# Which copula is the right one?

V. Durrleman, A. Nikeghbali & T. Roncalli*

Groupe de Recherche Opérationnelle
Crédit Lyonnais
France

August 25, 2000

**Abstract**

In this paper, we give a few methods for the choice of copulas in financial modelling.

## 1 Introduction

Copulas reveal to be a very powerful tool in the banking industry, more especially in the modelling of the different sorts of existing risks. It allows a multidimensional framework, giving up the gaussian assumption which we know is incorrect in asset modelling and the study of extremal events. Bouyé, Durrleman, Nikeghbali, Riboulet and Roncalli [2000] review different financial problems and show how copulas could help to solve them. For example, they used copulas for operational risk measurement and the study of multidimensional stress scenarios. But one of the difficulty is in general the choice of the copula. This article deals with this problem.

Let's recall some elementary facts about copulas. All these results are developped in Deheuvels [1981] and Nelsen [1998]. Let $(X_1, \ldots, X_N) \in \mathbb{R}^N$ be a random vector with cumulative distribution function

$$\mathbf{F}(x_1, \ldots, x_N) = \mathbb{P}(X_1 \leq x_1, \ldots, X_N \leq x_N) \tag{1}$$

and marginal functions

$$\mathbf{F}_n(x_n) = \mathbb{P}(X_n \leq x_n), \qquad 1 \leq n \leq N \tag{2}$$

A copula function $\mathbf{C}$ of $\mathbf{F}$ is defined as a cumulative distribution function of a probability measure with support in $[0,1]^N$ such that:

1. $\forall\, 1 \leq n \leq N$, $0 \leq u_n \leq 1$
$$\mathbf{C}_n(u_n) = \mathbf{C}(1, \ldots, 1, u_n, 1, \ldots, 1) = u_n \tag{3}$$

2. $\forall\, (x_1, \ldots, x_N)$ such that $x_n$ is a continuity point of $\mathbf{F}_n$ $(1 \leq n \leq N)$,
$$\mathbf{F}(x_1, \ldots, x_N) = \mathbf{C}(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_N(x_N)) \tag{4}$$

The following results about copulas are very useful. The complete proofs of these results are established in Deheuvels [1981].

* *Corresponding author:* Groupe de Recherche Opérationnelle, Bercy-Expo — Immeuble Bercy Sud — $4^{\text{è}}$ étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France; *E-mail adress:* thierry.roncalli@creditlyonnais.fr

1. Every distribution function $\mathbf{F}$ has at least one copula function, uniquely defined on the image set $(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_N(x_N))$ of the points $(x_1, \ldots, x_N)$ such that $\forall 1 \leq n \leq N$, $x_n$ is a continuity point of $\mathbf{F}_n$. If all the marginal functions are continuous, then the copula function of $\mathbf{F}$ is unique.

2. Every copula $\mathbf{C}$ is continuous and statisfies the following inequality ($\forall 1 \leq n \leq N$, $0 \leq u_n, v_n \leq 1$)

$$\left|\mathbf{C}(u_1, \ldots, u_N) - \mathbf{C}(v_1, \ldots, v_N)\right| \leq \sum_{n=1}^{N} |u_n - v_n| \tag{5}$$

3. The set $\mathcal{C}$ of all copula functions is convex, compact with any of the following topologies, equivalent on $\mathcal{C}$: punctual convergence, uniform convergence on $[0,1]^N$, weak convergence of the associated probability measure (Deheuvels [1978]).

4. If $h_1, \ldots, h_N$ are monotone, non decreasing mappings of $\mathbb{R}$ in itself, any copula function of $(X_1, \ldots, X_N)$ is also a copula function of $(h_1(X_1), \ldots, h_N(X_N))$.

5. If $\left\{\mathbf{F}^{(m)}, m \geq 1\right\}$ is a sequence of probability cumulative distribution functions in $\mathbb{R}^N$, the convergence of $\mathbf{F}^{(m)}$ to a distribution function $\mathbf{F}$ with continous margins $\mathbf{F}_n$, when $m \longrightarrow \infty$, is equivalent to both conditions:

   (a) $\forall 1 \leq n \leq N$, $\mathbf{F}_n^{(m)} \longrightarrow \mathbf{F}_n$ punctually;

   (b) if $\mathbf{C}$ is the unique copula function associated to $\mathbf{F}$, and if $\mathbf{C}^{(m)}$ is a copula function associated to $\mathbf{F}^{(m)}$, $\mathbf{C}^{(m)} \longrightarrow \mathbf{C}$ (with the topology of $\mathcal{C}$).

So it appears that copulas are in fact the dependence structure of the model. All the information about the dependence is contained in the copula function. Thus the choice of the copula that is going to fit the data is very important. Usually, one takes a parametric family of copulas among many existing others and fit it to the data by estimating the parameters of the family. Nevertheless, there does not exist a systematic rigorous method for the choice of copulas: nothing can tell us that the selected family of copula will converge to the real structure dependence underlying the data. This can provide biased results since according to the dependence structure selected the obtained results might be very different.

What we propose in this paper is different methods to cope with the uncertainty about the real underlying dependence structure of the studied phenomena. In a second section, we present the parametric estimation with a given family. Then, we give several methods to choose the 'optimal' copula, all based on the empirical copula introduced be Deheuvels [1979].

# 2 Parametric estimation with a given family

## 2.1 The method of the maximum likelihood

Let us first consider the case where both the copula $\mathbf{C}$ and the margins $\mathbf{F}_n$ are continuous. The density of the joint distribution $\mathbf{F}$ is given by the following expression

$$f(x_1, \ldots, x_n, \ldots, x_N) = c(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_n(x_n), \ldots, \mathbf{F}_N(x_N)) \prod_{n=1}^{N} f_n(x_n) \tag{6}$$

where $f_n$ is the density of the margin $\mathbf{F}_n$ and $c$ is the density of the copula

$$c(u_1, \ldots, u_n, \ldots, u_N) = \frac{\partial\, \mathbf{C}(u_1, \ldots, u_n, \ldots, u_N)}{\partial\, u_1 \cdots \partial\, u_n \cdots \partial\, u_N} \tag{7}$$

Let $\mathcal{X} = \{(x_1^t, \ldots, x_N^t)\}_{t=1}^T$ denote a sample. The expression of the log-likelihood is also

$$\ell(\theta) = \sum_{t=1}^{T} \ln c\left(\mathbf{F}_1\left(x_1^t\right), \ldots, \mathbf{F}_n\left(x_n^t\right), \ldots, \mathbf{F}_N\left(x_N^t\right)\right) + \sum_{t=1}^{T}\sum_{n=1}^{N} \ln f_n\left(x_n^t\right) \tag{8}$$

with $\theta$ the $K \times 1$ vector of parameters[1]. Let $\hat{\theta}_{\mathrm{ML}}$ be the maximum likelihood estimator. Then, it verifies the property of **asymptotic normality** and we have

$$\sqrt{T}\left(\hat{\theta}_{\mathrm{ML}} - \theta_0\right) \longrightarrow \mathcal{N}\left(\mathbf{0}, \mathcal{J}^{-1}(\theta_0)\right) \tag{9}$$

with $\mathcal{J}(\theta_0)$ the **information matrix** of Fisher (Davidson and MacKinnon [1993]). If the margins are discrete (the support is $\{x_n^-, \ldots, x_n^+\}$, the Radon-Nikodym density is[2]

$$\begin{aligned}\Pr\{(X_1, \ldots, X_N) = (x_1, \ldots, x_N)\} = \\ (-1)^N \sum_{i_1=0}^{1} \cdots \sum_{i_N=0}^{1} (-1)^{i_1+\cdots+i_N} \mathbf{C}\left(\mathbf{F}_1(x_1 - i_1), \ldots, \mathbf{F}_n(x_n - i_n), \ldots, \mathbf{F}_N(x_N - i_N)\right)\end{aligned} \tag{11}$$

The expression of the log-likelihood becomes

$$\ell(\theta) = \sum_{t=1}^{T} \ln \Pr\left\{(X_1, \ldots, X_N) = \left(x_1^t, \ldots, x_N^t\right)\right\} \tag{12}$$

In order to clarify the underlying methods and to help the reader to reproduce the results, we use the LME database[3]. We consider the joint distribution of the asset returns Aluminium Alloy (AL) and Copper (CU). We assume that the margins are gaussians and that the dependence structure is the Frank copula:

$$\mathbf{C}(u_1, u_2) = -\alpha^{-1} \ln\left(\frac{1}{1 - e^{-\alpha}}\left[\left(1 - e^{-\alpha}\right) - \left(1 - e^{-\alpha u_1}\right)\left(1 - e^{-\alpha u_2}\right)\right]\right) \tag{13}$$

The log-likelihood of the observation $t$ is then

$$\begin{aligned}\ell_t(\theta) = {} & \ln\left[\alpha\left(1 - \exp(-\alpha)\right)\exp\left(-\alpha\left(\Phi\left(\frac{x_1^t - m_1}{\sigma_1}\right) + \Phi\left(\frac{x_2^t - m_2}{\sigma_2}\right)\right)\right)\right] - \\ & \ln\left[\left(1 - \exp(-\alpha)\right) - \left(1 - \exp\left(-\alpha\Phi\left(\frac{x_1^t - m_1}{\sigma_1}\right)\right)\right)\left(1 - \exp\left(-\alpha\Phi\left(\frac{x_2^t - m_2}{\sigma_2}\right)\right)\right)\right]^2 + \\ & \left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_1^2\right) - \frac{1}{2}\left(\frac{x_1^t - m_1}{\sigma_1}\right)^2\right] + \left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_2^2\right) - \frac{1}{2}\left(\frac{x_2^t - m_2}{\sigma_2}\right)^2\right]\end{aligned} \tag{14}$$

We then obtain the following results

| Parameters | estimate | std.err. | t-statistic | p-value |
|---|---|---|---|---|
| $m_1$ | -0.000247 | 0.000258 | -0.95 | 0.34 |
| $m_2$ | 0.013700 | 0.000192 | 71.5 | 0.00 |
| $\sigma_1$ | -0.000210 | 0.000324 | -0.65 | 0.52 |
| $\sigma_2$ | 0.017175 | 0.000241 | 71.4 | 0.00 |
| $\alpha$ | 4.597625 | 0.158681 | 28.9 | 0.00 |

[1] $\theta$ is the set of parameters of the margins and the copula.

[2] If $N = 2$, we retreive the result

$$\begin{aligned}\Pr\{(X_1, X_2) = (x_1, x_2)\} = {} & \mathbf{C}\left(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)\right) - \mathbf{C}\left(\mathbf{F}_1(x_1 - 1), \mathbf{F}_2(x_2)\right) - \\ & \mathbf{C}\left(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2 - 2)\right) + \mathbf{C}\left(\mathbf{F}_1(x_1 - 1), \mathbf{F}_2(x_2 - 1)\right)\end{aligned} \tag{10}$$

[3] This database is on the web site of the London Metal Exchange (http://www.lme.co.uk) and it is available in a Gauss format on request.

## 2.2 The IFM method

The problem with the ML method is that it could be computational intensive in the case of high dimension, because it requires to estimate jointly the parameters of the margins and the parameters of the dependence structure. However, the copula representation splits the parameters into *specific* parameters for marginal distributions and *common* parameters for the dependence structure (or the parameters of the copula). The log-likelihood (8) could then be written as (Joe and Xu [1996])

$$\ell(\theta) = \sum_{t=1}^{T} \ln c\left(\mathbf{F}_1\left(x_1^t;\theta_1\right),\ldots,\mathbf{F}_n\left(x_n^t;\theta_n\right),\ldots,\mathbf{F}_N\left(x_N^t;\theta_N\right);\alpha\right) + \sum_{t=1}^{T}\sum_{n=1}^{N} \ln f_n\left(x_n^t;\theta_n\right) \tag{15}$$

with $\theta = (\theta_1,\ldots,\theta_N,\alpha)$. $\theta_n$ and $\alpha$ are the vectors of parameters of the parametric marginal distribution $\mathbf{F}_n$ and the copula $\mathbf{C}$. We could also perform the estimation of the univariate marginal distributions in a first time

$$\hat{\theta}_n = \arg\max \ell^n(\theta_n) := \arg\max \sum_{t=1}^{T} \ln f_n\left(x_n^t;\theta_n\right) \tag{16}$$

and then estimate $\alpha$ given the previous estimates

$$\hat{\alpha} = \arg\max \ell^c(\alpha) := \arg\max \sum_{t=1}^{T} \ln c\left(\mathbf{F}_1\left(x_1^t;\hat{\theta}_1\right),\ldots,\mathbf{F}_n\left(x_n^t;\hat{\theta}_n\right),\ldots,\mathbf{F}_N\left(x_N^t;\hat{\theta}_N\right);\alpha\right) \tag{17}$$

This two-step method is called the method of *inference functions for margins* or IFM method (Joe and Xu [1996]). The IFM estimator $\hat{\theta}_{\mathrm{IFM}}$ is then defined as the vector $\left(\hat{\theta}_1,\ldots,\hat{\theta}_N,\hat{\alpha}\right)$. Like the ML estimator, we could show that it verifies the property of **asymptotic normality** and we have

$$\sqrt{T}\left(\hat{\theta}_{\mathrm{IFM}} - \theta_0\right) \longrightarrow \mathcal{N}\left(\mathbf{0},\mathcal{V}^{-1}(\theta_0)\right) \tag{18}$$

with $\mathcal{V}(\theta_0)$ the **information matrix** of Godambe. Let us define a score function in the following way $\mathbf{g}(\theta) = \left(\partial_{\theta_1}\ell^1,\ldots,\partial_{\theta_N}\ell^N,\partial_\alpha\ell^c\right)$. The Godambe information matrix takes the form (Joe [1997]):

$$\mathcal{V}(\theta_0) = \mathbf{D}^{-1}\mathbf{M}\left(\mathbf{D}^{-1}\right)^\top \tag{19}$$

where $\mathbf{D} = \mathbf{E}\left[\partial\mathbf{g}(\theta)^\top/\partial\theta\right]$ and $\mathbf{M} = \mathbf{E}\left[\mathbf{g}(\theta)^\top\mathbf{g}(\theta)\right]$. The estimation of the covariance matrix requires to compute many derivatives. Joe and Xu [1996] suggest then to use the Jacknife method to estimate it. Note also that the IFM method could be viewed as a special case of the generalized method of moments with an identity weight matrix.

Using a close idea of the IFM method, we remark that the parameter vector $\alpha$ of the copula could be estimated without specifying the marginals. The method consists in transforming the data $(x_1^t,\ldots,x_N^t)$ into uniform variates $(\hat{u}_1^t,\ldots,\hat{u}_N^t)$ — using the empirical distributions — and then estimate the parameter in the following way:

$$\hat{\alpha} = \arg\max \sum_{t=1}^{T} \ln c\left(\hat{u}_1^t,\ldots,\hat{u}_n^t,\ldots,\hat{u}_N^t;\alpha\right) \tag{20}$$

In this case, $\hat{\alpha}$ could be viewed as the ML estimator given the *observed* margins (**without assumptions on the parametric form of the marginal distributions**). Because this method is based on the empirical distributions, we call it the *canonical maximum likelihood method* or CML. Note that the IFM method could be viewed as a CML method with $\hat{u}_n^t = \mathbf{F}_n\left(x_n^t;\hat{\theta}_n\right)$. In the rest of the paper, we use the notation $u_n^t$ to designe either $\hat{u}_n^t$ or $\mathbf{F}_n\left(x_n^t;\hat{\theta}_n\right)$.

We remark that the copula representation presents some very interesting properties for the estimation. Consider for example the multivariate generalized hyperbolic distribution (Prause [1999])

$$\begin{aligned} f(\mathbf{x}) &= A(\lambda,\alpha,\beta,\delta,\rho)\left(\delta^2+(\mathbf{x}-\mu)^\top\rho^{-1}(\mathbf{x}-\mu)\right)^{\frac{1}{4}(2\lambda-N)} \\ &\quad\times\mathbf{K}_{\lambda-\frac{N}{2}}\left(\alpha\sqrt{\delta^2+(\mathbf{x}-\mu)^\top\rho^{-1}(\mathbf{x}-\mu)}\right)\exp\left(\beta^\top(\mathbf{x}-\mu)\right) \end{aligned} \tag{21}$$

where $\mathbf{K}$ denotes the modified Bessel function of the third kind and

$$A(\lambda,\alpha,\beta,\delta,\rho)=\frac{\left(\alpha^2-\beta^\top\rho\beta\right)^{\frac{1}{2}\lambda}}{(2\pi)^{\frac{N}{2}}\alpha^{\lambda-\frac{N}{2}}\delta^\lambda\mathbf{K}_\lambda\left(\delta\sqrt{\alpha^2-\beta^\top\rho\beta}\right)} \tag{22}$$

The number of parameters is equal to $\frac{1}{2}N[(N+3)]+3$. Suppose that $N=25$, then we have 353 parameters. Bouyé, Durrleman, Nikeghbali, Riboulet and Roncalli [2000] present a multivariate version of the univariate generalized hyperbolic distribution based on the gaussian copula. Its density is

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{|\rho|^{\frac{1}{2}}}A(\lambda,\alpha,\beta,\delta,\rho)\left[\prod_{n=1}^N\left(\delta_n^2+(x_n-\mu_n)^2\right)^{\frac{1}{4}(2\lambda_n-1)}\right] \\ &\quad\times\left[\prod_{n=1}^N\mathbf{K}_{\lambda_n-\frac{1}{2}}\left(\alpha_n\sqrt{\delta_n^2+(x_n-\mu_n)^2}\right)\right]\exp\left(\beta^\top(\mathbf{x}-\mu)-\frac{1}{2}\varsigma^\top\left(\rho^{-1}-\mathbb{I}\right)\varsigma\right) \end{aligned} \tag{23}$$

where

$$A(\lambda,\alpha,\beta,\delta,\rho)=\frac{1}{(2\pi)^{\frac{N}{2}}}\prod_{n=1}^N\frac{\left(\alpha_n^2-\beta_n^2\right)^{\frac{1}{2}\lambda_n}}{\alpha_n^{\lambda_n-\frac{1}{2}}\delta_n^{\lambda_n}\mathbf{K}_{\lambda_n}\left(\delta_n\sqrt{\alpha_n^2-\beta_n^2}\right)} \tag{24}$$

and $\varsigma=(\varsigma_1,\ldots,\varsigma_N)^\top$, $\varsigma_n=\mathbf{\Phi}^{-1}(\mathbf{F}_n(x_n))$ and $\mathbf{F}_n$ the univariate GH distribution[4]. This distribution has $\frac{1}{2}N[(N+9)]$ parameters. For example, we have 425 parameters for $N=25$. Even if this distribution has more parameters than the previous one, we could perform the estimation more easily. The IFM method could not be apply to the distribution (21). In the case of the copula-based distribution, the IFM method consists in estimating the parameters of the margins ($N$ maximum likelihood with 5 parameters per estimation) and the estimation of the parameters of the copula (which is straightforward in the case of the gaussian copula — see below).

One of the important issue for the estimation is the existence of analytical solution of the IFM (or CML) estimator, because it reduces computational aspects. This is for example a key point in the Finance industry. Here are the IFM estimator for the gaussian and the student copulas:

- **Gaussian copula**
  Let $\varsigma_t=\left(\mathbf{\Phi}^{-1}(u_1^t),\ldots,\mathbf{\Phi}^{-1}(u_n^t),\ldots,\mathbf{\Phi}^{-1}(u_N^t)\right)$. Then we have

$$\hat{\rho}_{\mathrm{IFM}}=\frac{1}{T}\sum_{t=1}^T\varsigma_t^\top\varsigma_t \tag{27}$$

[4] The corresponding density function is given by

$$\begin{aligned} f(x) &= a(\lambda,\alpha,\beta,\delta)\ \ \delta^2+(x-\mu)^2\ \ ^{\frac{1}{4}(2\lambda-1)} \\ &\quad\times\mathbf{K}_{\lambda-\frac{1}{2}}\ \ \alpha\sqrt{\delta^2+(x-\mu)^2}\ \ \exp(\beta(x-\mu)) \end{aligned} \tag{25}$$

with

$$a(\lambda,\alpha,\beta,\delta)=\frac{\alpha^2-\beta^2\ \ ^{\frac{1}{2}\lambda}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^\lambda\mathbf{K}_\lambda\ \ \delta\sqrt{\alpha^2-\beta^2}} \tag{26}$$

- **Student copula**
  Let $\varsigma_t = \left(\mathbf{t}_\nu^{-1}\left(u_1^t\right), \ldots, \mathbf{t}_\nu^{-1}\left(u_n^t\right), \ldots, \mathbf{t}_\nu^{-1}\left(u_N^t\right)\right)$. Bouyé, Durrleman, Nikeghbali, Riboulet and Roncalli [2000] propose to estimate the $\rho$ matrix with the following algorithm:

  1. Let $\hat{\rho}_0$ be the IFM estimate of the $\rho$ matrix for the gaussian copula;
  2. $\hat{\rho}_{m+1}$ is obtained using the following equation

  $$\hat{\rho}_{m+1} = \frac{1}{T}\left(\frac{\nu+N}{\nu}\right)\sum_{t=1}^{T}\frac{\varsigma_t^\top \varsigma_t}{1+\frac{1}{\nu}\varsigma_t^\top \hat{\rho}_m^{-1}\varsigma_t} \tag{28}$$

  3. Repeat the second step until convergence — $\hat{\rho}_{m+1} = \hat{\rho}_m$ $(:= \hat{\rho}_\infty)$.
  4. The IFM estimate of the $\rho$ matrix for the Student copula is $\hat{\rho}_{\mathrm{IFM}} = \hat{\rho}_\infty$.

Let us now consider the previous example. We have for $n = 1, 2$

$$\ell^n\left(m_n, \sigma_n\right) = \sum_{t=1}^{T}\left[-\frac{1}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left(\sigma_n^2\right) - \frac{1}{2}\left(\frac{x_n^t - m_n}{\sigma_n}\right)^2\right] \tag{29}$$

We denote $\hat{m}_n$ and $\hat{\sigma}_n$ the corresponding ML estimates. We then define

$$u_n^t = \Phi\left(\frac{x_n^t - \hat{m}_n}{\hat{\sigma}_n}\right) \tag{30}$$

The paramater $\alpha$ is also estimated by maximising $\ell^c\left(\alpha\right) = \sum_{t=1}^{T}\ell_t^c\left(\alpha\right)$ with

$$\begin{aligned}\ell_t^c\left(\alpha\right) &= \ln\left[\alpha\left(1-\exp\left(-\alpha\right)\right)\exp\left(-\alpha\left(u_1^t + u_2^t\right)\right)\right] - \\ &\quad \ln\left[\left(1-\exp\left(-\alpha\right)\right) - \left(1-\exp\left(-\alpha u_1^t\right)\right)\left(1-u_2^t\right)\right]^2\end{aligned} \tag{31}$$

The Godambe covariance matrix is then computed with the score function

$$\mathbf{g}\left(\theta\right) = \begin{bmatrix}\partial\ell^1\left(m_1,\sigma_1\right)/\partial m_1 \\ \partial\ell^1\left(m_1,\sigma_1\right)/\partial\sigma_1 \\ \partial\ell^2\left(m_2,\sigma_2\right)/\partial m_2 \\ \partial\ell^2\left(m_2,\sigma_2\right)/\partial\sigma_2 \\ \partial\ell^c\left(\alpha\right)/\partial\alpha\end{bmatrix} \tag{32}$$

The results are the following and are very closed to those given by the ML method.

| Parameters | estimate | std.err. | t-statistic | p-value |
|---|---|---|---|---|
| $m_1$ | -0.000207 | 0.000256 | -0.81 | 0.42 |
| $m_2$ | 0.013316 | 0.000325 | 41.0 | 0.00 |
| $\sigma_1$ | -0.000256 | 0.000320 | -0.80 | 0.42 |
| $\sigma_2$ | 0.016680 | 0.000455 | 36.7 | 0.00 |
| $\alpha$ | 4.440463 | 0.156818 | 28.3 | 0.00 |

In the case of the CML method, we obtain $\hat{\alpha}_{\mathrm{CML}} = 3.578972$ — the IFM and CML estimates are not very closed, because the IFM method does take into account the marginal distributions.

## 2.3 Estimation based on the dependence measures

We could also estimate the parameters such that they fit the dependence measures, for example the Kendall's tau, the Spearman's rho or the upper tail dependence measure. Let $\mathcal{L}(\theta)$ be a loss function. We define the point estimator $\hat{\theta}$ as the solution of the following problem (LEHMANN and CASELLA [1998])

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \mathcal{L}(\theta) \tag{33}$$

with $\Theta$ the parameters space. In most cases, people use a quadratic loss distribution

$$\mathcal{L}(\theta) = [\hat{g} - g(\theta)]^{\top} W [\hat{g} - g(\theta)] \tag{34}$$

with $W$ a weight matrix, $\hat{g}$ the observed criterion values and $g$ the criterion functions.

In the case of one-parameter bivariate copula, we could perform the estimation with only one dependence measure. In some cases, analytical solutions are available (see the table below for some examples).

| | $\varrho$ | $\tau$ | $\lambda$ |
|---|---|---|---|
| Gaussian | $\rho = 2\sin\left(\frac{\pi}{6}\varrho\right)$ | $\rho = \sin\left(\frac{\pi}{2}\tau\right)$ | ✓ |
| Gumbel | numerical solution | $\alpha = (1-\tau)^{-1}$ | $\ln 2 \ln (2-\lambda)^{-1}$ |
| FGM | $3\varrho$ | $\frac{9}{2}\tau$ | ✓ |

Otherwise, we use a root finding procedure. This is the case of the Frank copula (FREES and VALDEZ [1997]):

$$\begin{aligned} \varrho &= 1 - \frac{12}{\alpha}\left(D_1(\alpha) - D_2(\alpha)\right) \\ \tau &= 1 - \frac{4}{\alpha}\left(1 - D_1(\alpha)\right) \end{aligned} \tag{35}$$

with $D_k(x)$ the Debye function (ABRAMOWITZ and STEGUN [1970]). With the LME example, we obtain $\hat{\varrho} = 0.49958$, so it comes that $\hat{\alpha} = 3.4390$. Note that this estimate is closer to the CML estimate than the ML and IFM estimates.

# 3 Non parametric estimation

In this section, we don't assume anymore that we have a parametric copula. We are now interested in modelling the dependence structure with consistency, this is to say thanks to a copula that is going to converge to the real underlying dependence structure.

## 3.1 The Deheuvels or empirical copula

We present in this paragraph the notion of empirical copula as introduced by DEHEUVELS [1979]. Let $\mathbf{X}^t = (X_1^t, \ldots, X_N^t) \in \mathbb{R}^N$ be an i.i.d. sequence with distribution $\mathbf{F}$ and margins $\mathbf{F}_n$. We assume that $\mathbf{F}$ is continuous so that the copula associated to $\mathbf{F}$ is unique.

If $\delta_{\mathbf{u}}$ stands for the Dirac measure with $\mathbf{u} \in \mathbb{R}^N$, we define $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} \delta_{\mathbf{X}^t}$ the empirical measure associated with a sample of $\mathbf{X}$, and $\hat{\mathbf{F}}(x_1, \ldots, x_N) = \hat{\mu}\left(\prod_{n=1}^{N} ]-\infty, x_n]\right)$ its empirical function. We note $\left\{x_1^{(t)}, \ldots, x_N^{(t)}\right\}$ the order statistic and $\{r_1^t, \ldots, r_N^t\}$ the rank statistic of the sample which are linked by the relationship $x_n^{(r_n^t)} = x_n^t$. It is possible to introduce the empirical copula of the sample as any copula $\hat{\mathbf{C}}$ of the empirical distribution $\hat{\mathbf{F}}$. But the problem is that $\hat{\mathbf{C}}$ is not unique, that's why DEHEUVELS [1981] proposes the following way to cope with the problem.

**Definition 1** *Any copula* $\hat{\mathbf{C}} \in \mathcal{C}$ *defined on the lattice*

$$\mathfrak{L} = \left\{ \left( \frac{t_1}{T}, \ldots, \frac{t_N}{T} \right) : 1 \leq n \leq N, t_n = 0, \ldots, T \right\} \tag{36}$$

*by*

$$\hat{\mathbf{C}} \left( \frac{t_1}{T}, \ldots, \frac{t_N}{T} \right) = \frac{1}{T} \sum_{t=1}^{T} \prod_{n=1}^{N} \mathbf{1}_{[r_n^t \leq t_n]} \tag{37}$$

*is an empirical copula.*

We introduce the notation $\hat{\mathbf{C}}_{(T)}$ in order to define the order of the copula, that is the dimension of the sample used to construct it. Deheuvels [1979,1981] obtains then the following conclusions:

1. The empirical measure $\hat{\mu}$ (or the empirical distribution function $\hat{\mathbf{F}}$) is uniquely and reciprocally defined by both
   (a) the empirical measures of each coordinate $\hat{\mathbf{F}}_n$;
   (b) the values of an empirical copula $\hat{\mathbf{C}}$ on the set $\mathfrak{L}$.
2. The empirical copula $\hat{\mathbf{C}}$ defined on $\mathfrak{L}$ is in distribution independent of the margins of $\mathbf{F}$.
3. If $\hat{\mathbf{C}}_{(T)}$ is any empirical copula of order $T$, then $\hat{\mathbf{C}}_{(T)} \rightarrow \mathbf{C}$ with the topology of $\mathcal{C}$ (uniform convergence for example).

We could now define the analog of the Radon-Nikodym density for the empirical copula $\hat{\mathbf{C}}$

$$\hat{c} \left( \frac{t_1}{T}, \ldots, \frac{t_n}{T}, \ldots, \frac{t_N}{T} \right) = \sum_{i_1=1}^{2} \cdots \sum_{i_N=1}^{2} (-1)^{i_1 + \cdots + i_N} \hat{\mathbf{C}} \left( \frac{t_1 - i_1 + 1}{T}, \ldots, \frac{t_n - i_n + 1}{T}, \ldots, \frac{t_N - i_N + 1}{T} \right) \tag{38}$$

$\hat{c}$ is called the *empirical copula frequency* (Nelsen [1998]). The relationships between empirical copula distribution and frequency are

$$\hat{\mathbf{C}} \left( \frac{t_1}{T}, \ldots, \frac{t_n}{T}, \ldots, \frac{t_N}{T} \right) = \sum_{i_1=1}^{t_1} \cdots \sum_{i_N=1}^{t_N} \hat{c} \left( \frac{i_1}{T}, \ldots, \frac{i_n}{T}, \ldots, \frac{i_N}{T} \right) \tag{39}$$

Note that empirical copulas permits to define the sample version of dependence measures. For example, the the Spearman's rho and Kendall's tau take the following form (Nelsen [1998])

$$\hat{\varrho} = \frac{12}{T^2 - 1} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \left( \hat{\mathbf{C}} \left( \frac{t_1}{T}, \frac{t_2}{T} \right) - \frac{t_1 t_2}{T^2} \right) \tag{40}$$

and

$$\hat{\tau} = \frac{2}{T - 1} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{i_2=1}^{t_1 - 1} \sum_{i_2=1}^{t_2 - 1} \left( \hat{\mathbf{c}} \left( \frac{t_1}{T}, \frac{t_2}{T} \right) \hat{\mathbf{c}} \left( \frac{i_1}{T}, \frac{i_2}{T} \right) - \hat{\mathbf{c}} \left( \frac{t_1}{T}, \frac{i_2}{T} \right) \hat{\mathbf{c}} \left( \frac{i_1}{T}, \frac{t_2}{T} \right) \right) \tag{41}$$

We could also define the sample version of the quantile dependence function

$$\lambda(u) = \Pr \{ U_1 > u | U_2 > u \} = \frac{\bar{\mathbf{C}}(u, u)}{1 - u} \tag{42}$$

and the upper tail dependence measure

$$\lambda = \lim_{u \longrightarrow 1^-} \lambda(u) \tag{43}$$

We have

$$\hat{\lambda}\left(\frac{t}{T}\right) = 2 - \frac{1 - \hat{\mathbf{C}}\left(\frac{t}{T}, \frac{t}{T}\right)}{1 - \frac{t}{T}} \tag{44}$$

and $\hat{\lambda} = \lim_{t \to T} \hat{\lambda}\left(t/T\right)$.

We consider now the example of the LME data. We have represented in the figure 1 the corresponding empirical copula[5]. In the figure 2, we have reported the empirical quantile dependence function $\hat{\lambda}$ and the dependence function with the Frank copula and the parameter equal to the previous ML estimate. By assuming a normal distribution, we could compute also a confidence interval for $\hat{\lambda}$ (Coles, Currie and Tawn [1999]).

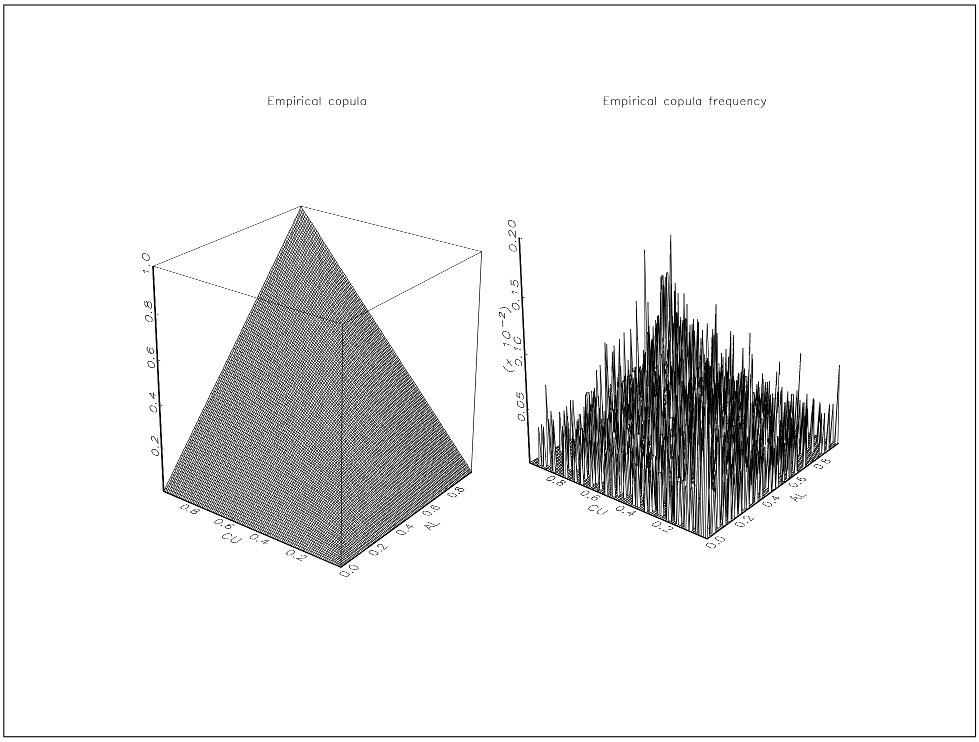


Figure 1: Empirical copula of the (AL,CU) data

## 3.2 Copula approximation

In this paragraph, we use the idea of Li, Mikusinski, Sherwood and Taylor [1997] on approximation of copulas. Their motivation was to present "*certain approximations that lead naturally to a stronger convergence than uniform convergence*". These approximations are very interesting to study the properties of the $*$-product of copulas defined by Darsow, Nguyen and Olsen [1992] to characterize markov processes. Below, we use these approximations to estimate non parametrically the underlying dependence structure. We present only the two dimensional case, but the generalization is straightforward.

[5] We have 2713 observations in the database. Nevertheless, we use a grid equal to $\frac{1}{25}$ for the plots. The order $T$ of the empirical data is then 109.

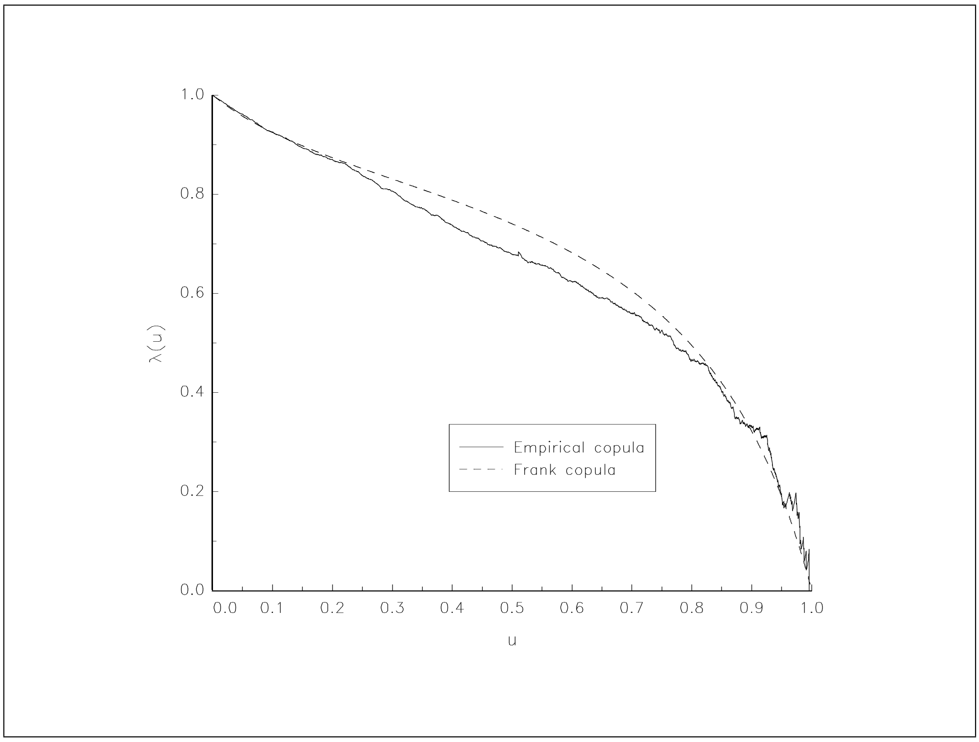


Figure 2: Quantile dependence function

### 3.2.1 Approximation by Bernstein polynomials

Let $B_{i,n}(x)$ denotes the Bernstein polynomial

$$B_{i,n}(x) = \binom{n}{i} x^{i} (1-x)^{n-i} \tag{45}$$

Li, Mikusinski, Sherwood and Taylor [1997] showed that for any copula $\mathbf{C}$, $\mathfrak{B}_n(\mathbf{C})$ defined by

$$\mathfrak{B}_n(\mathbf{C})(u,v) = \sum_{i=1}^{n} \sum_{j=1}^{n} B_{i,n}(u) B_{j,n}(v) \mathbf{C}\left(\frac{i}{n}, \frac{j}{n}\right) \tag{46}$$

is a copula too. Moreover, it is well known that the copula $\mathfrak{B}_n(\mathbf{C})$ converges to the copula $\mathbf{C}$ in the strong sense. Now, if we consider the two place function

$$\mathfrak{B}_T\left(\hat{\mathbf{C}}_{(T)}\right)(u,v) = \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} B_{t_1,T}(u) B_{t_2,T}(v) \hat{\mathbf{C}}\left(\frac{t_1}{T}, \frac{t_2}{T}\right) \tag{47}$$

we know thanks to what we have said previously that $\mathfrak{B}_T\left(\hat{\mathbf{C}}_{(T)}\right)$ is a copula that will (uniformly) converge to the real copula $\mathbf{C}$

$$\mathfrak{B}_T\left(\hat{\mathbf{C}}_{(T)}\right) \longrightarrow \mathbf{C} \tag{48}$$

### 3.2.2 The checkerboard approximation

Another consistent way for modelling the dependence structure is to consider a *checkerboard approximation*

$$\mathfrak{C}_n(\mathbf{C})(u,v) = n^2 \sum_{i=1}^{n}\sum_{j=1}^{n} \Delta_{i,j}(\mathbf{C}) \int_0^u \chi_{i,n}(x)\, \mathrm{d}x \int_0^v \chi_{j,n}(y)\, \mathrm{d}y \tag{49}$$

where

$$\Delta_{i,j}(\mathbf{C}) = \mathbf{C}\left(\frac{i}{n},\frac{j}{n}\right) - \mathbf{C}\left(\frac{i-1}{n},\frac{j}{n}\right) - \mathbf{C}\left(\frac{i}{n},\frac{j-1}{n}\right) + \mathbf{C}\left(\frac{i-1}{n},\frac{j-1}{n}\right) \tag{50}$$

and $\chi_{i,n}$ is the characteristic function of the interval $\left[\frac{i-1}{n},\frac{i}{n}\right]$. Li, Mikusinski, Sherwood and Taylor [1997] give some properties of $\mathfrak{C}_n(\mathbf{C})$:

1. $\mathfrak{C}_n(\mathbf{C})$ is the copula of the doubly stochastic measure whose mass in every square of the form $\left[\frac{i-1}{n},\frac{i}{n}\right] \times \left[\frac{j-1}{n},\frac{j}{n}\right]$ is uniformly distributed and equal to the total mass in that square in the doubly stochastic measure of the original copula
2. $\mathfrak{C}_n(\mathbf{C})$ approximate $\mathbf{C}$ uniformly

$$\sup_{(u,v)\in[0,1]^2} \left|\mathbf{C}(u,v) - \mathfrak{C}_n(\mathbf{C})(u,v)\right| \leq 2n^{-1} \tag{51}$$

If we now take the two place function

$$\mathfrak{C}_T\left(\hat{\mathbf{C}}_{(T)}\right)(u,v) = T^2 \sum_{t_1=1}^{T}\sum_{t_2=1}^{T} \Delta_{t_1,t_2}\left(\hat{\mathbf{C}}_{(T)}\right) \int_0^u \chi_{t_1,T}(x)\, \mathrm{d}x \int_0^v \chi_{t_2,T}(y)\, \mathrm{d}y \tag{52}$$

where

$$\Delta_{t_1,t_2}\left(\hat{\mathbf{C}}_{(T)}\right) = \hat{\mathbf{C}}_{(T)}\left(\frac{t_1}{T},\frac{t_2}{T}\right) - \hat{\mathbf{C}}_{(T)}\left(\frac{t_1-1}{T},\frac{t_2}{T}\right) - \hat{\mathbf{C}}_{(T)}\left(\frac{t_1}{T},\frac{t_2-1}{T}\right) + \hat{\mathbf{C}}_{(T)}\left(\frac{t_1-1}{T},\frac{t_2-1}{T}\right) \tag{53}$$

From what we have said previously, we can deduce that $\mathfrak{C}_T\left(\hat{\mathbf{C}}_{(T)}\right)$ is a copula which converges (uniformly) to the real underlying copula

$$\mathfrak{C}_T\left(\hat{\mathbf{C}}_{(T)}\right) \longrightarrow \mathbf{C} \tag{54}$$

### 3.2.3 Applications

We consider the previous empirical copula of the (AL,CU) data. We have reported in the figure 3 its approximation. One of the main interest of this construction is that **the associated measure is continuous**. Moreover, we know that the approximated copula converge to the underlying dependence structure. In the figure 4, we have represented the measure for different orders $T$. Another important point is that the convergence of the concordance measures is satisfied (Durrleman, Nikeghbali and Roncalli [2000]). However, these approximation methods do not preserve the upper tail dependence. This is a real problem in financial modelling. Nevertheless, one could use the perturbation method of Durrleman, Nikeghbali and Roncalli [2000] to obtain the desired upper tail dependence.

# 4 Selection criteria based on empirical copulas

In this section, we consider the selection problem of 'optimal' copula. In general, we assume that we have a finite subset of copulas $\tilde{\mathcal{C}} \subset \mathcal{C}$. And we would like to know which copula fits best the data. If $\tilde{\mathcal{C}}$ corresponds to the Archimedean family, we could solve the problem using results on Kendall's processes. In the general case, we could use a distance based on the discrete $L^p$ norm.

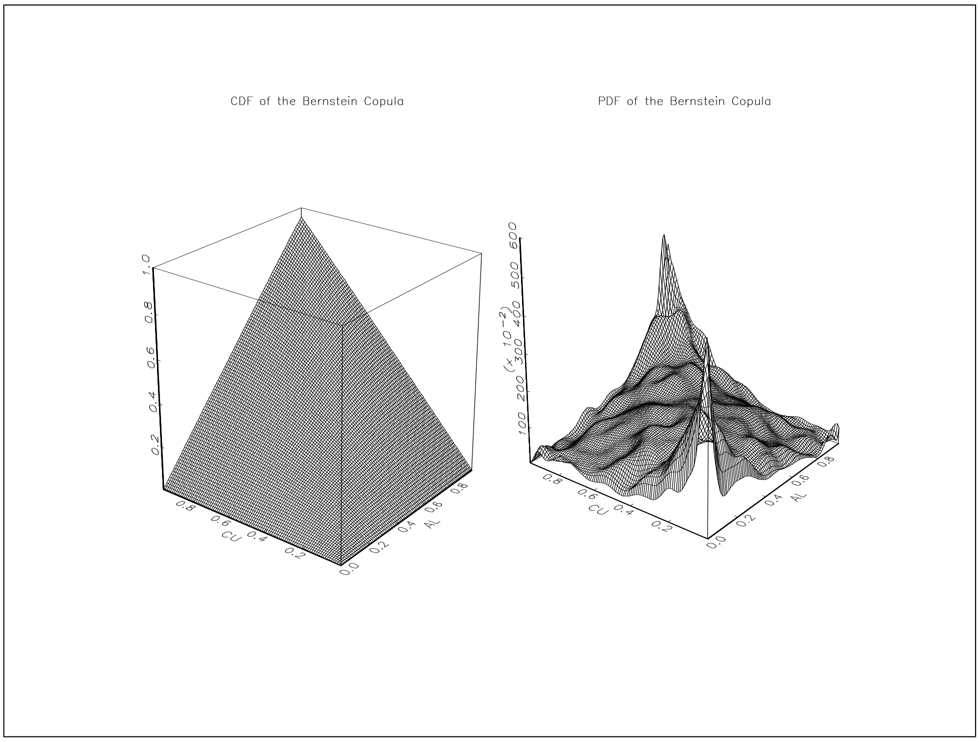


Figure 3: Approximation of the empirical copula of the (AL,CU) data with Bernstein polynomials

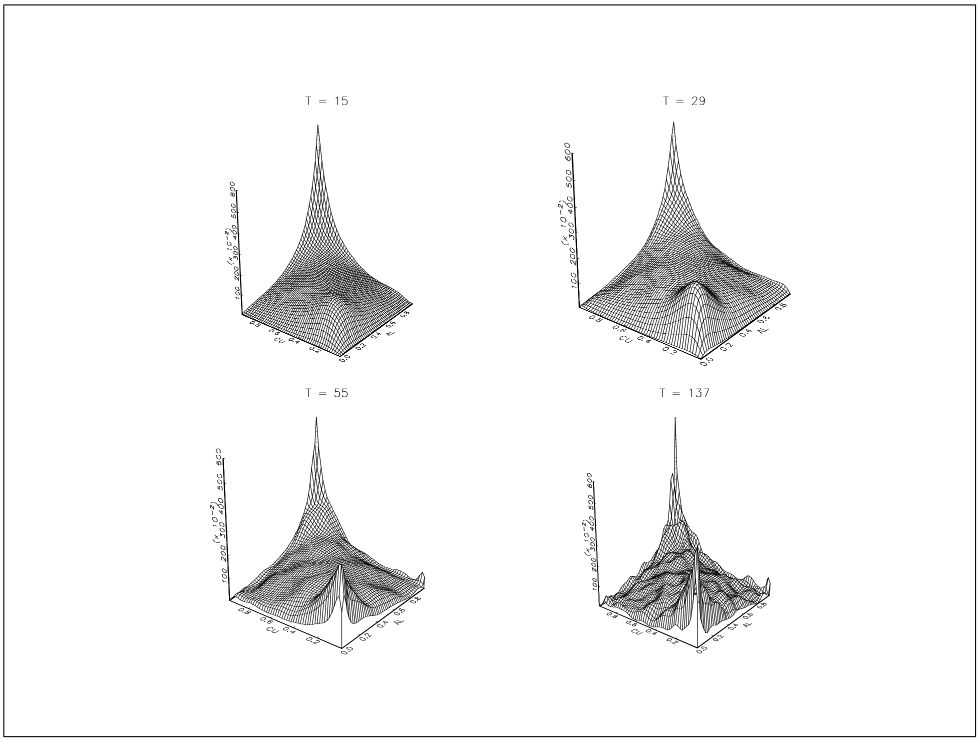


Figure 4: Probability density function of the Bernstein copula with different orders

### 4.1 The case of Archimedean copulas

Genest and MacKay [1996] define Archimedean copulas as the following:

$$\mathbf{C}(u_1,\ldots,u_n,\ldots,u_N)=\begin{cases}\varphi^{-1}\left(\varphi(u_1)+\ldots+\varphi(u_n)+\ldots+\varphi(u_N)\right) & \text{if } \sum_{n=1}^{N}\varphi(u_n)\leq\varphi(0)\\ 0 & \text{otherwise}\end{cases}$$

with $\varphi(u)$ a $C^2$ function with $\varphi(1)=0$, $\varphi'(u)<0$ and $\varphi''(u)>0$ for all $0\leq u\leq 1$. Genest and Rivest [1993] have developed an empirical method to identify the copula in the Archimedean case. Let $\mathbf{X}$ be a vector of $N$ random variables, $\mathbf{C}$ the associated copula with generator $\varphi$ and $K$ the function defined by

$$K(u)=\Pr\left\{\mathbf{C}(U_1,\ldots,U_N)\leq u\right\} \tag{55}$$

Barbe, Genest, Ghoudi and Rémillard [1996] showed that

$$K(u)=u+\sum_{n=1}^{N}(-1)^n\frac{\varphi^n(u)}{n!}\varkappa_{n-1}(u) \tag{56}$$

with $\varkappa_n(u)=\frac{\partial_u\,\varkappa_{n-1}(u)}{\partial_u\,\varphi(u)}$ and $\varkappa_0(u)=\frac{1}{\partial_u\,\varphi(u)}$. A non parametric estimate of $K$ is given by

$$\hat{K}(u)=\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{[\vartheta_i\leq u]} \tag{57}$$

with

$$\vartheta_i=\frac{1}{T-1}\sum_{t=1}^{T}\mathbf{1}_{\left[x_1^t<x_1^i,\ldots,x_N^t<x_N^i\right]} \tag{58}$$

The idea is then to choose the Archimedean copula which gives $\hat{K}$. Frees et Valdez [1997] propose to use a QQ-plot of $K$ and $\hat{K}$. Another procedure consists in comparing $u-K(u)$ and $u-\hat{K}(u)$ (Genest and Rivest [1993]).

We consider our LME example. Suppose that $\tilde{\mathcal{C}}$ is the set of four archimedean copulas, the Gumbel copula with parameters $\alpha=1$, $\alpha=1.5$, $\alpha=2$ and $\alpha=3$. We note them respectively $C_1$, $C_2$, $C_3$ and $C_4$. We have represented in the figures 5 and 6 the two graphical procedures. In general, we notice that the procedure of Genest and Rivest [1993] identifies more easily the 'optimal' copula and $C_2$ seems to be the best dependence structure.

In general, $\tilde{\mathcal{C}}$ consists not in one parametric family with different values of the parameters, but $\tilde{\mathcal{C}}$ is the set of different families, for example Gumbel, Frank, Cook-Johnson, etc. (see the table 4.1 of Nelsen [1998] for a list of Archimedean copulas). For each family, we could first estimate the parameters by IFM ou CML method in order to reduce the cardinality of $\tilde{\mathcal{C}}$. Note that if $\mathbf{C}$ is an absolutely two-dimensional copula, the expression of the log-likelihood is

$$\ell_t^c(\alpha)=\ln\varphi''\left(\mathbf{C}\left(u_1^t,u_2^t\right)\right)+\ln\left(\varphi'\left(u_1^t\right)\varphi'\left(u_2^t\right)\right)-3\ln\left(-\varphi'\left(\mathbf{C}\left(u_1^t,u_2^t\right)\right)\right) \tag{59}$$

We consider our previous example with three copula families. The CML estimation gives the following results

| | Copula | $\varphi(u)$ | $\mathbf{C}(u_1,u_2)$ | $\hat{\alpha}_{\mathsf{CML}}$ |
|---|---|---|---|---|
| $C_1$ | Gumbel | $(-\ln u)^{\alpha}$ | $\exp\left(-\left((-\ln u_1)^{\alpha}+(-\ln u_2)^{\alpha}\right)^{\frac{1}{\alpha}}\right)$ | 1.462803 |
| $C_2$ | Cook-Johnson | $\alpha^{-1}\left(u^{-\alpha}-1\right)$ | $\max\left(\left[u_1^{-\alpha}+u_2^{-\alpha}-1\right]^{-\frac{1}{\alpha}},0\right)$ | 0.708430 |
| $C_3$ | Frank | $-\ln\left(\frac{\exp(-\alpha u)-1}{\exp(-\alpha)-1}\right)$ | $-\alpha^{-1}\ln\left(1+\left(e^{-\alpha}-1\right)^{-1}\left(e^{-\alpha u}-1\right)\left(e^{-\alpha v}-1\right)\right)$ | 3.578972 |

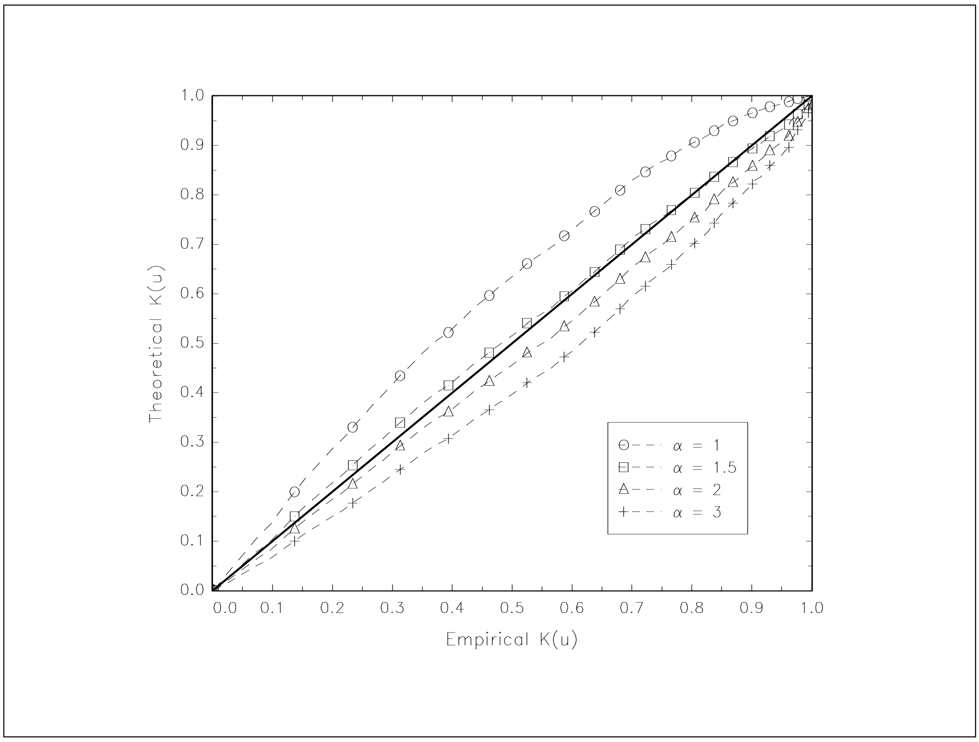


Figure 5: QQ-plot procedure of FREES and VALDEZ [1998]

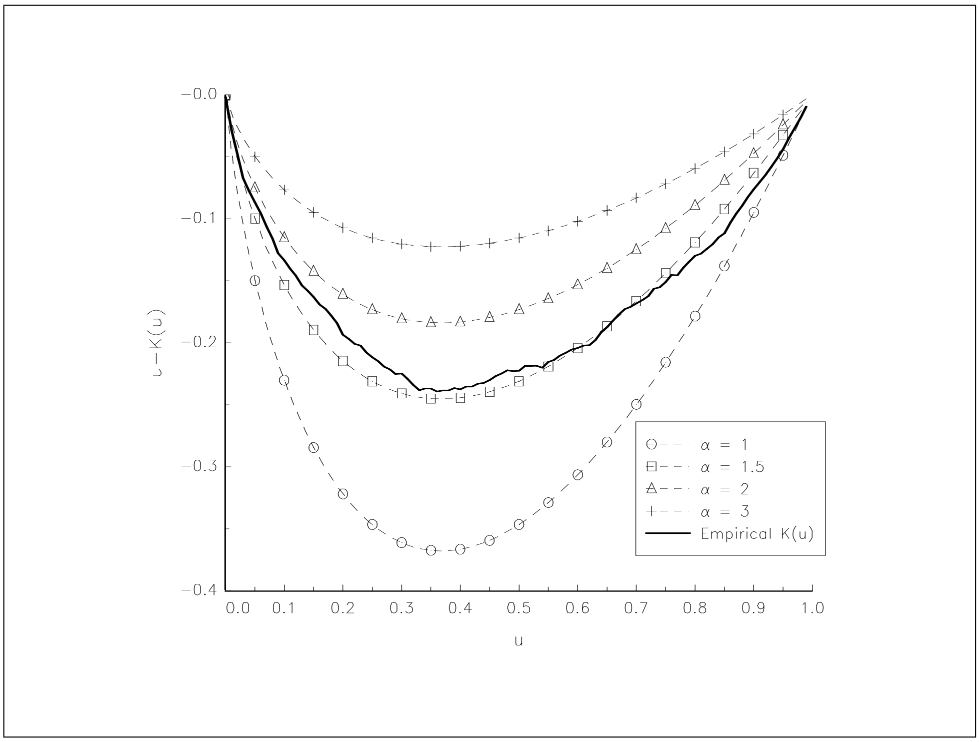


Figure 6: Graphical procedure of GENEST and RIVEST [1993]

We have reported in the figure 7 the values of $u - K(u)$ for these copulas. Because $K(u)$ is a distribution, we could define the 'optimal' copula as the copula which gives the minimum distance between $K(u)$ and $\hat{K}(u)$. For example, we could take the Kolmogorov distance, the Hellinger distance, etc. In the case of the $L^2$ norm, the distance is defined as $d_2\left(\hat{K}, K\right) = \int_0^1 \left[K(u) - \hat{K}(u)\right]^2 \mathrm{d}u$. With the LME data, we obtain the following values: $d_2\left(\hat{K}, K\right) = 2.63 \times 10^{-4}$ for the Gumbel copula, $d_2\left(\hat{K}, K\right) = 11.4 \times 10^{-4}$ for the Cook-Johnson copula and $d_2\left(\hat{K}, K\right) = 1.10 \times 10^{-4}$ for the Frank copula, which is then our 'optimal' copula.

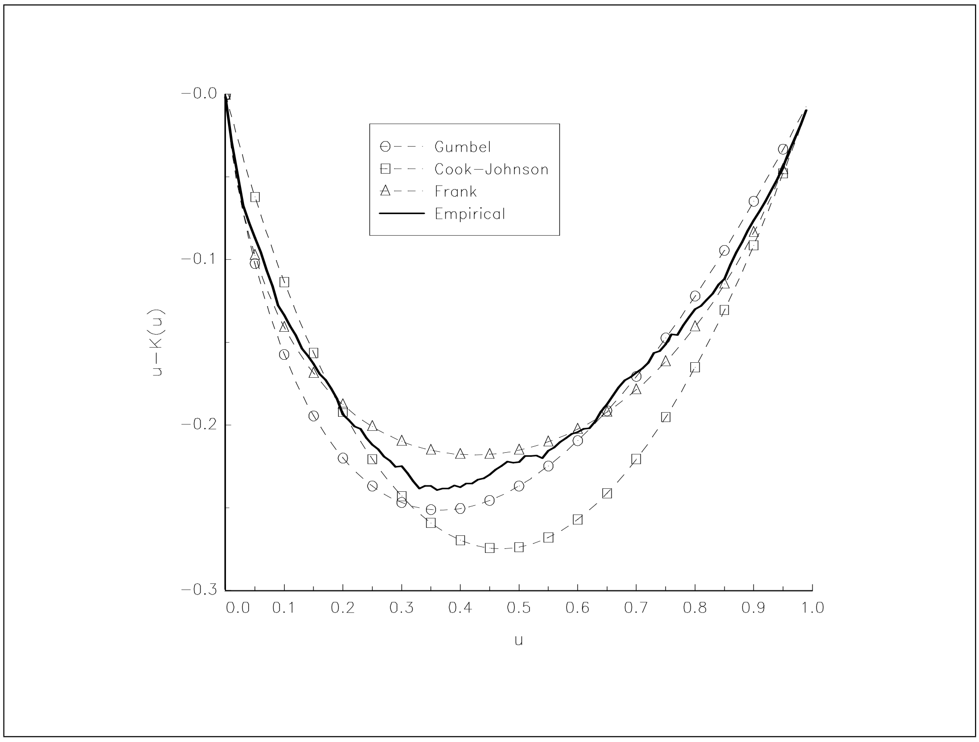


Figure 7: Comparison of the Gumbel, Cook-Johnson and Frank copulas

## 4.2 Selecting a copula among a given subset of copulas

We are not interested here in the different ways of constructing an empirical copula (by interpolation or approximation for example). All we know about any empirical copula is its values on the lattice $\mathfrak{L}$ that we could compute thanks to the data. Here we assume we have a finite subset of copulas $\tilde{\mathcal{C}} \subset \mathcal{C}$, and we are interested in knowing which one of the copulas in $\tilde{\mathcal{C}}$ fits best the data (there might be parametric copulas or non parametric copulas or Fréchet upper bounds for example).

What we suggest here is to consider the distance between each considered copula and the empirical copula. But we have to specify the distance we consider. As there exists more than one empirical copula, it appears that any distance based on the $L^p$ norm would not be proper. For example, let's assume that we have different empirical copulas $\hat{\mathbf{C}}_j$ $(j = 1, \ldots, J)$ and a finite subset of $K$ copulas $\tilde{\mathcal{C}} = \{\mathbf{C}_k\}_{1 \leq k \leq K}$. Let us consider the $L^2$ norm

$$d_2\left(\hat{\mathbf{C}}_j, \mathbf{C}_k\right) = \left\|\hat{\mathbf{C}}_j - \mathbf{C}_k\right\|_{L^2} := \left(\int \cdots \int_{[0,1]^N} \left|\hat{\mathbf{C}}_j(\mathbf{u}) - \mathbf{C}_k(\mathbf{u})\right|^2 \mathrm{d}\mathbf{u}\right)^{\frac{1}{2}} \tag{60}$$

If we consider that the best copula in $\tilde{\mathcal{C}}$ is the copula which minimizes $d_2\left(\hat{\mathbf{C}}_j, \mathbf{C}_k\right)$, we might encounter some problems since the 'optimal' copula depends on $j$ and thus the copula that minimize $d_2\left(\hat{\mathbf{C}}_j, \mathbf{C}_k\right)$ can be different for different values of $j$.

As all empirical copulas take the same values on $\mathfrak{L}$, we suggest to take a distance based on the **discrete** $L^p$ norm

$$\begin{aligned} \bar{d}_2\left(\hat{\mathbf{C}}_{(T)}, \mathbf{C}_k\right) &= \left\|\hat{\mathbf{C}}_{(T)} - \mathbf{C}_k\right\|_{L^2} \\ &= \left(\sum_{t_1=1}^{T} \cdots \sum_{t_n=1}^{T} \cdots \sum_{t_N=1}^{T} \left[\hat{\mathbf{C}}_{(T)}\left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}, \ldots, \frac{t_N}{T}\right) - \mathbf{C}_k\left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}, \ldots, \frac{t_N}{T}\right)\right]^2\right)^{\frac{1}{2}} \end{aligned} \tag{61}$$

We consider that the best copula in the family $\tilde{\mathcal{C}}$ is the copula which minimizes $\bar{d}_2\left(\hat{\mathbf{C}}_{(T)}, \mathbf{C}_k\right)$. The advantage of this method is that it does not depend on the behavior of the empirical copulas out of the lattice $\mathfrak{L} = \left\{\left(\frac{t_1}{T}, \ldots, \frac{t_N}{T}\right) : 1 \leq n \leq N, t_n = 0, \ldots, T\right\}$.

We use the LME example. And we suppose that the set $\tilde{\mathcal{C}}$ corresponds to the set of the previous paragraph (Gumbel, Cook-Johnson and Frank copulas) and the Gaussian copula. We obtain the following results:

| | Copula | $\hat{\alpha}_{\mathsf{CML}}$ | $\frac{\bar{d}_2\left(\hat{\mathbf{C}}_{(T)}, \mathbf{C}_k\right)}{T}$ |
|---|---|---|---|
| $C_1$ | Gumbel | 1.462803 | $7.39 \times 10^{-3}$ |
| $C_2$ | Cook-Johnson | 0.708430 | $16.5 \times 10^{-3}$ |
| $C_3$ | Frank | 3.578972 | $4.47 \times 10^{-3}$ |
| $C_4$ | Normal | 0.491794 | $4.79 \times 10^{-3}$ |

The Frank copula is then always our 'optimal' copula.

## 4.3 Estimating parameters in 'untractable' cases

We propose here a method for estimating parameters (of some parametric families of copulas) which are difficult to compute by maximum likelihood. This is for example the case for some asymmetric extreme copulas because of the shape of the likelihood function. Moreover, the maximum likelihood method requires the cross derivatives of the copula. For high dimensions and some copulas (for example, the MM1-MM5 families in Joe [1997]), an explicit expression is very difficult to obtain. The cross derivatives could then be computed with a numerical algorithm based on finite differences. However, these algorithms produces important numerical errors in high dimension. That's why we suggest in these cases to estimate the parameter of the parametric copula $\mathbf{C}(\mathbf{u};\theta)$ by taking the loss function $\mathcal{L}(\theta)$ equal to the distance (61). We then define the point estimator $\hat{\theta}$ as the solution of the following minimization problem

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left(\sum_{\mathbf{u} \in \mathfrak{L}} \left[\hat{\mathbf{C}}_{(T)}(\mathbf{u}) - \mathbf{C}(\mathbf{u};\theta)\right]^2\right)^{\frac{1}{2}} \tag{62}$$

To perform the numerical optimization, we could use a quasi-Newton algorithm like the BFGS method so that we don't need the Hessian matrix anymore, but just the first order derivatives. Moreover, this sort of estimation is coherent with what we have previously said about the choice of the dependence struture for modelling. For example, we obtain the following results for the LME data:

| | Copula | $\hat{\alpha}_{\mathsf{CML}}$ | $\frac{\bar{d}_2(\hat{\mathbf{C}}_{(T)},\mathbf{C}_k)}{T}$ | $\hat{\alpha}_{\mathsf{L2M}}$ | $\frac{\bar{d}_2(\hat{\mathbf{C}}_{(T)},\mathbf{C}_k)}{T}$ |
|---|---|---|---|---|---|
| $C_1$ | Gumbel | 1.462803 | $7.39 \times 10^{-3}$ | 1.554518 | $5.01 \times 10^{-3}$ |
| $C_2$ | Cook-Johnson | 0.708430 | $16.5 \times 10^{-3}$ | 1.056350 | $12.1 \times 10^{-3}$ |
| $C_3$ | Frank | 3.578972 | $4.47 \times 10^{-3}$ | 3.389611 | $3.93 \times 10^{-3}$ |
| $C_4$ | Normal | 0.491794 | $4.79 \times 10^{-3}$ | 0.528938 | $3.06 \times 10^{-3}$ |

$\hat{\alpha}_{\mathsf{L2M}}$ is the point estimator based on the discrete $L^2$ norm. We remark that we improve significantly the distance. Moreover, in this case, this is the Normal copula which appears to be 'optimal'.

### 4.4 The influence of the margins on the choice of the dependence structure

In order to define a 'tractable' set $\tilde{\mathcal{C}}$, we have to reduce the cardinality of the subset of $\mathcal{C}$ in a first time. This is done by choosing different parameter families and by doing a ML optimization for each family. In the previous example, we have perform the estimation step using the CML method, that is we have assumed that the margins correpond to the empirical ones. However, we are going to see that the margins play an important role to define the optimal copula.

In the second section, we assume that the margins of AL and CU are gaussians. In the case of the Frank copula, we then obtain the following values for the distances:

| Estimation method | $\frac{\bar{d}_2(\hat{\mathbf{C}}_{(T)},\mathbf{C}_k)}{T}$ |
|---|---|
| CML | $4.47 \times 10^{-3}$ |
| IFM | $11.5 \times 10^{-3}$ |
| ML | $12.8 \times 10^{-3}$ |
| L2M | $3.93 \times 10^{-3}$ |

We remark that the discrete $L^2$ norm for IFM and ML methods are larger. If we compare the level curves of the Deheuvels copula, CML Frank copula and ML Frank copula, we obtain the figure 8. We could certainly notice that the gaussian margins are not appropriate. With this example, we see clearly that the impact of the margins is very important. If the margins are not well specified, we could then find an optimal copula into the subset $\tilde{\mathcal{C}}$ which is in fact irrelevant to give the good dependence structure. That's why estimation based on CML method is very important. If we notice that there exist significant differences between CML and ML (or IFM) methods, that indicates that the margins are not well appropriate.

To clarify this point, we consider a simulation study. We assume that the bivariate distribution $\mathbf{F}$ correspond to the Normal copula with parameter 0.5 and two margins $\mathbf{F}_1$ and $\mathbf{F}_2$ which are a student distribution ($\mathbf{F}_1 = t_2$ and $\mathbf{F}_2 = t_3$). We suppose now that we fit a distribution with a Normal copula and two gaussian margins. Because the margins are wrong, the IFM and ML estimators are biased (see the figure 9). We remark that this is not the case of the CML estimator.

## 5 Conclusion

Copulas are a powerful tool in financial modelling. But one of the difficulty is in general the choice of the copula. In this article, we give a few methods to solve this problem. They are all based (sometimes not directly) on the Deheuvels or empirical copula.

## References

[1] Abramowitz, M. and I.A. Stegun [1970], Handbook of Mathematical Functions, ninth edition, Dover, Handbook of Mathematical Functions, ninth edition, Dover

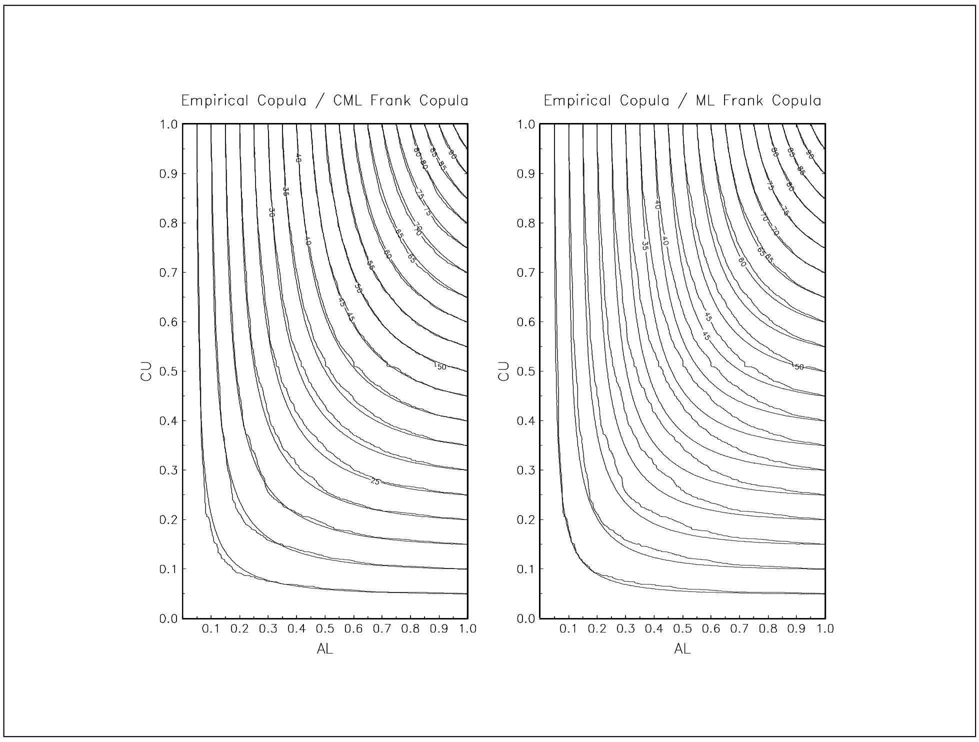


Figure 8: Comparison of empirical copula, CML Frank copula and ML Frank copula

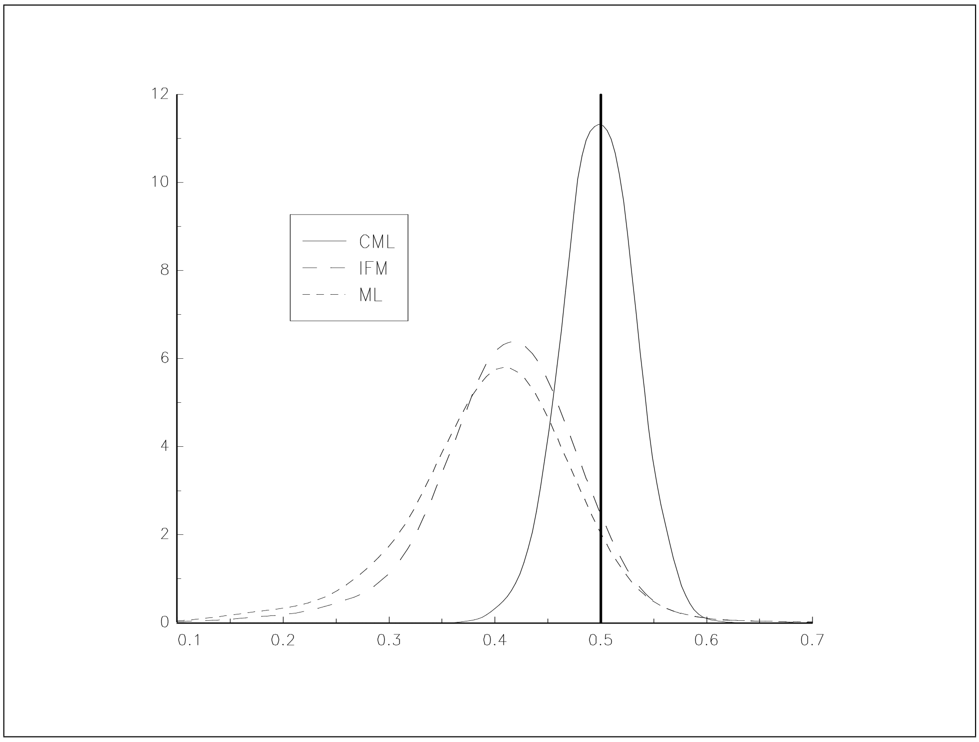


Figure 9: Comparison of the density of the CML IFM and ML estimators when the margins are wrong

[2] Barbe, P., C. Genest, K. Ghoudi and B. Rémillard [1996], On Kendall's process, *Journal of Multivariate Analysis*, **58**, 197-229

[3] Beneš, V. and J. Štěpán [1997], Distributions with Given Marginals and Moment Problems, Kluwer Academic Publishers, Dordrecht

[4] Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet and T. Roncalli [2000], Copulas for finance — A reading guide and some applications, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper*

[5] Coles, S., J. Currie and J. Tawn [1999], Dependence measures for extreme value analyses, Department of Mathematics and Statistics, Lancaster University, Working Paper

[6] Darsow, W.F., B. Nguyen and E.T. Olsen [1992], Copulas and markov processes, *Illinois Journal of Mathematics*, **36-4**, 600-642

[7] Davidson, R. and J. MacKinnon [1993], Estimation and Inference in Econometrics, Oxford University Press, Oxford

[8] Deheuvels, P. [1978], Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes, *Publications de l'Institut de Statistique de l'Université de Paris*, **23**, 1-36

[9] Deheuvels, P. [1979], La fonction de dépendance empirique et ses propriétés — Un test non paramétrique d'indépendance, *Académie Royale de Belgique – Bulletin de la Classe des Sciences – 5e Série*, **65**, 274-292

[10] Deheuvels, P. [1981], A non parametric test for independence, *Publications de l'Institut de Statistique de l'Université de Paris*, **26**, 29-50

[11] Durrleman, V., A. Nikeghbali, and T. Roncalli [2000], Copulas approximation and new families, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper*

[12] Embrechts, P., McNeil, A.J. and D. Straumann [1999], Correlation and dependency in risk management: properties and pitfalls, Departement of Mathematik, ETHZ, Zürich, *Working Paper*

[13] Frees, E.W. and E.A. Valdez, [1998], Understanding relationships using copulas, *North American Actuarial Journal*, **2**, 1-25

[14] Genest, C. and J. MacKay [1986], The joy of copulas: Bivariate distributions with uniform marginals, *American Statistician*, **40**, 280-283

[15] Genest, C. and L. Rivest [1993], Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association*, **88**, 1034-1043

[16] Joe, H. [1997], Multivariate Models and Dependence Concepts, *Monographs on Statistics and Applied Probability*, **73**, Chapmann & Hall, London

[17] Joe, H. and J.J. Xu [1996], The estimation method of inference functions for margins for multivariate models, Department of Statistics, University of British Columbia, *Technical Report*, **166**

[18] Lehmann, E.L. and G. Casella [1998], Theory of Point Estimation, second edition, Springer Verlag, New York

[19] Li, X., P. Mikusiński, H. Sherwood and M.D. Taylor [1997], On approximation of copulas, in V. Beneš and J. Štěpán (Eds.), *Distributions with Given Marginals and Moment Problems*, Kluwer Academic Publishers, Dordrecht

[20] Nelsen, R.B. [1998], An Introduction to Copulas, *Lectures Notes in Statistics*, **139**, Springer Verlag, New York

[21] Prause, K. [1999], The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures, *Dissertation*, University of Freiburg