G_dist1,pMvdc,xlim=c(-4,5),ylim=c(0,3),main="CDF",
xlab="CDAX",ylab="CACT",zlab=" ")
contour(G_dist1,pMvdc,xlim=c(-2,2),ylim=c(-2,2),main="Contour plot",
xlab="CDAX",ylab="CACT")
```

BIBLIOGRAPHY

K. Aas (2004) “Modelling the dependence structure of financial assets: A survey of four copulas”, Norwegian Computing Center, Applied Research and Development

K. Aas and X. K. Dimakos (2004) “Statistical modelling of financial time series: An introduction”, Nowegian Computing Center, Applied research and development

T. Bollerslev (1986) “Generalized Autoregressive Conditional Heteroskedasticity”, Journal of Econometrics, **31**, 307- 327

U. Cherubini, E. Luciano and W. Vecchiato (2004) “Copula methods in finance”, John Wiley & Sons, New York

R. Engle (1982) “Autoregressive Conditional Heteroskedasticity with Estimates of Variance of United Kingdom Inflation”, Econometrica, **50**, 987- 1008

P. Embrechts (2009) “Copulas: A personal view”, Department of Mathematics, ETH Zurich

P. Embrechts, A.J. McNeil and D. Straumann (1999) “Correlation and Dependency in Risk Management: Properties and Pitfalls”. Department of Mathematik, ETHZ, Zurich, working paper.

Y. Fang and L. Madsen (2013) “Modified Gaussian pseudo-copula: Applications in insurance and finance”, Insurance: Mathematics and Economics 53, 292- 301.

C. Genest and A. Favre (2007) “Everything You Always Wanted to Know about Copula Modelling but Were Afraid to Ask”, McGill University. Montréal, Canada.

C. Genest and J. MacKay (1986) “The joy of copulas: bivariate distributions with uniform marginals”, Amer. Statist., **40**, 280-283.

Olivier Guéant “Computing the Value at Risk of a Portfolio: Academic Literature and Practitioners’ response”

P. Hougaard (1986) “A class of multivariate failure time distributions”, Biometrika, **73**, 671-678.

H. Joe (1997) “Multivariate Models and Dependence Concepts”. Chapman & Hall, London.

H. Joe and J.J. Xu (1996) “The Estimation Method of Inference Functions for Margins for Multivariate Models”, Department of Statistics, University of British Columbia, Vancouver, Canada, Technical report 166.

P. Jorion (1997) “Value at risk”, McGraw- Hill

J.F. Jouanin, G. Rapuch, G. Riboulet and T. Roncalli (2001) “Modelling Dependence for Credit Derivatives with Copulas”. Groupe de Recherche Opèrationelle, Crédit Lyonnais, working paper.

JF. Jouanin, G. Riboulet and T. Roncalli (2011) “Financial Applications of Copula Functions “, Groupe de Recherche Opérationelle, Crédit Lyonnais, France.

C. Kharoubi-Ratokomalala, and F. Maurer (2013) “Copulas in Finance Ten Years Later”, The Journal of Applied Business Research, September/ October 2013, Volume 29, number 5.

C.H. Kimberling (1974) “A probabilistic interpretation of complete monotonicity”, Aequationes Math., **10**, 152- 164.

Thomas J. Linsmeier and Neil D. Pearson (1996) “Risk Measurement: an Introduction to Value at Risk”, University of Illinois at Urbana- Champaign.

R. Lucchetti “Metodo della massima verosimiglianza”

Y. Malevergne and D. Sornette (2001) “Testing the Gaussian Copula Hypothesis for Financial Assets Dependences”, Laboratoire de Physique de la Matière Condensée, Université de Nice-Sophia Antipolis, Institut de Science Financière et d’Assurances, Université Lyon I, Institute of Geophysics and Planetary Physics and Department of Earth and Space Science, University of California, Los Angeles, California.

A.M. Mineo (2003) “Una Guida all’utilizzo dell’Ambiente Statistico R”, Dipartimento di Scienze Statistiche e Matematiche “S. Vianelli”, Università degli Studi di Palermo

J. Myung, (2003) “Tutorial on maximum likelihood estimation”, Journal of Mathematical Psychology, **47**, 90- 100

Roger B. Nelsen (1999) “An introduction to copulas”, Lecture Notes in Statistics. Springer-Verlag, New York.

A. Patton (2007) “Copula- Based Models for Financial Time Series”, Oxford- Man Institute of Quantitative Finance, University of Oxford, working paper.

P. Posedel (2005) “Properties and Estimation of GARCH(1,1) Model”, Metodoloski zvezki, Vol. 2, No. 2, 243- 257.

S.T. Rachev, M. Stein and W. Sun (2015) “Copula Concepts in Financial Markets”, University of Karlsruhe, KIT & University of Santa Barbara & FinAnalytica

V. Ricci (2005) “Analisi delle serie storiche con R”

T. Roncalli (2002) “Gestion des Risques Multiples. Cours ENSAI de 3^{e} année”. Groupe de Recherche Opèrationelle, Crédit Lyonnais, working paper.

T. Schmidt (2006) “Coping with Copulas”, Department of Mathematics, University of Leipzig.

B. Schweizer and E. Wolff (1981) “On non-parametric measures of dependence for random variables”, Ann. Statist., **9**, 879-885.

J. Yan (2007) “Enjoy the Joy of Copulas: with a Package Copula”, Journal of Statistical Software, Volume 21, Issue 4

A. Zilio (2005) “Il metodo delle copule: inferenza, calcolo del VaR ed implementazione in R”, Università degli Studi di Padova.

E. Zivot (2013) “Copulas”, Computational Finance & Risk Management, Department of Applied Mathematics, University of Washington

Databank Eikon, Thomson Reuters Eikon Datastream, Università degli Studi di Padova