A Short, Comprehensive, Practical Guide to Copulas

Attilio Meucci[1]
attilio_meucci@symmys.com

published[2]: October 1 2011
this revision: September 9, 2011
last revision and code available at `http://symmys.com/node/351`

Abstract

We provide a visual primer on copulas. We highlight and prove the most important theoretical results. We discuss implementation issues. Documented code and the latest revision of this article are available at http://symmys.com/node/351.

JEL Classification: C1, G11

Keywords: marginal, Sklar's theorem, grade, co-monotonic, linear returns, compounded returns, unit cube, scenarios-probabilities, pdf, cdf, quantile, Copula-Marginal-Algorithm.

[1]The author is grateful to Garli Beibi
[2]This article appears as Meucci, A. (2011), *A Short, Comprehensive, Practical Guide to Copulas* - GARP Risk Professional, October 2011, p. 22-27

1 Introduction

The stochastic behavior of one single financial variable, say prices, or implied volatilities, etc., is fully described by its probability distribution, which is called the marginal distribution. In a market of multiple financial variables, all the information on the stochastic behavior of the market is fully described by the joint probability distribution.

The multivariate distribution of a set of financial variables is fully specified by the separate marginal distributions of the variables and by their copula, or, loosely speaking, the correlations among the variables.

Modeling the marginals and the copula separately provides greater flexibility for the practitioner to model randomness. Therefore, copulas have been used extensively in finance, both on the sell-side to price derivatives, see e.g. Li (2000), and on the buy-side to model portfolio risk, see e.g. Meucci, Gan, Lazanas, and Phelps (2007).

Here we provide a review of the theory of copulas proving the most useful results. For more background on the subject, the reader is referred to articles such as Embrechts, A., and Straumann (2000), Durrleman, Nikeghbali, and Roncalli (2000), Embrechts, Lindskog, and McNeil (2003), or monographs such as to Nelsen (1999), Cherubini, Luciano, and Vecchiato (2004), Brigo, Pallavicini, and Torresetti (2010) and Jaworski, Durante, Haerdle, and Rychlik (2010). We also provide a guide to the practical implementation of copulas. For a detailed discussion of implementation issues and for the code, please refer to the companion paper Meucci (2011).

This article is organized as follows. In Section 2 we review strictly univariate results that nonetheless naturally lead to the multivariate concept of copulas. In Section 3 we introduce copulas, highlighting and proving the most important theoretical results. In Section 4 we address copulas implementation issues.

2 Univariate results

In this section we cover well-known results that prepare the ground for the definition of, and the intuition behind, copulas.

Consider an arbitrary random variable X, with a fully arbitrary distribution. All the features of the distribution of X are described by its probability density function (pdf) f_X, defined in such a way that, for any set of potential values $\mathcal{X}$ for the variable X, the following identity for the pdf holds

$$\mathbb{P}\{X \in \mathcal{X}\} \equiv \int_{\mathcal{X}} f_X(x)\, dx. \tag{1}$$

Accordingly, with mild abuse of notation we write $X \sim f_X$ to denote that X has a distribution whose pdf is f_X. An alternative way to represent the distribution of X is its cumulative distribution function (cdf), defined as follows

$$F_X(x) \equiv \int_{-\infty}^{x} f_X(z)\, dz. \tag{2}$$

If we feed the random variable X into a generic function g we obtain another random variable $Y \equiv g(X)$. For instance, if we set $g(x) \equiv \sin(x)$ we obtain a random variable $Y \equiv \sin(X)$, which is bound in the interval $[-1, 1]$. A special situation arises when we transform X with its own cdf, i.e. $g \equiv F_X$.

Key concept. If we feed the arbitrary variable X through its own cdf, we obtain a very special transformed random variable, which is called the grade of X

$$U \equiv F_X(X). \tag{3}$$

The distribution of the grade is uniform on the unit interval regardless of the original distribution f_X

$$U \sim \mathsf{U}_{[0,1]}, \tag{4}$$

The simple proof of this result proceeds as follows

$$\begin{aligned} F_U(u) &\equiv \mathbb{P}\{U \leq u\} = \mathbb{P}\{F_X(X) \leq u\} \\ &= \mathbb{P}\left\{X \leq F_X^{-1}(u)\right\} = F_X\left(F_X^{-1}(u)\right) = u. \end{aligned} \tag{5}$$

Therefore $F_U(u) = u$, which is the cdf of a uniform distribution.

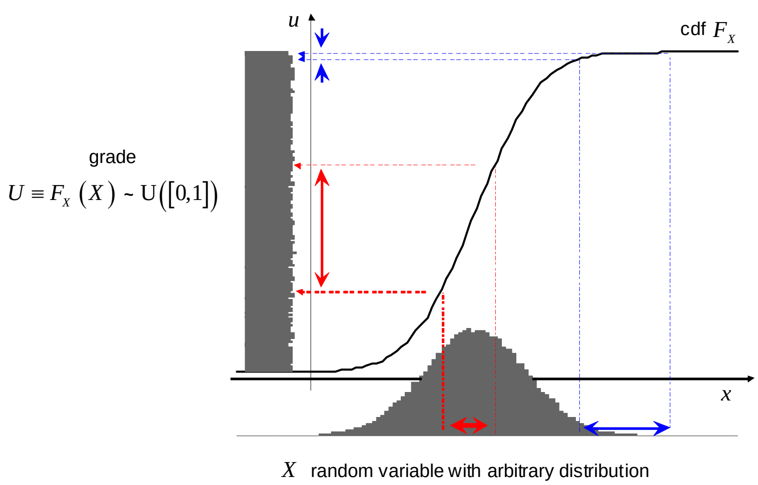

Figure 1: The cdf maps an arbitrary random variable into a uniform variable

In Figure 1 we sketch the intuition behind (4). First of all the variable U lives in the interval $[0, 1]$ because the cdf satisfies $0 \leq F_X(x) \leq 1$. Furthermore, the variable U is uniform on $[0, 1]$ because the cdf F_X is steeper where there are more potential outcomes for X and thus these outcomes are spread out over a wider interval; on the other hand, the cdf F_X is flatter where there are few outcomes, and thus the cdf concentrates all the outcomes occurring over a large interval into a small set. The balance between these two effects, i.e. dilution of abundant scenarios and concentration of scarce scenarios, gives rise to a uniform distribution.

The above result (3)-(4) also works backwards: if we feed a uniform random variable U into the inverse cdf F_X^{-1} we obtain a random variable $X \equiv F_X^{-1}(U)$ with distribution f_X

$$U \sim \mathsf{U}_{[0,1]} \quad \mapsto \quad X \equiv F_X^{-1}(U) \sim f_X. \tag{6}$$

Even better, we can choose any desired target distribution $\bar{f}_X$, say Student t or chi-square, etc.; then we compute the cdf $\bar{F}_X$ and the inverse cdf $\bar{F}_X^{-1}$; and finally transform a uniform variable U to generate a random variable X with the desired arbitrary distribution.

Key concept. Starting from an arbitrary target distribution and a uniform random variable, we can transform the uniform variable into a variable with the desired target distribution

$$\left.\begin{array}{l} \bar{f}_X \\ \\ U \sim \mathsf{U}_{[0,1]} \end{array}\right\} \quad \mapsto \quad X \equiv \bar{F}_X^{-1}(U) \sim \bar{f}_X. \tag{7}$$

The proof of (7) reads

$$\mathbb{P}\{X \leq x\} = \mathbb{P}\left\{\bar{F}_X^{-1}(U) \leq x\right\} = \mathbb{P}\left\{U \leq \bar{F}_X(x)\right\} = \bar{F}_X(x). \tag{8}$$

This result is extremely useful to generate Monte Carlo scenarios from arbitrary desired distributions using as input only a uniform number generator.

3 Copulas: theory

Now we are fully equipped to introduce the copula, by extending to the multivariate framework the univariate results (3)-(4) and (7).

Consider a N-dimensional vector of random variables $\mathbf{X} \equiv (X_1, \ldots, X_N)'$ with a fully general multivariate distribution represented by its pdf $\mathbf{X} \sim f_{\mathbf{X}}$. We recall that in the multivariate case the pdf $f_{\mathbf{X}}$ is defined in such a way that, for any set of potential joint values $\mathcal{X} \in \mathbb{R}^N$ for $(X_1, \ldots, X_N)$ the following identity holds

$$\mathbb{P}\{(X_1, \ldots, X_N) \in \mathcal{X}\} \equiv \int_{\mathcal{X}} f_{\mathbf{X}}(x_1, \ldots, x_N)\, dx_1 \cdots dx_N. \tag{9}$$

From the joint distribution $f_{\mathbf{X}}$ we can in principle extract all the N marginal distributions $X_n \sim f_{X_n}$, where $n = 1, \ldots, N$, by computing the marginal pdf's as follows

$$f_{X_n}(x_n) = \int_{\mathbb{R}^{N-1}} f_{\mathbf{X}}(x_1, \ldots, x_N)\, dx_1 \cdots dx_{n-1} dx_{n+1} \cdots dx_N. \tag{10}$$

Then we can compute the marginal cdf's F_{X_n} as in (2). Finally, we can feed each cdf F_{X_n}, which is a function, with the respective entry of the vector $\mathbf{X}$, namely the random variable X_n. The outcome of this operation are the grades, which we know from (4) have a uniform distribution on the unit interval

$$U_n \equiv F_{X_n}(X_n) \sim \mathsf{U}_{[0,1]}. \tag{11}$$

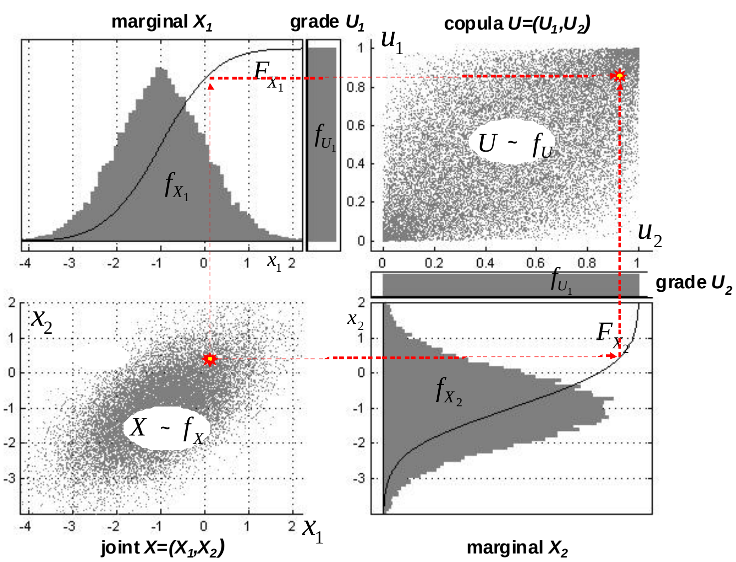

Figure 2: Copulas are non-linear standardizations of multivariate distributions

However, the main point to remember is that the entries of $\mathbf{U} \equiv (U_1, \ldots, U_N)'$ are *not* independent. Therefore, the joint distribution $f_{\mathbf{U}}$ of the grades is not uniform on its domain, which is the unit cube $[0,1] \times \cdots \times [0,1]$, see Figure 2.

Key concept. The copula of an arbitrary distribution $f_{\mathbf{X}}$ is the joint distribution $f_{\mathbf{U}}$ of its grades

$$\begin{pmatrix} U_1 \equiv F_{X_1}(X_1) \\ \vdots \\ U_N \equiv F_{X_N}(X_N) \end{pmatrix} \sim f_{\mathbf{U}} \tag{12}$$

The grades of the distribution $f_{\mathbf{X}}$ can be interpreted as a sort of non-linear z-score, which forces all the entries X_n to have a uniform distribution on the unit interval $[0,1]$. By feeding each random variable X_n into its own cdf, all the information contained in each marginal distribution f_{X_n} is swept away, and what is left is the pure joint information amongst the X_n's, i.e. the copula $f_{\mathbf{U}}$. We summarize this statement in an alternative, intuitive formulation of the copula

> **Key concept**. the copula is the information missing from the individual marginals to complete the joint distribution
>
> $$\text{" joint = copula + marginals "} \tag{13}$$

The intuitive definition (13) can be made rigorous. From the definition of copula (12)

$$\begin{aligned} F_{\mathbf{U}}(\mathbf{u}) &\equiv \mathbb{P}\{U_1 \leq u_1, \ldots, U_N \leq u_N\} \\ &= \mathbb{P}\{F_{X_1}(X_1) \leq u_1, \ldots, F_{X_N}(X_N) \leq u_N\} \\ &= \mathbb{P}\left\{X_1 \leq F_{X_1}^{-1}(u_1), \ldots, X_N \leq F_{X_N}^{-1}(u_N)\right\} \\ &= F_X\left(F_{X_1}^{-1}(u_1), \ldots, F_{X_N}^{-1}(u_N)\right). \end{aligned} \tag{14}$$

Differentiating this expression we obtain Sklar's theorem (for two variables, the general case follows immediately)

$$\begin{aligned} f_{\mathbf{U}}(u_1, u_2) &= \partial^2_{u_1 u_2} F_{\mathbf{U}}(u_1, u_2) = \partial^2_{u_1 u_2} F_{\mathbf{X}}\left(F_{X_n}^{-1}(u_1), F_{X_2}^{-1}(u_2)\right) \\ &= \partial^2_{x_1 x_2} F_{\mathbf{X}}\left(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)\right) d_{u_1} F_{X_1}^{-1}(u_1)\, d_{u_2} F_{X_2}^{-1}(u_2) \\ &= \frac{\partial^2_{x_1 x_2} F_{\mathbf{X}}\left(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)\right)}{d_{x_1} F_{X_1}\left(F_{X_1}^{-1}(u_1)\right) d_{x_2} F_{X_2}\left(F_{X_2}^{-1}(u_2)\right)} \\ &= \frac{f_{\mathbf{X}}\left(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)\right)}{f_{X_1}\left(F_{X_1}^{-1}(u_1)\right) f_{X_2}\left(F_{X_2}^{-1}(u_2)\right)}. \end{aligned} \tag{15}$$

> **Key concept**. Sklar's theorem links the original joint distribution $f_{\mathbf{X}}$, the copula $f_{\mathbf{U}}$, and the marginals f_{X_n}, or equivalently F_{X_n}, as follows
>
> $$\begin{aligned} \underbrace{f_{\mathbf{X}}\left(F_{X_1}^{-1}(u_1), \ldots, F_{X_N}^{-1}(u_N)\right)}_{\text{joint}} &= \\ \underbrace{f_{\mathbf{U}}(u_1, \ldots, u_N)}_{\text{pure joint}} &\times \underbrace{f_{X_1}\left(F_{X_1}^{-1}(u_1)\right) \times \cdots \times f_{X_N}\left(F_{X_N}^{-1}(u_N)\right)}_{\text{pure marginal}}. \end{aligned} \tag{16}$$
>
> Sklar's theorem justifies the intuitive copula definition (13).

Sklar's theorem provides the pdf of the copula from the joint pdf and the marginal pdf's. This allows us to use maximum likelihood to fit copulas to empirical data.

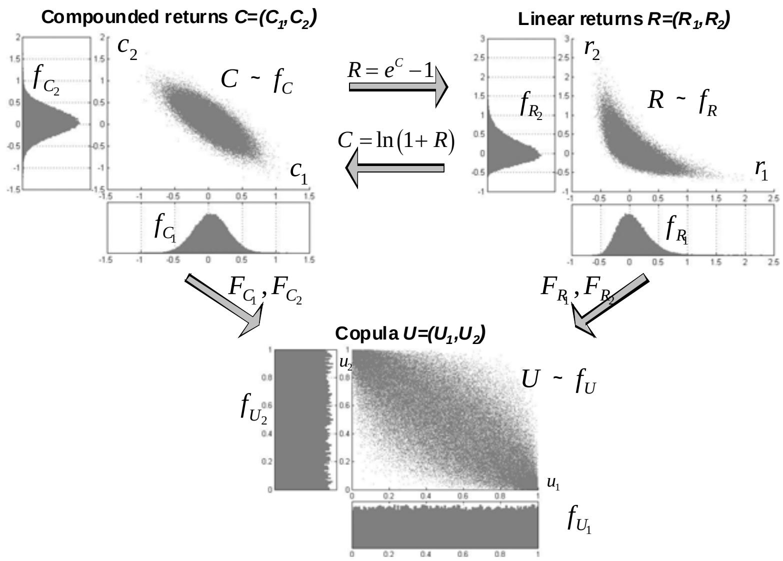

Figure 3: Linear and compounded returns have same copula

We now derive another useful result. If we feed the grades (12) into arbitrary inverse cdf's $F_{Y_n}^{-1}$, we obtain new transformed random variables with a given joint distribution, which we denote by $f_{\mathbf{Y}}$

$$\left(\begin{array}{c} Y_1 \equiv F_{Y_1}^{-1}\left(F_{X_1}\left(X_1\right)\right) \\ \vdots \\ Y_N \equiv F_{Y_N}^{-1}\left(F_{X_N}\left(X_N\right)\right) \end{array}\right) \sim f_{\mathbf{Y}} \tag{17}$$

The joint distribution f_Y has marginals whose cdf's are F_{Y_n}, which follows from applying (11) and (6) in sequence. Furthermore, the copula of $\mathbf{Y}$ is the same as the copula of $\mathbf{X}$, because from the definition of $\mathbf{Y}$ in (17) and the definition of the copula (12) we obtain

$$\left(\begin{array}{c} F_{Y_1}\left(Y_1\right) = F_{X_1}\left(X_1\right) \\ \vdots \\ F_{Y_N}\left(Y_N\right) = F_{X_N}\left(X_N\right) \end{array}\right) \sim f_{\mathbf{U}}. \tag{18}$$

Therefore, we derive that the copula of an arbitrary random variable $\mathbf{X} \equiv \left(X_1, \ldots, X_N\right)'$ does not change when we transform each X_n into a new variable $Y_n \equiv g_n\left(X_n\right)$ by means of functions $g_n\left(x\right) \equiv F_{Y_n}^{-1}\left(F_{X_n}\left(x\right)\right)$, where F_{Y_n} are *arbitrary* cdf's. It is easy to verify that such functions g_n are a very broad class, namely all the increasing transformations, also known as co-monotonic transformations. Thus we obtain the following

> **Key concept**. Co-monotonic transformations $Y_n = g_n(X_n)$ of the entries of X do not alter the copula of X
>
> $$\begin{array}{ccccc} \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \sim f_{\mathbf{X}} & & \overset{g_1,\ldots,g_N}{\longrightarrow} & & \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix} \sim f_{\mathbf{Y}} \\ & \overset{F_{X_1},\ldots,F_{X_N}}{\searrow} & & \overset{F_{Y_1},\ldots,F_{Y_N}}{\swarrow} & \\ & & \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \sim f_{\mathbf{U}} & & \end{array} \tag{19}$$

To illustrate how the copula is not affected by increasing transformations, consider the linear returns and the compounded returns between time t and time $t+1$ for the *same* securities

$$R_n \equiv \frac{P_{n,t+1}}{P_{n,t}} - 1, \quad C_n \equiv \ln(\frac{P_{n,t+1}}{P_{n,t}}). \tag{20}$$

These two types of returns, though calculated on the same securities prices $P_{n,t+1}$, are different. Therefore, their distributions are different, refer to Meucci (2010) for more details on the pitfalls of disregarding such differences. For example, if the prices distribution $f_{\mathbf{P}}$ is multivariate log-normal, the linear returns distribution $f_{\mathbf{R}}$ is multivariate shifted-lognormal and the compounded returns distribution $f_{\mathbf{C}}$ is multivariate normal, as illustrated in Figure 3 for two negatively correlated stocks.

However, the copula of the linear returns and the copula of the compounded returns are identical, see again Figure 3. This result is not surprising, because $R_n = e^{C_n} - 1$ is an increasing transformation of C_n and $C_n = \ln(1 + R_n)$ is an increasing transformation of R_n.

4 Copulas: practice

The implementation of the copula-marginal decomposition in practice relies on two distinct processes, which appear in multiple steps in the theoretical discussion of Section 3.

First, the separation process, which led us to the definition of the copula (12).

Key concept. The separation process $\mathcal{S}$ strips an arbitrary distribution $f_{\mathbf{X}}$ into its marginals f_{X_n} and its copula $f_{\mathbf{U}}$

$$\mathcal{S}: \quad \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \sim f_{\mathbf{X}} \quad \mapsto \quad \begin{cases} f_{X_1}, \ldots, f_{X_N} \\ \\ \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \sim f_{\mathbf{U}}, \end{cases} \tag{21}$$

where U_n are the grades

$$U_n \equiv F_{X_n}(X_n). \tag{22}$$

The separation process can be reverted, similarly to the univariate case (6). By feeding each grade U_n back into the respective inverse cdf $F_{X_n}^{-1}$ we obtain a random variable $\mathbf{X} \equiv (X_1, \ldots, X_N)$ whose joint distribution is exactly the original distribution $f_{\mathbf{X}}$:

$$\begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \sim f_{\mathbf{U}} \quad \mapsto \quad \begin{pmatrix} X_1 \equiv F_{X_1}^{-1}(U_1) \\ \vdots \\ X_N \equiv F_{X_N}^{-1}(U_N) \end{pmatrix} \sim f_{\mathbf{X}} \tag{23}$$

However, we do not need to limit ourselves to reverting to the original distribution $f_{\mathbf{X}}$. Copulas are so powerful because they can be glued with arbitrary marginal distributions, with a technique that generalizes the univariate case (7), and which we used to prove the co-monotonic invariance of the copula (19).

Accordingly, we start with two ingredients: an arbitrary copula $\bar{f}_{\mathbf{U}}$, i.e. grades $\mathbf{U} \equiv (U_1, \ldots, U_N)'$ each of which has a uniform distribution, and joint distribution structure specified by $\bar{f}_{\mathbf{U}}$; and arbitrary marginal distributions $\bar{f}_{X_n}$. Then we compute the marginal cdf's $\bar{F}_{X_n}$, their inverses $\bar{F}_{X_n}^{-1}$ and we feed each grade into the respective marginal cdf. The output is a N-variate random variable $\mathbf{X} \equiv (X_1, \ldots, X_N)'$ that has the desired copula $\bar{f}_{\mathbf{U}}$ and the desired marginals $\bar{f}_{X_n}$.

We summarize as follows this second process.

Key concept. The combination process $\mathcal{C}$ glues arbitrary marginals $\bar{f}_{X_n}$ and an arbitrary copula $\bar{f}_{\mathbf{U}}$ into a new joint distribution $\bar{f}_{\mathbf{X}}$

$$\mathcal{C}: \quad \left. \begin{array}{c} \bar{f}_{X_1}, \ldots, \bar{f}_{X_N} \\ \\ \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \sim \bar{f}_{\mathbf{U}} \end{array} \right\} \quad \mapsto \quad \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \sim \bar{f}_{\mathbf{X}}, \tag{24}$$

where

$$X_n \equiv \bar{F}_{X_n}^{-1}(U_n). \tag{25}$$

To implement copulas in practice, we must be able to implement the separation process (21), in order to obtain suitable copulas, and the combination process (24), in order to glue those copulas with suitable marginals.

Implementing all the steps involved in such processes analytically is impossible, except in trivial cases, such as within the normal family.

Therefore, all practical applications of copulas rely on numerical techniques, most notably the representation of distributions by means of Monte Carlo scenarios.

Within the Monte Carlo framework, scenarios that represent copulas are obtained by feeding joint distributions scenarios into the respective marginal cdf's as in (22); and joint scenarios with a given copula are obtained by feeding grades scenarios into the respective inverse cdf's as in (25).

However, even the above numerical operations present difficulties. First, the computation of the cdf's in (22) requires univariate integrations as in (2). Second, the computation of the inverse cdf in (25) requires univariate integrations followed by search algorithms. Finally the extraction of the marginal distributions from the joint distribution in (21) requires multivariate integrations as in (10).

The first two problems, namely the univariate integration and inversion, have been addressed for a broad variety of parametric distributions, for which cdf and quantile are available either analytically of in terms of efficient quadratures.

On the other hand, the multivariate integration to extract the marginals from the joint distribution represents a significant hurdle, because numerical integration is practical only in markets of very low dimension N. For larger markets, one must resort to analytical or quasi-analytical formulas. Such formulas are available only for a handful of distributions, most notably the elliptical family, which includes the normal and the Student t distributions.

An alternative to avoid the multivariate integration is to draw scenarios directly from parametric copulas. However, the parametric specifications that allow for direct simulation are limited to the Archimedean family, see Genest and Rivest (1993), and few other extensions. Furthermore, the parameters of the Archimedean family are not immediate to interpret. Finally, simulating grades scenarios from the Archimedean family when the dimension N is large is computationally challenging.

To summarize, traditional implementations of copulas mainly proceed as follows: first, Monte Carlo scenarios are drawn from elliptical or related distributions; next, the scenarios are channelled through the respective (quasi-)analytical marginal cdf's as in (22), thereby obtaining grade scenarios; then, the grade scenarios are fed into flexible parametric quantiles as in (25), thereby obtaining the desired joint scenarios.

To avoid the restrictive assumptions of the traditional copula implementation and circumvent all the above problems Meucci (2011) proposes the Copula-Marginal Algorithm (CMA), which simulates Monte Carlo scenarios with *flexible probabilities* from *arbitrary* distributions; computes the marginal cdf's *without integrations*; and *avoids the quantile* computation. CMA is numerically extremely efficient. For all the details, the code, and an application to stress-testing with panic markets, please refer to Meucci (2011).

References

Brigo, D., A. Pallavicini, and R. Torresetti, 2010, *Credit Models and the Crisis: A Journey Into CDOs, Copulas, Correlations and Dynamic Models* (Wiley).

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

Durrleman, V., A. Nikeghbali, and T. Roncalli, 2000, Which copula is the right one?, *Working Paper.*

Embrechts, P., McNeil A., and D. Straumann, 2000, Correlation: Pitfalls and alternatives, *Working Paper.*

Embrechts, P., F. Lindskog, and A. J. McNeil, 2003, Modelling dependence with copulas and applications to risk management, *Handbook of heavy tailed distributions in finance.*

Genest, C., and R. Rivest, 1993, Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association* 88, 1034– 1043.

Jaworski, P., F. Durante, W. Haerdle, and T. (Editors) Rychlik, 2010, *Copula Theory and its Applications* (Springer, Lecture Notes in Statistics - Proceedings).

Li, D. X., 2000, On default correlation: A copula function approach, *Journal of Fixed Income* 9, 43–54.

Meucci, A., 2010, Linear vs. compounded returns - common pitfalls in portfolio management, *GARP Risk Professional* April, 52–54 Article and code available at http://symmys.com/node/141.

———, 2011, A new breed of copulas for risk and portfolio management, *Risk* September, 122–126 Article and code available at http://symmys.com/node/335.

———, Y Gan, A. Lazanas, and B. Phelps, 2007, A portfolio mangers guide to Lehman Brothers tail risk model, *Lehman Brothers Publications.*

Nelsen, R. B., 1999, *An Introduction to Copulas* (Springer).