How to get bounds for distribution convolutions? A simulation study and an application to risk management

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

Groupe de Recherche Opérationnelle
Crédit Lyonnais
France

September 10, 2000

Abstract

In this paper, we consider the problem of bounds for distribution convolutions and we present some applications to risk management. We show that the upper Fréchet bound is not always the more risky dependence structure. It is in contradiction with the belief in finance that maximal risk correspond to the case where the random variables are comonotonic.

1 Introduction

We consider here two problems applied to finance. The first one is the Kolmogorov problem and we show how Makarov inequalities could be used in risk management. The second one is based on the Kantorovich distance between two distributions. For this two problem, we show that maximal risk does not correspond to the case where the random variables are comonotonic.

2 Distribution convolutions and Makarov inequalities

In this section, we apply the Makarov inequalities to risk aggregation.

2.1 The triangle functions and $\sigma-$operations

We note Δ^+ the space of probability distribution functions whose support is contained in $\mathbb{R}^+$. Schweizer [1991] gives the following definition for a *triangle function*:

Definition 1 *Given a function $\tau : \Delta^+ \times \Delta^+ \longrightarrow \Delta^+$, τ is a triangle function if it satisfies*[1]

1. $\tau(\mathbf{F}, \varepsilon_0) = \mathbf{F}$;
2. $\tau(\mathbf{F}_1, \mathbf{F}_2) = \tau(\mathbf{F}_2, \mathbf{F}_1)$;
3. $\tau(\mathbf{F}_1, \mathbf{F}_2) \leq \tau(\mathbf{G}_1, \mathbf{G}_2)$ *whenever* $\mathbf{F}_1(x) \leq \mathbf{G}_1(x)$ *and* $\mathbf{F}_2(x) \leq \mathbf{G}_2(x)$ *for* $x \in \mathbb{R}^+$;

* *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] ε_0 is the unit step-function

$$\varepsilon_0(x) = \begin{array}{lll} 0 & \text{if} & x \leq 0 \\ 1 & \text{if} & x > 0 \end{array} \tag{1}$$

4. $\tau(\tau(\mathbf{F}_1,\mathbf{F}_2),\mathbf{F}_3)=\tau(\mathbf{F}_1,\tau(\mathbf{F}_2,\mathbf{F}_3))$*;*

SCHWEIZER and SKLAR [1983] define also a probabilistic metric space as the ordered triple $(S,\digamma,\tau)$ where S is a set, $\digamma$ is a mapping from $S\times S$ into Δ^+ and τ a triangle function. SCHWEIZER [1991] does the following interpretation:

> *In short, probabilistic metric spaces are generalizations of ordinary metric spaces in which R^+ is replaced by the set Δ^+ and the operation of addition on R^+ is replaced by a triangle function.*

Two triangle functions play a special role. Let $\mathcal{B}$ be the class of Borel-measurable two place functions. Let X_1 and X_2 be two random variables with distributions $\mathbf{F}_1$ and $\mathbf{F}_2$. We then consider the random variable $X=L(X_1,X_2)$ with distribution $\mathbf{G}$ and $L\in\mathcal{B}$. The *supremal convolution* $\tau_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)$ is

$$\tau_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)(x)=\sup_{L(x_1,x_2)=x}\mathbf{C}(\mathbf{F}_1(x_1),\mathbf{F}_2(x_2)) \tag{2}$$

whereas the *infimal convolution* $\rho_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)$ corresponds to

$$\rho_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)(x)=\inf_{L(x_1,x_2)=x}\tilde{\mathbf{C}}(\mathbf{F}_1(x_1),\mathbf{F}_2(x_2)) \tag{3}$$

with $\tilde{\mathbf{C}}$ the dual of the copula $\mathbf{C}$. Another important function is the $\sigma-$*convolution* (FRANK [1991]) defined by

$$\sigma_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)(x)=\int_{L(x_1,x_2)<x}\mathrm{d}\mathbf{C}(\mathbf{F}_1(x_1),\mathbf{F}_2(x_2)) \tag{4}$$

In general, we have

$$\tau_{\mathbf{C}^-,L}(\mathbf{F}_1,\mathbf{F}_2)\le\tau_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)\le\sigma_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)\le\rho_{\mathbf{C},L}(\mathbf{F}_1,\mathbf{F}_2)\le\rho_{\mathbf{C}^-,L}(\mathbf{F}_1,\mathbf{F}_2) \tag{5}$$

2.2 The dependency bounds

Dependency bounds are related to a Kolmogorov's problem (MAKAROV [1981]): find distribution function $\mathbf{G}_\vee$ and $\mathbf{G}_\wedge$ such that for all x in $\mathbb{R}$, $\mathbf{G}_\vee(x)=\inf\Pr\{X_1+X_2<x\}$ and $\mathbf{G}_\wedge(x)=\sup\Pr\{X_1+X_2<x\}$ where the infimum and supremum are taken over all possible joint distribution functions having margins $\mathbf{F}_1$ and $\mathbf{F}_2$. The solution is then $\mathbf{G}_\vee=\tau_{\mathbf{C}^-,+}(\mathbf{F}_1,\mathbf{F}_2)$ and $\mathbf{G}_\wedge=\rho_{\mathbf{C}^-,+}(\mathbf{F}_1,\mathbf{F}_2)$.

This solution has a long history (see the section eight of SCHWEIZER [1991]). Note that it holds even for non positive random variables and moreover these bounds cannot be improved (FRANK, NELSEN and SCHWEIZER [1987]). WILLIAMSON [1989] extends this result when $\mathbf{C}_-\prec\mathbf{C}$ and L denotes the four arithmetic operators ($+$, $-$, $\times$ and $\div$).

Theorem 2 *Let X_1 and X_2 be two positive random variables with distributions $\mathbf{F}_1$ and $\mathbf{F}_2$ and dependence structure $\mathbf{C}$ such that $\mathbf{C}_-\prec\mathbf{C}$. Then, the distribution $\mathbf{G}$ of $X=L(X_1,X_2)$ is contained within the bounds $\mathbf{G}_\vee(x)\le\mathbf{G}(x)\le\mathbf{G}_\wedge(x)$ with $\mathbf{G}_\vee(x)=\tau_{\mathbf{C}_-,L}(\mathbf{F}_1,\mathbf{F}_2)(x)$ and $\mathbf{G}_\wedge(x)=\rho_{\mathbf{C}_-,L}(\mathbf{F}_1,\mathbf{F}_2)(x)$.* ***These bounds are the pointwise best possible.***

WILLIAMSON and DOWNS [1990] remark that knowing a tighter lower bound than the lower Fréchet bound ($\mathbf{C}^-\prec\mathbf{C}_-\prec\mathbf{C}$) provide tighter bounds $\mathbf{G}_\vee$ and $\mathbf{G}_\wedge$, but knowing a tighter upper bound than the upper Fréchet bound ($\mathbf{C}\prec\mathbf{C}_+\prec\mathbf{C}^+$) has **curiously** no effect on the dependency bounds!

2.3 Duality of infimal and supremal convolutions

Let $\mathbf{F}$ be a distribution. We note $\mathbf{F}^{(-1)}$ the inverse of $\mathbf{F}$, that is $\mathbf{F}^{(-1)}(u) = \sup\{x \mid \mathbf{F}(x) < u\}$. Frank and Schweizer [1979] show then that $\tau_{\mathbf{C},L}(\mathbf{F}_1, \mathbf{F}_2) = \breve{\tau}^{(-1)}_{\mathbf{C},L}\left(\mathbf{F}_1^{(-1)}, \mathbf{F}_2^{(-1)}\right)$ and $\rho_{\mathbf{C},L}(\mathbf{F}_1, \mathbf{F}_2) = \breve{\rho}^{(-1)}_{\mathbf{C},L}\left(\mathbf{F}_1^{(-1)}, \mathbf{F}_2^{(-1)}\right)$ with

$$\breve{\tau}_{\mathbf{C},L}\left(\mathbf{F}_1^{(-1)}, \mathbf{F}_2^{(-1)}\right)(u) = \inf_{\mathbf{C}(u_1,u_2)=u} L\left(\mathbf{F}_1^{(-1)}(u_1), \mathbf{F}_2^{(-1)}(u_2)\right) \tag{6}$$

and

$$\breve{\rho}_{\mathbf{C},L}\left(\mathbf{F}_1^{(-1)}, \mathbf{F}_2^{(-1)}\right)(u) = \sup_{\tilde{\mathbf{C}}(u_1,u_2)=u} L\left(\mathbf{F}_1^{(-1)}(u_1), \mathbf{F}_2^{(-1)}(u_2)\right) \tag{7}$$

Williamson and Downs [1990] derive then numerical algorithms to compute the dependency bounds $\mathbf{G}_\vee$ and $\mathbf{G}_\wedge$. For example, if $\mathbf{C}_- = \mathbf{C}^-$ and L is the operation $+$, we have

$$\begin{aligned}\mathbf{G}_\vee^{(-1)}(u) &= \inf_{\max(u_1+u_2-1,0)=u} \mathbf{F}_1^{(-1)}(u_1) + \mathbf{F}_2^{(-1)}(u_2) \\ &= \inf_{v\in[u,1]} \mathbf{F}_1^{(-1)}(v) + \mathbf{F}_2^{(-1)}(u-v+1)\end{aligned} \tag{8}$$

and

$$\begin{aligned}\mathbf{G}_\wedge^{(-1)}(u) &= \sup_{\min(u_1+u_2,1)=u} \mathbf{F}_1^{(-1)}(u_1) + \mathbf{F}_2^{(-1)}(u_2) \\ &= \sup_{v\in[0,u]} \mathbf{F}_1^{(-1)}(v) + \mathbf{F}_2^{(-1)}(u-v)\end{aligned} \tag{9}$$

2.4 Some simulations

Frank, Nelsen and Schweizer [1987] derive analytical expressions of the dependency bounds in the case of gaussian, Cauchy and exponential distributions when $\mathbf{C}_- = \mathbf{C}^-$ and L is the operation $+$. We present now some simulations based on the numerical algorithms of Williamson and Downs [1990].

We consider the example when X_1 and X_2 are two gaussian distributions $\mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{N}(\mu_2, \sigma_2)$. If the dependence structure is Normal with parameter ρ, it comes that the distribution $\mathbf{G}$ of $X_1 + X_2$ is gaussian and we have $\mathbf{G} = \mathcal{N}\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2}\right)$. The lower and upper dependency bounds are given by equations (8) and (9) when $\mathbf{C}_- = \mathbf{C}^-$. If $\mathbf{C}_- = \mathbf{C}^\perp$, $\mathbf{G}_\vee$ and $\mathbf{G}_\wedge$ becomes

$$\mathbf{G}_\vee^{(-1)}(u) = \inf_{v\in[0,1]} \mathbf{F}_1^{(-1)}(v) + \mathbf{F}_2^{(-1)}\left(\frac{u}{v}\right) \tag{10}$$

and

$$\mathbf{G}_\wedge^{(-1)}(u) = \sup_{v\in[0,1]} \mathbf{F}_1^{(-1)}(v) + \mathbf{F}_2^{(-1)}\left(\frac{u-v}{1-v}\right) \tag{11}$$

In the case where $\mathbf{C}_- = \mathbf{C}^+$, we obtain the following results

$$\begin{aligned}\mathbf{G}_\vee^{(-1)}(u) &= \inf_{\min(u_1,u_2)=u} \mathbf{F}_1^{(-1)}(u_1) + \mathbf{F}_2^{(-1)}(u_2) \\ &= \sup_{\max(u_1,u_2)=u} \mathbf{F}_1^{(-1)}(u_1) + \mathbf{F}_2^{(-1)}(u_2) \\ &= \sup_{u_1+u_2-\min(u_1,u_2)=u} \mathbf{F}_1^{(-1)}(u_1) + \mathbf{F}_2^{(-1)}(u_2) \\ &= \mathbf{G}_\wedge^{(-1)}(u)\end{aligned} \tag{12}$$

We have reported in the figure 1 the dependency bounds when $\mu_1 = 1$, $\mu_2 = 2$, $\sigma_1 = 2$ and $\sigma_2 = 1$ in the case $\mathbf{C}_- = \mathbf{C}^-$. Moreover, we have represented the distribution $\mathbf{G}$ when the dependence structure is respectively

$\mathbf{C}^{-}$, $\mathbf{C}^{\perp}$ and $\mathbf{C}^{+}$. The figure 2 shows the impact of a tighter lower copula bound ($\mathbf{C}_{-} = \mathbf{C}^{\perp}$). Finally, we consider the influence of the standard error ($\sigma_1 = \sigma_2 = \sigma$) in the figure 3 — the solid lines correspond to the case $\mathbf{C}_{-} = \mathbf{C}^{-}$ whereas the dashed line correspond to the case $\mathbf{C}_{-} = \mathbf{C}^{\perp}$.

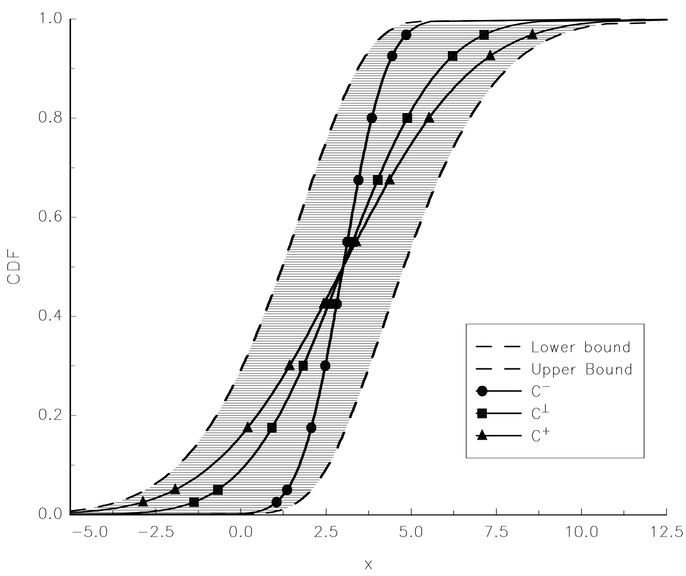

Figure 1: Dependency bounds for two gaussian margins

We suppose now that the margins are student t_ν with ν degrees of freedom. Note that the limit distribution when ν is $+\infty$ is the standardized normal distribution. In the figure 4, we have represented the dependency bounds when ν is respectively 2 (dashed lines) and $+\infty$ (solid lines). We remark clearly the impact of the fat tails on $\mathbf{G}_\vee$ and $\mathbf{G}_\wedge$.

2.5 Financial applications

2.5.1 A first illustration

Embrechts, McNeil and Straumann [1999] apply the previous results of dependency bounds on the aggregation problem of value-at-risk. They highlight the following fallacy:

> **The worst case VaR for a linear portfolio X_1+X_2 occurs when the Pearson correlation $\rho(X_1, X_2)$ is maximal, i.e. X_1 and X_2 are comonotonic.**

The value-at-risk is a measure of economic capital which corresponds to the α quantile of the (potential) loss distribution $\mathbf{F}$:

$$\mathrm{VaR}_\alpha(X) = \mathbf{F}^{-1}(\alpha) \tag{13}$$

Using the duality theorem, we could then show that the corresponding dependency bounds are

$$\mathbf{G}_\wedge^{(-1)}(\alpha) \leq \mathrm{VaR}_\alpha(X_1 + X_2) \leq \mathbf{G}_\vee^{(-1)}(\alpha) \tag{14}$$

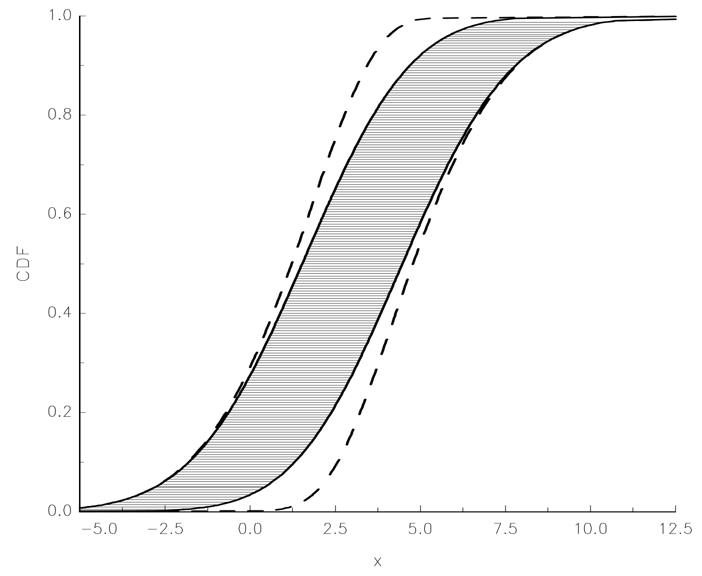

Figure 2: Impact of a positive quadrant dependence ($\mathbf{C} \succ \mathbf{C}^{\perp}$) on the dependency bounds

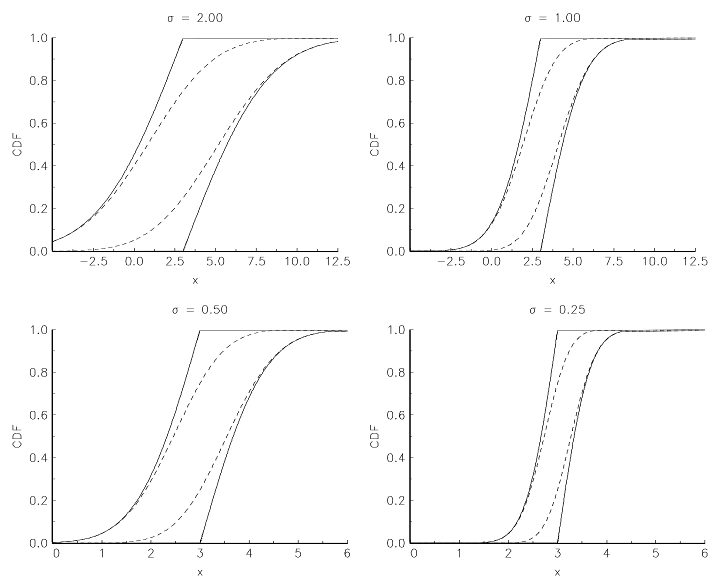

Figure 3: Impact of σ on the dependency bounds

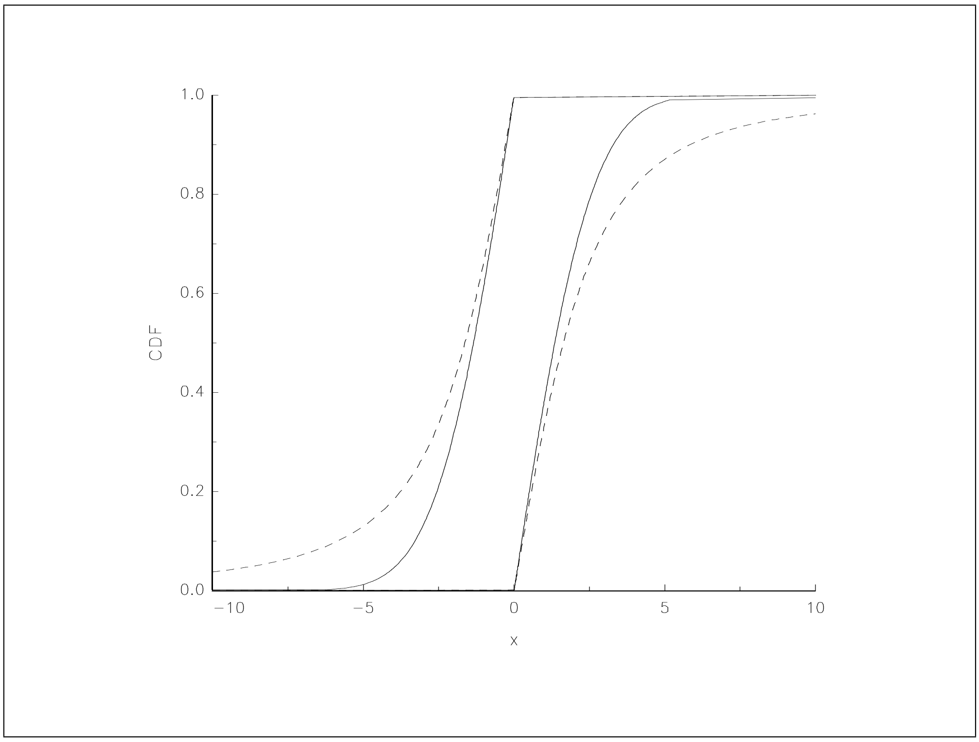

Figure 4: Dependency bounds for two student margins

where $\mathbf{G}_{\wedge}^{(-1)}$ and $\mathbf{G}_{\vee}^{(-1)}$ are given by equations (8) and (9). As noted by EMBRECHTS, MCNEIL and STRAUMANN [1999], these bounds are best-possible.

EMBRECHTS, MCNEIL and STRAUMANN [1999] consider the example where the random variables of the losses are two Gamma distribution with parameter equal to 3. They compare the dependency bounds of the aggregated *VaR* with *VaR* computed under the hypothesis of comonotonicity — $\mathbf{F}^{-1}(\alpha) = \mathbf{F}_1^{-1}(\alpha) + \mathbf{F}_2^{-1}(\alpha)$ — and independence — $x = \mathbf{F}^{-1}(\alpha)$ is then computed using a numerical procedure to solve the non-linear equation $\sigma_{\mathbf{C}^{\perp},+}(\mathbf{F}_1, \mathbf{F}_2)(x) = \alpha$. We have reproduced their results in the figure 5. We remark that the maximal value of $\mathrm{VaR}_{\alpha}(X_1 + X_2)$, i.e. the value taken by $\mathbf{G}_{\vee}^{(-1)}(\alpha)$, could be considerably larger than the corresponding value when $\mathbf{C} = \mathbf{C}^{+}$. EMBRECHTS, MCNEIL and STRAUMANN [1999] conclude that "this is not surprising since we know that *VaR* is not a subadditive risk measure (ARTZNER, DELBAEN, EBER and HEATH [1999]) and there are situations where $\mathrm{VaR}_{\alpha}(X_1 + X_2) > \mathrm{VaR}_{\alpha}(X_1) + \mathrm{VaR}_{\alpha}(X_2)$".

2.5.2 The diversification effect

Let $\varrho(X)$ be a risk measure associated to the "final net worth of a position" X. ARTZNER, DELBAEN, EBER and HEATH [1999] show that if ϱ is a coherent measure of risk, then we verify that the subadditivity axiom

$$\varrho(X_1 + X_2) \leq \varrho(X_1) + \varrho(X_2) \tag{15}$$

We generally define the diversification effect as follows

$$D = \frac{\varrho(X_1) + \varrho(X_2) - \varrho(X_1 + X_2)}{\varrho(X_1) + \varrho(X_2)} \tag{16}$$

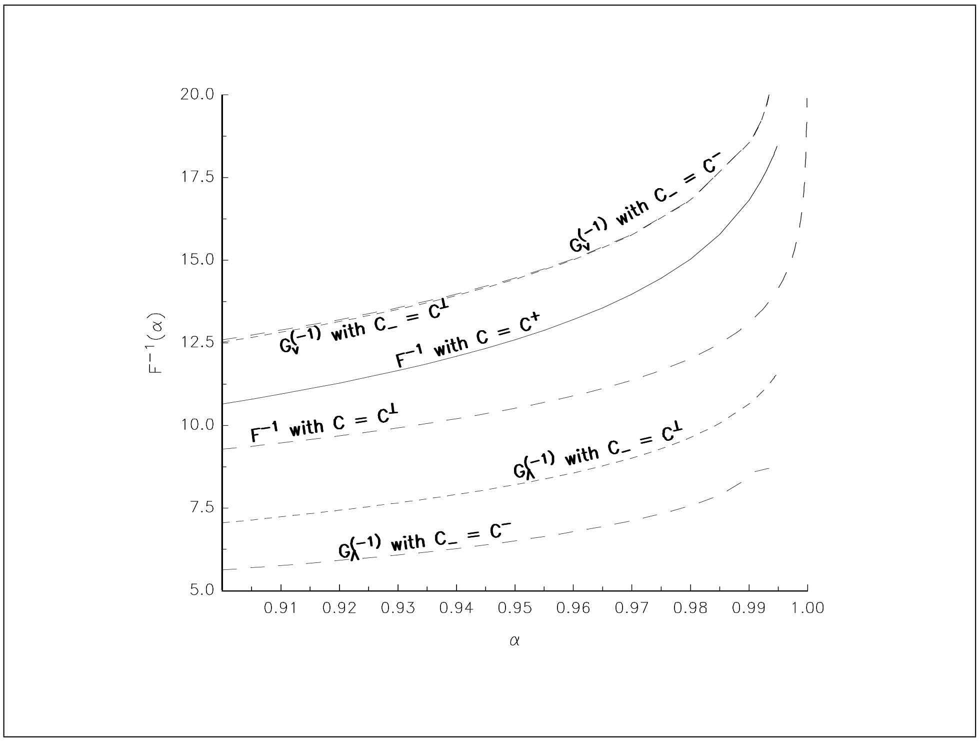

Figure 5: Dependency bounds for Value-at-Risk with Gamma marginals

In an economic capital approach (ref. [1]), it comes that $D \in [0,1]$. In the case of the value-at-risk, it seems natural to transpose the definition, and we have

$$D = \frac{\mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2) - \mathrm{VaR}_\alpha(X_1 + X_2)}{\mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)} \tag{17}$$

However, we have seen that there are situations where $\mathrm{VaR}_\alpha(X_1 + X_2) > \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$. A more appropriate definition is then

$$\bar{D} = \frac{\mathbf{G}_\vee^{(-1)}(\alpha) - \mathrm{VaR}_\alpha(X_1 + X_2)}{\mathbf{G}_\vee^{(-1)}(\alpha)} \tag{18}$$

It comes that

$$\begin{aligned}
\bar{D} &= \frac{\mathbf{G}_\vee^{(-1)}(\alpha) - \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)}{\mathbf{G}_\vee^{(-1)}(\alpha)} + \frac{\mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)}{\mathbf{G}_\vee^{(-1)}(\alpha)} \times D \\
&= \chi\left(\mathbf{C}_\vee^{(\alpha)}, \mathbf{C}^+; \alpha\right) + \left[1 - \chi\left(\mathbf{C}_\vee^{(\alpha)}, \mathbf{C}^+; \alpha\right)\right] \times D
\end{aligned} \tag{19}$$

with

$$\chi\left(\mathbf{C}_\vee^{(\alpha)}, \mathbf{C}^+; \alpha\right) = \frac{\mathbf{G}_\vee^{(-1)}(\alpha) - \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)}{\mathbf{G}_\vee^{(-1)}(\alpha)}$$

Embrechts, McNeil and Straumann [1999] interpret $\chi\left(\mathbf{C}_\vee^{(\alpha)}, \mathbf{C}^+; \alpha\right)$ as "*the amount by which VaR fails to be subadditive*".

We consider the example of two student margins t_{ν_1} and t_{ν_2}. We suppose that we aggregate the risks using a Normal copula with parameter ρ (ρ takes respectively the value -0.5, 0, 0.5 et 0.75). For each aggregation,

		$\rho = -0.5$		$\rho = 0$		$\rho = 0.5$		$\rho = 0.75$	
		$\hat{\rho} = \checkmark$		$\hat{\rho} = \checkmark$		$\hat{\rho} = \checkmark$		$\hat{\rho} = \checkmark$	
α	χ	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$
0.80	0.44	0.20	0.55	0.057	0.47	0.010	0.44	0.010	0.44
0.85	0.39	0.24	0.54	0.098	0.45	0.026	0.41	0.010	0.39
0.90	0.35	0.28	0.54	0.12	0.43	0.045	0.38	0.019	0.37
0.95	0.32	0.32	0.54	0.16	0.43	0.058	0.36	0.024	0.34
0.98	0.30	0.32	0.53	0.19	0.43	0.051	0.34	0.082	0.36

Table 1: Results with $\nu_1 = 2$ and $\nu_2 = 2$

		$\rho = -0.5$		$\rho = 0$		$\rho = 0.5$		$\rho = 0.75$	
		$\hat{\rho} = -0.18$		$\hat{\rho} = 0$		$\hat{\rho} = 0.18$		$\hat{\rho} = 0.28$	
α	χ	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$
0.80	0.39	0.27	0.56	0.11	0.46	0.062	0.43	0.010	0.40
0.85	0.34	0.30	0.54	0.18	0.46	0.098	0.40	0.057	0.38
0.90	0.29	0.34	0.53	0.21	0.44	0.085	0.35	0.054	0.33
0.95	0.25	0.38	0.53	0.24	0.43	0.11	0.33	0.064	0.29
0.98	0.22	0.40	0.53	0.27	0.43	0.14	0.32	0.074	0.28

Table 2: Results with $\nu_1 = 3$ and $\nu_2 = 4$

		$\rho = -0.5$		$\rho = 0$		$\rho = 0.5$		$\rho = 0.75$	
		$\hat{\rho} = -0.39$		$\hat{\rho} = 0$		$\hat{\rho} = 0.39$		$\hat{\rho} = 0.59$	
α	χ	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$
0.80	0.36	0.26	0.53	0.21	0.49	0.092	0.42	0.035	0.38
0.85	0.30	0.32	0.52	0.22	0.46	0.13	0.39	0.042	0.33
0.90	0.24	0.38	0.53	0.24	0.42	0.13	0.34	0.056	0.29
0.95	0.19	0.40	0.51	0.26	0.40	0.12	0.29	0.067	0.24
0.98	0.15	0.43	0.52	0.29	0.39	0.14	0.27	0.075	0.21

Table 3: Results with $\nu_1 = 9$ and $\nu_2 = 10$

		$\rho = -0.5$		$\rho = 0$		$\rho = 0.5$		$\rho = 0.75$	
		$\hat{\rho} = -0.48$		$\hat{\rho} = 0$		$\hat{\rho} = 0.48$		$\hat{\rho} = 0.72$	
α	χ	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$	D	$\bar{D}$
0.80	0.36	0.31	0.56	0.19	0.48	0.080	0.41	0.080	0.41
0.85	0.29	0.35	0.54	0.21	0.44	0.12	0.38	0.073	0.35
0.90	0.24	0.37	0.52	0.26	0.43	0.11	0.32	0.073	0.29
0.95	0.18	0.41	0.51	0.26	0.40	0.12	0.28	0.067	0.23
0.98	0.14	0.45	0.53	0.30	0.39	0.14	0.26	0.077	0.21

Table 4: Results with $\nu_1 = 50$ and $\nu_2 = 50$

we compute D, $\bar{D}$ and $\chi\left(\mathbf{C}_{\vee}^{(\alpha)}, \mathbf{C}^{+}; \alpha\right)$ for different confidence levels. The results are given in the tables 1–4. Moreover, we have reported also the Pearson correlation $\hat{\rho}$. With this example, we could do the following remarks:

1. the diversification effect does not depend only on the correlation, but also on marginal distributions;
2. the dependence structure (i.e. the copula used to perform the aggregation) plays a more important role than the correlation for determining the diversification effect;
3. a lower correlation does not imply systematically a greater diversification effect.

2.5.3 The 'square root' rule

In the finance industry, the aggregation is often done using the 'square root' rule

$$\mathrm{VaR}_{\alpha}\left(X_1 + X_2\right) = \sqrt{\left[\mathrm{VaR}_{\alpha}\left(X_1\right)\right]^2 + 2\rho\left(X_1, X_2\right)\mathrm{VaR}_{\alpha}\left(X_1\right)\mathrm{VaR}_{\alpha}\left(X_2\right) + \left[\mathrm{VaR}_{\alpha}\left(X_2\right)\right]^2} \tag{20}$$

with $\rho\left(X_1, X_2\right)$ the Pearson correlation between the two random variables $X_1 + X_2$. When $\rho\left(X_1, X_2\right) = 1$, we obtain of course $\mathrm{VaR}_{\alpha}\left(X_1 + X_2\right) = \mathrm{VaR}_{\alpha}\left(X_1\right) + \mathrm{VaR}_{\alpha}\left(X_2\right)$. Sometimes, risk managers have no ideas about the value of $\rho\left(X_1, X_2\right)$ and suppose $\rho\left(X_1, X_2\right) = 0$. This could be justified in operational risk measurement. For market risk measurement, the justification is more difficult in the case of the risk aggregation by desks. However, this rule is sometimes used when the aggregation is performed by markets (for example, bond, equity and commodities markets) because of the segmentation assumption.

We suppose that X_1 and X_2 are two student random variables with 5 and 7 degrees of freedom. We consider different dependence structures to perform the aggregation and we compare the obtained values with the values computed with the 'square root' rule. We have also reported the upper dependency bound. With this example, we remark that the relevance of this rule depends on the confidence level. We have the impression that it works better when α is closed to one and $\rho\left(X_1, X_2\right)$ is closed to 0.

2.5.4 The VaR aggregation in practice

In practice, risk managers use three methods to compute value-at-risk: analytical, historical and Monte Carlo. In the first method, we know the distribution of the losses X, whereas in the two others, we know some realizations of the random variable. Nevertheless, aggregation (or $\sigma-$*convolution*) of different types of *VaR* is not a numerical problem (see Williamson [1989] for a survey of the different algorithms: Skinner/Ackroyd method, spline based methods, Laguerre and Mellin transforms, etc.). In particular, Williamson [1989] suggests to use the condensation procedure (algorithm 3.4.28). For the dependency bounds, Williamson presents also an algorithm based on the *uniform quantisation method*.

We use the LME example of Bouyé, Durrleman, Nikeghbali, Riboulet and Roncalli [2000]. We consider the spot prices of the commodities Aluminium Alloy (AL), Copper (CU), Nickel (NI), Lead (PB) and the 15 months forward prices of Aluminium Alloy (AL-15), dating back to January 1988. We constitute two portfolios with the compositions

	AL	AL-15	CU	NI	PB
P_1	5	3			
P_2			5	2	-3

Computations give us the following values for the value-at-risk with a 99% confidence level:

	Analytical *VaR*	Historical *VaR*
P_1	363.05	445.74
P_2	1026.03	1274.64

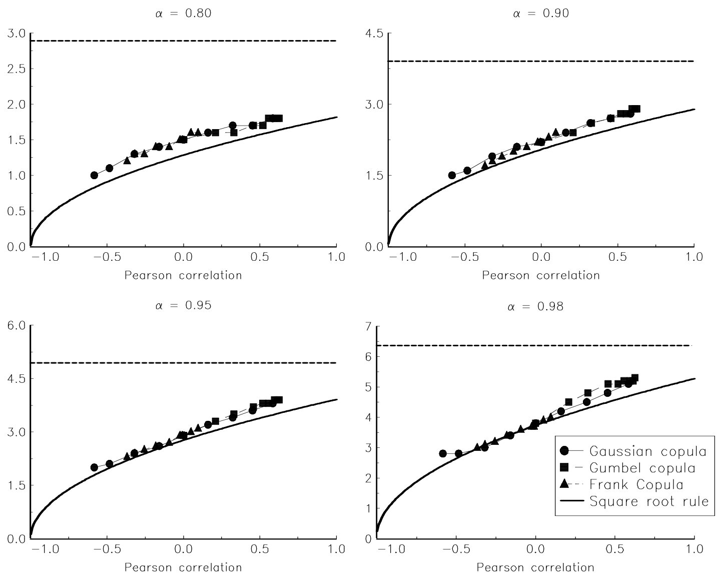

Figure 6: Illustration of the 'square root' rule

We remark that the method has a great influence on the dependency bound for the value-at-risk (see the figure 7). For illustration, here are the values of $\mathbf{G}_{\vee}^{(-1)}(\alpha)$ for α equal to 99%:

		P_1 Analytical *VaR*	P_1 Historical *VaR*
P_2	Analytical *VaR*	1507.85	1680.77
P_2	Historical *VaR*	1930.70	2103.67

2.5.5 Bivariate WCS methodology

Dependency bounds could also be used to propose a multivariate version of the "*worst-case scenario*" (or WCS) defined by BOUDOUKH, RICHARDSON et WHITELAW [1995]:

> *WCS asks the following question: what is the worst that can happen to the value of the firm's trading portfolio over a given period (eg, 20 trading days)?*

As explained by BAHAR, GOLD, KITTO et POLIZU [1997], the use of the WCS method could interest risk managers:

> *...VaR has been criticised at a more fundamental level — that it asks the wrong question... VaR requires a probability that loss shocks, over a short horizon, will not exceed a certain figure. This is often interpreted as expressing the number of times a loss shock in excess of the subjective threshold would occur over a long time horizon. But managers are more likely be concerned with the probability of realising a specific loss shock, rather than the probability of landing in a region of large loss shocks.*

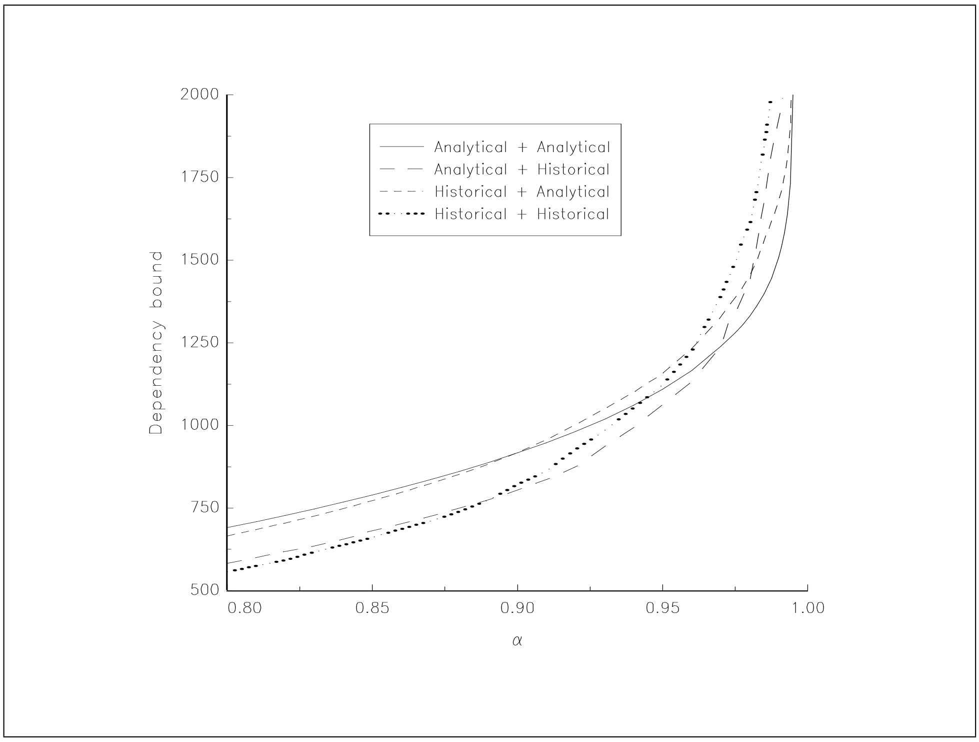

Figure 7: Plot of $\mathbf{G}_{\vee}^{(-1)}(\alpha)$

We would define the distribution $\mathbf{F}_{-}$ of $\chi_N^{-} = \min(X_1, \ldots, X_n, \ldots, X_N)$. If we suppose that the random variables X_n are independent with the same distribution $\mathbf{F}$, we have $\mathbf{F}_{-}(x) = 1 - [1 - \mathbf{F}(x)]^N$. Let f_{-} and f be the corresponding densities. Bahar, Gold, Kitto et Polizu [1997] show that $f_{-}(x) = N \times [1 - \mathbf{F}(x)]^{N-1} f(x)$ and $\mathbf{F}_{-}^{-1}(\alpha) = \mathbf{F}^{-1}\left(1 - (1-\alpha)^{\frac{1}{N}}\right)$. $\mathbf{F}_{-}^{-1}(\alpha)$ is then *the quantile of the worst*. Without knowledge on the dependency function, computing a bivariate WCS could be done using the previous dependency bounds.

Note that this analysis could also be extended in the case where the random variables are not independent. Let $\mathbf{C}$ be the copula of the random vector $(X_1, \ldots, X_n, \ldots, X_N)$ and $\mathbf{F}_n$ be the distribution of the random variable X_n. We have

$$\begin{aligned}
\mathbf{F}_{-}(x) &= \Pr\left(\chi_N^{-} \leq x\right) \\
&= 1 - \Pr(X_1 > x, \ldots, X_n > x, \ldots, X_N > x) \\
&= 1 - \bar{\mathbf{C}}(\mathbf{F}_1(x), \ldots, \mathbf{F}_n(x), \ldots, \mathbf{F}_N(x)) \qquad (21)
\end{aligned}$$

with $\bar{\mathbf{C}}$ the joint survival function (see [7] for an explicit form of $\bar{\mathbf{C}}$). In the case where $\mathbf{C} = \mathbf{C}^{\perp}$ and $\mathbf{F}_n = \mathbf{F}$, we obtain the previous expression. In the general case, *the quantile of the worst* is obtained thanks to a numerical root finding procedure.

2.5.6 Correlation stress-testing programs

We mention here a last example of financial applications. A simple definition of stress-testing is the following

> *The art of stress testing should give the institution a deeper understanding of the specific portfolios that could be put in jeopardy given a certain situation. The question then would be: "Would*

this be enough to bring down the firm?" That way, each institution can know exactly what scenario it is they do not want to engage in (Michael Ong, ABN Amro, *in* Dunbar and Irving [1998]).

In a quantitative stress-testing program, one of the big difficulty is to stress the correlation and to measure its impact. The effect of the worst situation could then be computed with the dependency bounds.

3 The Kantorovich distance based method

3.1 The Dall'aglio problem

Barrio, Giné and Matrán [1999] define the Kantorovich distance or L_1-Wasserstein distance between two probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$ as

$$\mathrm{d}_1(\mathbb{P}_1,\mathbb{P}_2) := \inf\left\{\iint_{\mathbb{R}^2} |x_1-x_2|\,\mathrm{d}\mu(x_1,x_2) : \mu \in \mathcal{P}\left(\mathbb{R}^2\right) \text{ with marginals } \mathbf{F}_1 \text{ and } \mathbf{F}_2\right\} \tag{22}$$

We can write it in another way:

$$\mathrm{d}_1(\mathbb{P}_1,\mathbb{P}_2) = \inf_{\mathbf{F}\in\mathcal{F}(\mathbf{F}_1,\mathbf{F}_2)} \mathbb{E}\,|X_1-X_2| \tag{23}$$

where $\mathbf{F}$ is taken on the set of all probability with marginals $\mathbf{F}_1$ and $\mathbf{F}_2$ — $\mathbf{F}$ belongs to the Fréchet class $\mathcal{F}(\mathbf{F}_1,\mathbf{F}_2)$. Shorack and Wellner [1986] showed that if $\mathbf{F}_1$ and $\mathbf{F}_2$ are the distribution functions associated to $\mathbb{P}_1$ and $\mathbb{P}_2$, then we have

$$\mathrm{d}_1(\mathbb{P}_1,\mathbb{P}_2) = \int_{-\infty}^{+\infty} |\mathbf{F}_2(x)-\mathbf{F}_1(x)|\,\mathrm{d}x \tag{24}$$

or

$$\mathrm{d}_1(\mathbb{P}_1,\mathbb{P}_2) = \int_0^1 \left|\mathbf{F}_2^{-1}(x)-\mathbf{F}_1^{-1}(x)\right|\,\mathrm{d}x \tag{25}$$

We can remark that when $\mathbb{P}_1$ is equal to $\mathbb{P}_2$, the Frechet upper bound copula $\mathbf{C}^+$ is the solution of the problem. More generally, Dall'aglio [1991] showed that there is **a set of minimizing joint distribution functions**:

- the largest is $\mathbf{F}_+(x_1,x_2) = \mathbf{C}^+(\mathbf{F}_1(x_1),\mathbf{F}_2(x_2))$;
- the smallest is

$$\begin{aligned}\mathbf{F}_-(x_1,x_2) \quad = \quad & \mathbf{1}_{[x_1\le x_2]}\left[\mathbf{F}_1(x_1)-\left(\inf_{x_1\le x\le x_2}[\mathbf{F}_1(x)-\mathbf{F}_2(x)]\right)^+\right]+ \\ & \mathbf{1}_{[x_1>x_2]}\left[\mathbf{F}_2(x_2)-\left(\inf_{x_2\le x\le x_1}[\mathbf{F}_2(x)-\mathbf{F}_1(x)]\right)^+\right]\end{aligned} \tag{26}$$

- any convex combination of $\mathbf{F}_+$ and $\mathbf{F}_-$ is still a solution of (23).

3.2 Financial applications

This expression (23) of the Kantorovich distance is very helpful when we consider that the random variates X_1 and X_2 represent losses for a two dimensional portofolio. We see that the joint probability measures which minimize (23) will be those with "maximum risks". They can be considered as the more risky distributions. In this case, the upper Fréchet bound is a solution, but it is **not unique**. This is in contradiction with the belief in finance, that the maximal risk corresponds to the case where the random variables are comonotonic.

Let us consider two examples. In the first example, we assume that the margins are gaussian ($X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim \mathcal{N}(0,2)$). In the second one, we use two Beta distributions ($X_1 \sim \beta(2,4)$ and $X_2 \sim \beta(0.5,0.5)$). We have reported the values taken by $\mathbf{F}_+(x_1,x_2)$ and $\mathbf{F}_-(x_1,x_2)$ in the figure 8. We remark that the two bounds are very different.

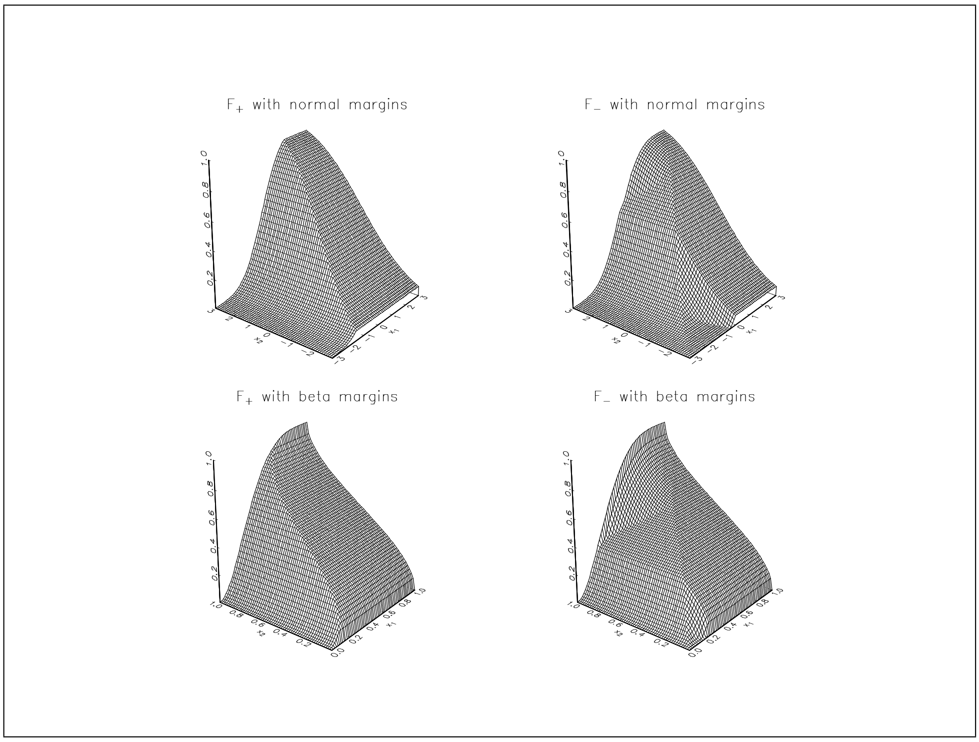

Figure 8: Illustration of the Dall'aglio problem

4 Conclusion

In this paper, we apply the dependency bounds to the value-at-risk problem, and show that the more risky distribution is not always the upper Fréchet copula.

References

[1] Amendment to the capital accord to incorporate market risks, *Basle Committee on Banking Supervision*, January 1996, N° **24**

[2] Artzner, A., F. Delbaen, J-M. Eber and D. Heath [1999], Coherent measures of risk, *Mathematical Finance*, **9**, 203-228

[3] Bahar, R., M. Gold, T. Kitto and C. Polizu [1997], Making the best of the worst, *Risk Magazine*, **10**, August

[4] Barrio E. (del), E. Giné and C. Matrán [1999], Central limit theorems for the Wasserstein distance between the empirical and the true distributions, *Annals of Probability*, **27**, 1009-1071

[5] Barrio E. (del), J.A. Cuesta Albertos, C. Matrán and J. Rodríguez Rodríguez [1999], Tests of goodness of fit based on the L_2-Wasserstein distance, *Annals of Statistics*, **27**, 1230-1239

[6] Boudoukh, J., M. Richardson et R. Whitelaw [1995], Expect the worst, *Risk Magazine*, **8**, September

[7] 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*

[8] Dall'Aglio, G. [1991], Fréchet classes: the beginnings, in G. Dall'Aglio, S. Kotz and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals (Beyond the Copulas)*, Kluwer Academic Publishers, Dordrecht

[9] Dall'Aglio, G., S. Kotz and G. Salinetti [1991], Advances in Probability Distributions with Given Marginals (Beyond the Copulas), Kluwer Academic Publishers, Dordrecht

[10] Dunbar, N. and R. Irving [1998], This is the way the world ends, *Risk Magazine*, **11**, December

[11] 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*

[12] Frank, M.J. [1991], Convolutions for dependent random variables, in G. Dall'Aglio, S. Kotz and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals (Beyond the Copulas)*, Kluwer Academic Publishers, Dordrecht

[13] Frank, M.J., R.B. Nelsen, and B. Schweizer [1987], Best-possible bounds for the distribution of a sum — a problem of Kolmogorov, *Probability Theory and Related Fields*, **74**, 199-211

[14] Frank, M.J. and B. Schweizer [1979], On the duality of generalized infimal and supremal convolutions, *Rendiconti di Matematica*, **12**, 1-23

[15] Hillali, Y. [1998], Analyse et Modélisation des Données Probabilistes : Capacités et Lois Multidimensionnelles, PhD Thesis, University of Paris IX-Dauphine

[16] Li, H., M. Scarsini and M. Shaked [1996], Bounds for the distribution of a multivariate sum, in L. Rushendorf, B. Schweizer and M.D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics*, Institute of Mathematical Statistics, Hayward, CA

[17] Makarov, G. [1981], Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed, *Theory of Probability and its Applications*, **26**, 803-806

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

[19] Nelsen, R.B. and B. Schweizer [1991], Bounds for distribution functions of sums of squares and radial errors, *International Journal of Mathematics and Mathematical Sciences*, **14**, 561-570

[20] Rushendorf, L., B. Schweizer and M.D. Taylor [1993], Distributions with Fixed Marginals and Related Topics, Institute of Mathematical Statistics, Hayward, CA

[21] Schweizer, B. [1991], Thirty years of copulas, in G. Dall'Aglio, S. Kotz and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals (Beyond the Copulas)*, Kluwer Academic Publishers, Dordrecht

[22] Schweizer, B. and A. Sklar [1983], Probabilistic Metric Spaces, Elsevier North-Holland, New York

[23] Shorack, G. and J.A. Wellner [1986], Empirical Processes with Applications to Statistics, John Wiley & Sons, New York

[24] Williamson, R.C [1989], Probabilistic Arithmetic, PhD Thesis, University of Queensland

[25] Williamson, R.C. and T. Downs [1990], Probabilistic arithmetic I. — Numerical methods for calculating convolutions and dependency bounds, *International Journal of Approximate Reasoning*, **4**, 89-158