

Erdely, Arturo

Article

Value at risk and the diversification dogma

Revista de Métodos Cuantitativos para la Economía y la Empresa

Provided in Cooperation with:

Universidad Pablo de Olavide, Sevilla

Suggested Citation: Erdely, Arturo (2017) : Value at risk and the diversification dogma, Revista de Métodos Cuantitativos para la Economía y la Empresa, ISSN 1886-516X, Universidad Pablo de Olavide, Sevilla, Vol. 24, pp. 209-219

This Version is available at:

<https://hdl.handle.net/10419/195388>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>

Value at Risk and the Diversification Dogma

ERDELY, ARTURO

Facultad de Estudios Superiores Acatlán

Universidad Nacional Autónoma de México (México)

E-mail: arturo.erdely@comunidad.unam.mx

ABSTRACT

The so-called risk diversification principle is analyzed, showing that its convenience depends on individual characteristics of the risks involved and the dependence relationship among them.

Keywords: value at risk; loss aggregation; comonotonicity; diversification.

JEL classification: G11; G22.

MSC2010: 91G10; 91G40.

Valor en riesgo y el dogma de la diversificación

RESUMEN

Se analiza el principio de diversificación de riesgos y se demuestra que no siempre resulta mejor que no diversificar, pues esto depende de características individuales de los riesgos involucrados, así como de la relación de dependencia entre los mismos.

Palabras claves: valor en riesgo; agregación de pérdidas; comonotonicidad; diversificación.

Clasificación JEL: G11; G22.

MSC2010: 91G10; 91G40.



1 Introduction

A popular proverb states *don't put all your eggs in one basket* and it is implicitly based on a *principle* (let's call it that way momentarily) of *risk diversification* which could have the following “justification”: Suppose it is needed to take $2n$ eggs from point A to point B, walking distance, and that there are only two alternatives available, either one person carrying all the eggs in one basket, or two people with n eggs each in separate (and independent) baskets. The proverb suggests that there is a higher risk with the single-person alternative since if he/she happens to stumble and fall we would have a total loss, while with the second alternative only half of the eggs would be lost, and in a worst case scenario (with lower probability) where the two people fall the loss would be the same as in the first alternative, anyway.

Let X be a random variable which counts how many eggs are lost under the first alternative (one basket), and let Y account for the same but for the second alternative (two baskets). Let $0 < \theta < 1$ be the probability of falling and breaking the eggs in a basket while walking from point A to B. Then X and Y are discrete random variables such that $\mathbb{P}(X \in \{0, 2n\}) = 1$ and $\mathbb{P}(Y \in \{0, n, 2n\}) = 1$, with point probabilities $\mathbb{P}(X = 2n) = \theta$, $\mathbb{P}(X = 0) = 1 - \theta$, $\mathbb{P}(Y = 2n) = \theta^2$, $\mathbb{P}(Y = n) = 2\theta(1 - \theta)$, and $\mathbb{P}(Y = 0) = (1 - \theta)^2$. Certainly the probability of facing the maximum loss of $2n$ eggs has a higher probability under the first alternative, but it is also true that the *no loss* probability is also higher under such alternative. Moreover:

$$\mathbb{P}(Y > 0) = \theta^2 + 2\theta(1 - \theta) = \theta(2 - \theta) > \theta = \mathbb{P}(X > 0),$$

which means that there is a higher probability of suffering a (partial or total) loss under the second alternative. Therefore... does it mean that it is better to put all the eggs in one basket? If a single trip is going to take place, the answer would be yes, but if the same trip is going to be repeated a large number of times we should analyze the *long run* average loss, which would be $\mathbb{E}(X) = 2n\theta$ for the first alternative, and $\mathbb{E}(Y) = 2n\theta^2 + 2n\theta(1 - \theta) = 2n\theta$ for the second alternative; that is, in the long run there is no difference between the two alternatives.

Is it never more convenient to diversify in two baskets? If the probability of stumbling and falling with $2n$ eggs is the same as with half of them (which might be true up to certain value of n) then the proverb is certainly wrong, but maybe for a sufficiently large value of n we should consider different probabilities of falling and breaking the eggs, say θ_1 for the first alternative and θ_2 for the second one, with $\theta_1 > \theta_2$. This last condition leads to $\mathbb{E}(X) > \mathbb{E}(Y)$ and in such case it is more convenient to diversify if a large number of trips are going to be made. But for a single trip decision the condition $\theta_1 > \theta_2$ is not enough to prefer diversification unless $\theta_2(2 - \theta_2) < \theta_1$, since $\theta_2 < \theta_2(2 - \theta_2)$.

The main purpose of the present work is to show that the common belief that risk diversification is **always** better, is more a **dogma**¹ rather than a general principle that has been proved, and that the correct view is to state that risk diversification may be better, as good as, or worse than lack thereof, depending on the risks involved and the dependence relationship among them.

2 Risk measures

Let X be a continuous random variable, with strictly increasing distribution function F_X , that represents an economic loss generated by certain events covered by insurance or related to investments. Without loss of generality, we consider amounts of constant value over time (inflation indexed, for example). As a point estimation for a potential loss, we may use the mean or the median. In the present work, the median is preferred since it always exists for continuous random variables and it is robust, in contrast with the mean that may not exist or could be numerically unstable under heavy-tailed probability distributions. Using the quantile function (inverse of F_X), we calculate the median as $\mathbb{M}(X) = F_X^{-1}(\frac{1}{2})$ since $\mathbb{P}(X \leq \mathbb{M}(X)) = \frac{1}{2}$.

¹A system of principles or tenets; doctrine. A specific principle of a doctrine put forth, such as by a church. *Source:* WordReference Random House Learner's Dictionary of American English © 2016.

Definition 2.1. The *excess of loss* for a continuous loss random variable X is the random variable:

$$L := X - \mathbb{M}(X).$$

As suggested by McNeil *et al.* (2015) one way to interpret a *risk measure* is as the required additional risk capital $\varrho(L)$ to cover a loss in excess of what was originally estimated. In the specialized literature on this subject, there are many properties for risk measures that are considered as “desirable” or “reasonable”, though some concerns have been raised for some of them.

Definition 2.2. A risk measure ϱ is *monotone* if for any excess of loss random variables L_1 and L_2 such that $\mathbb{P}(L_1 \leq L_2) = 1$, we have that $\varrho(L_1) \leq \varrho(L_2)$.

McNeil *et al.* (2015) and several other authors consider monotonicity as a clearly desirable property since financial positions that involve higher risks under any circumstance should be covered by more risk capital. Positions such that $\varrho(L) \leq 0$ do not require additional capital.

Definition 2.3. A risk measure ϱ is *translation invariant* if for any excess of loss random variable L and any constant c we have that $\varrho(L + c) = \varrho(L) + c$.

This property is also considered as desirable by McNeil *et al.* (2015) and other authors under the following argument: The uncertainty associated to $L' := L + c$ totally depends on L since c is fixed, $\varrho(L)$ is the additional risk capital required to cover an excess of loss under L and therefore it would be enough to add the fixed amount c in order to cover for L' .

Definition 2.4. A risk measure ϱ is *subadditive* if for any excess of loss random variables L_1 and L_2 we have that $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$.

This property cannot be considered as generally acceptable since there is some debate around it. One argument in favor is that diversification **always** reduces risk, which is more a **dogma** rather than something proved to be true under all circumstances. We may counterargue that for some risks there could be some sort of pernicious interaction that generates additional risk to the individual ones; so it may be also argued that it is better for a risk measure not to be subadditive, so that whenever it happens that $\varrho(L_1 + L_2) > \varrho(L_1) + \varrho(L_2)$ then it becomes clear that diversification is not convenient in such case.

Definition 2.5. A risk measure ϱ is *positively homogeneous* if for any excess of loss random variable L and any constant $\lambda > 0$ we have that $\varrho(\lambda L) = \lambda \varrho(L)$.

With regard to this property McNeil *et al.* (2015) and other authors mention that, in case that subadditivity has been accepted as reasonable, then for any positive integer n , it should be accepted that

$$\varrho(nL) = \varrho(L + \dots + L) \leq n\varrho(L) \quad (1)$$

and since there is no diversification “benefit” (because just a single risk source is involved), then the highest value would be attained in (1), that is equality. The same authors acknowledge there is some criticism about this property since for sufficiently large values of λ we should have $\varrho(\lambda L) > \lambda \varrho(L)$ to penalize for a high concentration of risk in a single source of it.

Definition 2.6. ϱ is a *coherent risk measure* if it satisfies Definitions 2.2 to 2.5.

The adjective “coherent” in this definition is somehow overbearing since it implicitly suggests that any risk measure that does not satisfy this definition would be incoherent despite the fact that there is some debate and concerns about two of the four properties to be required. There are other additional properties that have been proposed in some contexts, see McNeil *et al.* (2015) or Denuit *et al.* (2005), but for the purpose of this article the above mentioned ones are enough.

3 Value at risk

As suggested by McNeil *et al.* (2015), we may interpret $\varrho(L)$ as the additional risk capital to cover for a potential excess of loss with L , but in practice such interpretation could be easily unachievable. Consider, for example, an insurance portfolio with certain face amounts for each issued policy. The only way to guarantee that the insurance company has enough resources to pay the claims under all possible scenarios would require the total reserve to be equal to the sum of all the face amounts in such portfolio.

In practice, specially under the *Basel Accords* and *Solvency II* frameworks, what is calculated is the amount of risk capital that has an *acceptable* high probability (but strictly less than 1) of covering an excess of loss that might face an insurance or financial institution. Who determines how much is “acceptable”? Typically the regulatory authority, but each company may decide to use probability levels even higher than the regulatory ones.

Definition 3.1. *Value at Risk* of level $0 < \alpha < 1$ for an excess of loss random variable L is a risk measure defined as

$$\text{VaR}_\alpha(L) := F_L^{-1}(\alpha)$$

where F_L^{-1} is the quantile function of L , that is the inverse of the probability distribution function of L .

In other words, a level α Value at Risk associated to a continuous random variable is the amount that such variable would not exceed with probability α . It should be noticed that the median is a Value at Risk of level $\alpha = \frac{1}{2}$.

Proposition 3.1. *VaR is a monotone, translation invariant, and positively homogeneous risk measure.*

Proof:

- a) Let X and Y be random variables such that $\mathbb{P}(X \leq Y) = 1$. Then for any value $x \in \mathbb{R}$:

$$\mathbb{P}(X \leq x) = \mathbb{P}(X \leq x < Y) + \mathbb{P}(X \leq Y \leq x) \geq \mathbb{P}(\{X \leq Y\} \cap \{Y \leq x\}) = \mathbb{P}(Y \leq x),$$

that is $F_X(x) \geq F_Y(x)$. Let $x_\alpha := \text{VaR}_\alpha(X)$ and $y_\alpha := \text{VaR}_\alpha(Y)$. Then $\alpha = F_X(x_\alpha) \geq F_Y(x_\alpha)$ and since $\alpha = F_Y(y_\alpha)$ and distribution functions are non decreasing, necessarily $x_\alpha \leq y_\alpha$ and therefore $\text{VaR}_\alpha(X) \leq \text{VaR}_\alpha(Y)$.

- b) Let X be a continuous random variable with strictly increasing distribution function F_X and let $c \in \mathbb{R}$ be any given constant. Define the random variable $Y := X + c$, its probability distribution function is:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X + c \leq y) = \mathbb{P}(X \leq y - c) = F_X(y - c).$$

Let $x_\alpha := \text{VaR}_\alpha(X)$ and $y_\alpha := \text{VaR}_\alpha(Y)$. Then:

$$F_X(x_\alpha) = \alpha = F_Y(y_\alpha) = F_X(y_\alpha - c),$$

and since F_X is strictly increasing then $x_\alpha = y_\alpha - c$ which is equivalent to $\text{VaR}_\alpha(X) + c = \text{VaR}_\alpha(Y) = \text{VaR}_\alpha(X + c)$.

- c) Let X be a continuous random variable with strictly increasing distribution function F_X and let $\lambda > 0$ be a given constant. Define the random variable $Y := \lambda X$, its probability distribution function is:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\lambda X \leq y) = \mathbb{P}(X \leq y/\lambda) = F_X(y/\lambda).$$

Let $x_\alpha := \text{VaR}_\alpha(X)$ and $y_\alpha := \text{VaR}_\alpha(Y)$. Then:

$$F_X(x_\alpha) = \alpha = F_Y(y_\alpha) = F_X(y_\alpha/\lambda),$$

and since F_X is strictly increasing then $x_\alpha = y_\alpha/\lambda$, which is equivalent to $\lambda \text{VaR}_\alpha(X) = \text{VaR}_\alpha(Y) = \text{VaR}_\alpha(\lambda X)$. \square

It should be noticed that VaR is proved to be positively homogeneous without a subadditivity argument as in (1). In fact, VaR is not generally subadditive as it will become clear in a following section, but it will be also argued that this should not be considered as a disadvantage.

Example 3.1. Let X be a *Pareto* continuous random variable with parameters $\beta > 0$ y $\delta > 0$. Its probability density function is given by:

$$f_X(x | \beta, \delta) = \frac{\delta \beta^\delta}{x^{\delta+1}}, \quad x > \beta,$$

and therefore its probability distribution function:

$$F_X(t) = \int_{-\infty}^t f_X(x | \beta, \delta) dx = \delta \beta^\delta \int_{\beta}^t \frac{dx}{x^{\delta+1}} = 1 - \left(\frac{\beta}{t}\right)^\delta, \quad t > \beta.$$

The quantile function of X is the inverse of F_X , that is $F_X^{-1}(u) = \beta(1 - u)^{-1/\delta}$ for $0 < u < 1$, and consequently the median is $\mathbb{M}(X) = \text{VaR}_{1/2}(X) = F_X^{-1}(\frac{1}{2}) = 2^{1/\delta} \beta$. The level $\alpha > \frac{1}{2}$ VaR for the excess of loss $L = X - \mathbb{M}(X)$ is given by:

$$\text{VaR}_\alpha(L) = \text{VaR}_\alpha(X - \mathbb{M}(X)) = \text{VaR}_\alpha(X) - \mathbb{M}(X) = \beta[(1 - \alpha)^{-1/\delta} - 2^{1/\delta}].$$

Thus, with probability α , the excess of loss will not exceed the amount $\text{VaR}_\alpha(L)$. Notice that if $\alpha \rightarrow 1^-$ then $\text{VaR}_\alpha(L) \rightarrow +\infty$, which would require an infinite risk capital, something impossible in practice, and instead a value $\alpha < 1$ sufficiently close to 1 is arbitrarily set by the regulatory authority, for example $\alpha = 0.995$, though it is not clear how a particular value of α is considered “safe enough” in some sense.

As an additional comment for this last example, the mean for the *Pareto* model may no exist; it only does when $\delta > 1$ and even, in such a case, $\mathbb{E}(X) = \beta\delta/(\delta - 1)$, which implies that for values of δ sufficiently close to 1, it is possible to have $\mathbb{E}(X) > \text{VaR}_\alpha(X)$ for any given value $\alpha < 1$ because $\lim_{\delta \rightarrow 1^+} \mathbb{E}(X) = +\infty$. Since parameter δ controls tail heaviness of this probability distribution (lower values for δ imply heavier right tail), this exemplifies a comment at the beginning of the previous section in the sense that it is better to use the median instead of the mean.

4 Loss aggregation

Consider n excess of loss random variables L_1, \dots, L_n where $L_i = X_i - \mathbb{M}(X_i)$ for $i \in \{1, \dots, n\}$ as in Definition 2.1. It is of interest to calculate VaR of the aggregation of such random variables:

$$L = L_1 + \dots + L_n = \sum_{i=1}^n X_i - \sum_{i=1}^n \mathbb{M}(X_i) = S - c, \quad (2)$$

where the random variable $S := \sum_{i=1}^n X_i$ and the constant $c := \sum_{i=1}^n \mathbb{M}(X_i)$. In this case, we get $\text{VaR}_\alpha(L) = \text{VaR}_\alpha(S) - c$; so this last calculation essentially depends on obtaining or estimating the probability distribution function of S , that is F_S , because $\text{VaR}_\alpha(S) = F_S^{-1}(\alpha)$. Since S is a transformation of the n -dimensional random vector (X_1, \dots, X_n) , it is necessary to know either the joint probability distribution function $F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$ or its joint probability density function $f_{X_1, \dots, X_n}(x_1, \dots, x_n) \geq 0$ such that

$$\mathbb{P}[(X_1, \dots, X_n) \in B] = \int_B \dots \int f_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \dots dx_n.$$

A very popular probabilistic model is the *multivariate Normal* distribution, which undoubtedly has very nice mathematical properties that makes it very attractive for analysis and simplified calculations, but in practice it is usually inappropriate for the following reasons:

- All the univariate marginal distributions have to be *Normal*. Very often excess of loss random variables exhibits such a probabilistic behavior that is easily rejected by standard statistical normality tests, specially for heavier tails than those of the *Normal* distribution.
- The *multivariate Normal* is completely unable to consider *tail dependence* that very often is present among risks in finance and insurance, which consists in an important increase of the dependence degree under extreme values of the random variables involved.

These two flaws combined usually lead to a significant underestimation of the total aggregated risk. Instead, more flexible models have been explored, such as the ones built by means of *copula functions* which allow for any kind and distinct marginal univariate distributions and also account for tail dependence. Getting into the details of copula modeling is beyond the scope of the present article; the interested reader should refer to Nelsen (2006) for a book on basic copula theory, and the books by McNeil *et al.* (2015) and Denuit *et al.* (2005) for applications of copulas in finance and insurance risk modeling.

In two following sections, calculation of aggregated VaR will be considered in two extreme cases: Perfect positive dependence (comonotonicity) and complete absence of dependence (that is, independence). For simplicity, but without loss of generality, it is considered the aggregation of two excess of loss random variables, that is $L = L_1 + L_2$ where $L_1 = X - \mathbb{M}(X)$ and $L_2 = Y - \mathbb{M}(Y)$, which is equivalent to $L = S - c$ with $S := X + Y$ and $c := \mathbb{M}(X) + \mathbb{M}(Y)$ and therefore $\text{VaR}_\alpha(L) = \text{VaR}_\alpha(S) - c$.

5 Comonotonicity

The following result comes from the works by Hoeffding (1940) and Fréchet (1951) and it is known as the *Fréchet-Hoeffding bounds* for joint probability distribution functions, which for simplicity is stated for the bivariate case:

Lemma 5.1 (Fréchet–Hoeffding). *If (X, Y) is a random vector with joint probability distribution function $F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y)$ and marginal distribution functions $F_X(x) = \mathbb{P}(X \leq x)$ and $F_Y(y) = \mathbb{P}(Y \leq y)$, then:*

$$H_*(x, y) := \max\{F_X(x) + F_Y(y) - 1, 0\} \leq F_{X,Y}(x, y) \leq \min\{F_X(x), F_Y(y)\} =: H^*(x, y),$$

where the lower bound H_* and the upper bound H^* are both joint distribution functions and therefore infimum and supremum for all bivariate joint distribution functions.

Definition 5.1. Two random variables X and Y are *comonotone* or *perfectly positively dependent* if there exists a strictly increasing function g such that $\mathbb{P}[Y = g(X)] = 1$.

Proof of the following lemma may be found in Nelsen (2006) as Theorem 2.5.4 and following comment thereof:

Lemma 5.2 (Nelsen, 2006). *Let X and Y be continuous random variables with marginal distribution functions F_X and F_Y , respectively, and joint distribution function $F_{X,Y}$. Then X and Y are comonotone if and only if $F_{X,Y}$ is equal to the Fréchet-Hoeffding upper bound.*

Now the main result for this section:

Theorem 5.1. *If X and Y are continuous comonotone random variables, then:*

$$\text{VaR}_\alpha(X + Y) = \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y).$$

Proof:

Since X and Y are comonotone, there exists a strictly increasing function g such that $\mathbb{P}[Y = g(X)] = 1$. Hence, the distribution function of Y may be expressed as:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}[g(X) \leq y] = \mathbb{P}[X \leq g^{-1}(y)] = F_X(g^{-1}(y)).$$

By Lemma 5.2, we get:

$$F_{X,Y}(x, y) = \min\{F_X(x), F_Y(y)\} = \min\{F_X(x), F_X(g^{-1}(y))\}.$$

Define $S := X + Y$. Then, its distribution function satisfies:

$$F_S(s) = \mathbb{P}(S \leq s) = \mathbb{P}(X + Y \leq s) = \mathbb{P}(X + g(X) \leq s) = \mathbb{P}(Y \leq s - X).$$

Since $\mathbb{P}[Y = g(X)] = 1$, then $F_{X,Y}$ is a singular distribution because all the probability is distributed along the curve $y = g(x)$ and therefore, $F_S(s)$ is equal to the value of $F_{X,Y}$ at the intersection point (x_*, y_*) between the increasing curve $y = g(x)$ and the decreasing line $y = s - x$, for all $s \in \text{Ran } g$, which requires $g(x) = s - x$ and hence the intersection point is $(x_*, g(x_*))$ where x_* is the solution to the equation $x + g(x) = s$ which will be denoted as $x_* = h(s)$. Since g is strictly increasing so it is h which has inverse $h^{-1}(x) = x + g(x)$. Then:

$$F_S(s) = F_{X,Y}(x_*, g(x_*)) = \min\{F_X(x_*), F_X(g^{-1}(g(x_*)))\} = F_X(h(s)),$$

and consequently:

$$\begin{aligned} \text{VaR}_\alpha(X + Y) &= \text{VaR}_\alpha(S) = F_S^{-1}(\alpha) = h^{-1}(F_X^{-1}(\alpha)) \\ &= F_X^{-1}(\alpha) + g(F_X^{-1}(\alpha)) = \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) \quad \square \end{aligned}$$

Corollary 5.1. *If X and Y are continuous comonotone random variables, then for the excess of loss random variables $L_1 := X - \mathbb{M}(X)$ and $L_2 := Y - \mathbb{M}(Y)$, we have that:*

$$\text{VaR}_\alpha(L_1 + L_2) = \text{VaR}_\alpha(L_1) + \text{VaR}_\alpha(L_2).$$

Proof:

$$\begin{aligned} \text{VaR}_\alpha(L_1 + L_2) &= \text{VaR}_\alpha(X + Y - \mathbb{M}(X) - \mathbb{M}(Y)) = \text{VaR}_\alpha(X + Y) - \mathbb{M}(X) - \mathbb{M}(Y) \\ &= \text{VaR}_\alpha(X) - \mathbb{M}(X) + \text{VaR}_\alpha(Y) - \mathbb{M}(Y) = \text{VaR}_\alpha(L_1) + \text{VaR}_\alpha(L_2) \quad \square \end{aligned}$$

Example 5.1. Let X be a *Pareto* random variable with parameters $\beta = 1$ and $\delta > 0$ and define the random variable $Y := X^2$. Since $Y = g(X)$ with $g(x) = x^2$ a strictly increasing function on $\text{Ran } X =]1, +\infty[$ then X and Y are comonotone, with $\text{Ran } Y =]1, +\infty[$ also. Making use of the formulas in Example 3.1, we obtain:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) = 1 - \left(\frac{1}{y}\right)^{\delta/2}, \quad y > 1, \end{aligned}$$

which implies that Y is also a *Pareto* random variable but with parameters $\beta = 1$ and $\delta/2$. Therefore:

$$\text{Var}_\alpha(X) = (1 - \alpha)^{-1/\delta}, \quad \text{Var}_\alpha(Y) = (1 - \alpha)^{-2/\delta}.$$

Now let $S := X + Y = X + X^2$ where $\text{Ran } S =]2, +\infty[$ and we get:

$$\begin{aligned} F_S(s) &= \mathbb{P}(S \leq s) = \mathbb{P}(X + X^2 \leq s) = \mathbb{P}(X \leq (\sqrt{1 + 4s} - 1)/2) \\ &= F_X((\sqrt{1 + 4s} - 1)/2) = 1 - (2/(\sqrt{1 + 4s} - 1))^\delta, \quad s > 2, \end{aligned}$$

from where we obtain for any $0 < \alpha < 1$ the following:

$$\text{VaR}_\alpha(X + Y) = \text{VaR}_\alpha(S) = F_S^{-1}(\alpha) = (1 - \alpha)^{-1/\delta} + (1 - \alpha)^{-2/\delta} = \text{Var}_\alpha(X) + \text{Var}_\alpha(Y),$$

as expected. \square

6 Independence

In contrast with the comonotonicity case where such property always implies that the VaR of the sum is equal to sum of the individual VaRs, under lack of dependence (independence), it is not possible to establish a general formula that relates the VaR for a sum of independent random variables to the individual VaRs, it will depend on each particular case, as it is shown in the following three examples:

Example 6.1. Let X and Y be independent and identically distributed *Pareto* random variables with parameters $\beta = 1$ and $\delta = 1$ such that the right tail of their distributions is heavy enough for non existence of a mean. Again applying formulas from Example 3.1, we get $\text{VaR}_\alpha(X) = (1 - \alpha)^{-1} = \text{VaR}_\alpha(Y)$ where $0 < \alpha < 1$ and, by independence, the joint density function for the random vector (X, Y) is the product of the marginal densities:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = \frac{1}{x^2y^2}, \quad x > 1, y > 1.$$

Let $S := X + Y$ then $\text{Ran } S =]2, +\infty[$ and its distribution function:

$$\begin{aligned} F_S(s) &= \mathbb{P}(S \leq s) = \mathbb{P}(X + Y \leq s) = \mathbb{P}(Y \leq s - X) = \iint_{y \leq s-x} f_{X,Y}(x, y) dx dy \\ &= \int_1^{s-1} x^{-2} \int_1^{s-x} y^{-2} dy dx = 1 - \frac{2}{s} - \frac{2}{s^2} \log(s-1), \quad s > 2. \end{aligned}$$

Let $s_* := \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) = 2/(1 - \alpha) > 2$. Then:

$$F_S(s_*) = \alpha - \frac{(1 - \alpha)^2}{2} \log\left(\frac{1 + \alpha}{1 - \alpha}\right) < \alpha,$$

which implies for any $0 < \alpha < 1$:

$$\text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) = s_* < F_S^{-1}(\alpha) = \text{VaR}_\alpha(S) = \text{VaR}_\alpha(X + Y).$$

Despite total absence of dependence between the random variables, the right tails of their distributions are heavy enough such that the diversification effect is definitely not convenient: The VaR of the sum is greater than the sum of the individual VaRs, in this particular case. \square

Example 6.2. Now let X and Y be independent and identically distributed *Normal* $(0, 1)$ random variables. Their distribution function is expressed as:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt.$$

The tails of this distribution are not as heavy as in the previous example, and it has finite mean and variance. Then the random variable $S := X + Y$ has *Normal* $(0, 2)$ distribution, which is the same as $\sqrt{2} X$ since a linear transformation of a *Normal* random variable is still *Normal* and $\mathbb{E}(\sqrt{2} X) = \sqrt{2} \mathbb{E}(X) = 0$ and $\mathbb{V}(\sqrt{2} X) = 2\mathbb{V}(X) = 2$. Therefore the distribution function of S may be expressed as:

$$F_S(s) = \mathbb{P}(S \leq s) = \mathbb{P}(\sqrt{2} X \leq s) = \mathbb{P}(X \leq s/\sqrt{2}) = \Phi(s/\sqrt{2}),$$

and its quantile function as $F_S^{-1}(u) = \sqrt{2} \Phi^{-1}(u)$, $0 < u < 1$. Consequently, for any $0 < \alpha < 1$:

$$\text{VaR}_\alpha(X + Y) = \text{VaR}_\alpha(S) = F_S^{-1}(\alpha) = \sqrt{2} \Phi^{-1}(\alpha) < 2\Phi^{-1}(\alpha) = \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y).$$

In contrast with the previous example, the VaR of this sum of random variables is strictly less than the sum of the individual VaRs, and therefore in this particular case diversification is clearly convenient. \square

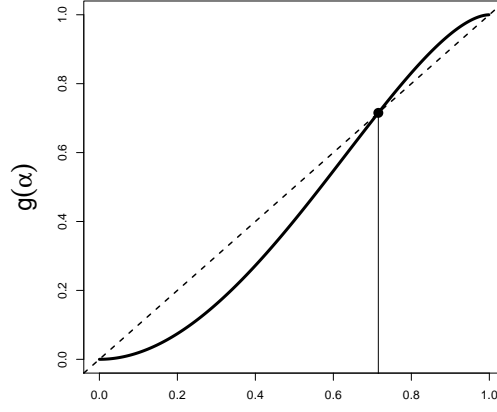


Figure 1: Graph of $g(\alpha) = F_S\left(\frac{\alpha}{-2 \log(1 - \alpha)}\right)$ in Example 6.3.

Example 6.3. Lastly, let X and Y be independent and identically distributed *Exponential* random variables with parameter equal to 1. The right tail of this distribution is not as heavy as in Example 6.1 but certainly heavier than in Example 6.2, with finite mean and variance. Their marginal probability density function is $f(x) = e^{-x}$, $x > 0$, and the corresponding distribution function $F(x) = 1 - e^{-x}$, $x > 0$, hence $\text{VaR}_\alpha(X) = -\log(1 - \alpha) = \text{VaR}_\alpha(Y)$ where $0 < \alpha < 1$. By independence the joint density function of the random vector (X, Y) is the product of the marginal densities:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = e^{-(x+y)}, \quad x > 0, y > 0.$$

Let $S := X + Y$, then $\text{Ran } S =]0, +\infty[$ and its distribution function is:

$$\begin{aligned} F_S(s) &= \mathbb{P}(X + Y \leq s) = \iint_{y \leq s-x} f_{X,Y}(x, y) dx dy \\ &= \int_0^s e^{-x} \int_0^{s-x} e^{-y} dy dx = 1 - e^{-s}(1 + s), \quad s > 0. \end{aligned}$$

By the way, calculating the derivative of $F_S(s)$, we get $f_S(s) = se^{-s}$, $s > 0$, which is a density of a *Gamma*(2, 1) random variable. Let $s_* := \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) = -2 \log(1 - \alpha)$. Then:

$$g(\alpha) := F_S(s_*) = 1 - (1 - \alpha)^2(1 - 2 \log(1 - \alpha)), \quad 0 < \alpha < 1.$$

By numerical approximation it is obtained that $g(\alpha) = \alpha$ if and only if $\alpha \approx 0.7153319$, see Figure 1, $g(\alpha) < \alpha$ if $\alpha < 0.7153319$ and $g(\alpha) > \alpha$ if $\alpha > 0.7153319$, which implies that

$$\text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) \begin{cases} < \text{VaR}_\alpha(X + Y) & \text{if } \alpha < 0.7153319 \\ = \text{VaR}_\alpha(X + Y) & \text{if } \alpha \approx 0.7153319 \\ > \text{VaR}_\alpha(X + Y) & \text{if } \alpha > 0.7153319 \end{cases}$$

This is an example where diversification convenience depends on the desired α level for VaR, in contrast with the two previous examples. \square

7 Final remarks

The main conclusion in the present work is that diversification is **not always** convenient. As shown in the examples, risk diversification may result better, worse or equivalent to lack thereof, depending on the

individual risks involved and the dependence relationship between them, and even on the desired risk level. In particular, as a consequence of Theorem 5.1, if two continuous random variables are comonotone then we can guarantee that the VaR is always equal to the sum of the individual VaRs. But for independent random variables everything may happen.

Moreover, it is argued that the fact of VaR being not subadditive is better and advantageous: When the VaR of a sum is greater than the sum of individual VaRs, we would be detecting a specially pernicious combination of risks on which is not convenient to diversify; while under “coherent” risk measures as in Definition 2.6 where subadditivity is always present, it would not be possible to detect such a harmful risk combination.

Bibliography

- Denuit, M.; Dhaene, J.; Goovaerts, M. and Kaas, R. (2005): *Actuarial Theory for Dependent Risks*. Chichester: Wiley.
- Fréchet, M. (1951): “Sur les tableaux de corrélation dont les marges sont données”. *Annales de l’Université de Lyon. Section A: Sciences, Mathématiques et Astronomie* **14**, 53–77.
- Hoeffding, W. (1940): “Masstabinvariante Korrelationstheorie”. *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin* **5**, 179–223.
- McNeil, A.J.; Frey, R. and Embrechts, P. (2015): *Quantitative Risk Management*. New Jersey: Princeton University Press.
- Nelsen, R.B. (2006): *An Introduction to Copulas*. New York: Springer.