Correlation: the most common myths in financial risk management practice

Giovanni Puccetti,* Giacomo Cagliani [†]

January 28, 2026

Abstract

Financial institutions are required to hold capital reserves to protect against future losses, and many rely on mathematical models to estimate the necessary amounts. While these models can make risk measurement more precise, they are often used by practitioners who may not fully grasp their assumptions or limitations. One recurring source of error lies in the treatment of correlation—the statistical link between different risks—which plays a crucial role in determining how losses might combine. Misjudging these relationships has, in the past, amplified systemic vulnerabilities, most famously in the misuse of the Gaussian copula model during the subprime crisis. This paper revisits the problem of correlation mistaken beliefs from a pedagogical standpoint, using simplified examples and reproducible R code to show how small misunderstandings can have large consequences. The discussion aims to encourage greater critical awareness of quantitative tools among risk managers, regulators, and students alike.

Keywords: Correlation, Statistical misconceptions, Rank correlation, Dependence modeling, Risk aggregation, Value-at-Risk.

1 Introduction and motivation

Under the current Basel and Solvency regulatory frameworks, banks and insurance companies are required to maintain capital reserves to offset potential future losses. To determine the minimum amount of such reserves for specific categories of risk—such as market, credit, and operational risk—financial institutions generally have two methodological options: either use a mathematically elementary standardized approach or choose a more advanced approach based on an internal stochastic model.

The use of internal models for determining risk capital is typically authorized by supervisory authorities only for large, internationally active institutions that possess a robust risk management framework, employ professionals skilled in advanced mathematical and statistical techniques, and demonstrate a proven track record of accuracy in measuring risk. The advanced approach allows institutions to better tailor their capital requirements to their specific risk profiles, but it also offers financial incentives: under certain conditions, banks may, in principle, reduce their capital requirements to as low as 72.5% of the amount computed under the standardized approach; see Section RBC20.4 in Basel Committee on Banking Supervision (2026).

---

*Giovanni Puccetti (`giovanni.puccetti@unimi.it`), Department of Economics, Management and Quantitative Methods, University of Milan, Italy.

†Giacomo Cagliani (`giacomo.cagliani@outlook.com`).

The Basel Committee on Banking Supervision first opened the door to the use of quantitative models in the financial industry with the introduction of the Basel II Accord (Basel Committee on Banking Supervision, 2004), which formally recognized the possibility of adopting internal advanced models. Mathematics undoubtedly brings rigor and structure to the calculation of capital requirements, but it must be handled with care. Mathematical models are sometimes operated by individuals who do not fully understand or control the underlying assumptions required for those models to make sense. In some cases, practitioners are permitted to compute capital reserves without a solid grasp of the mathematical premises on which their models depend.

One area where this problem appears frequently is risk aggregation, the process of determining the capital reserve for a portfolio or sum of risks. This task is commonly performed (for example, in market risk capital calculations) using correlation matrices or copula functions. A well-known case that received worldwide media coverage and illustrates the dangers of misusing such models was the misuse of the Gaussian copula model (with a single correlation parameter) applied to Collateralized Debt Obligation (CDO) pricing; see Salmon (2012). The model's reliance on an overly simplistic correlation assumption was blamed for contributing to one of the most significant modeling failures in recent financial history. An interview (Puccetti and Scherer, 2018) with David Li, who introduced a (not specifically Gaussian) copula model for CDO pricing (Li, 2000), provides useful context for this episode. We recommend this interview to practitioners and regulators interested in understanding the risks associated with the misuse of copulas and correlation structures in risk modeling.

This paper aims to highlight some of the most common mistaken beliefs, or "myths", regarding the use of correlation financial risk management. A first attempt in this direction was made in the milestone paper Embrechts et al. (2002). However, more than two decades after it first appeared, that "correlation pitfalls" paper feels too technical for a master's student or a practitioner without a strong mathematical or statistical background. To reach a broader audience, this paper takes a deliberately less technical approach, reducing mathematical detail and avoiding copulas; interested readers are referred to the aforementioned reference. This manuscript should be easily understood by an undergraduate student, a practitioner, or even a regulator or auditor without any training in mathematics or statistics beyond what is typically covered in introductory courses. Moreover, pedagogical (counter)examples are illustrated with the corresponding R code, allowing readers to reproduce or modify them easily and thereby immediately grasp the rationale and warnings behind them. Although the paper has a focus on financial risk management, the general cautions regarding the use of correlation are universally valid across all domains where correlation is applied (and possibly misunderstood).

First, we recall the definition of Pearson's linear correlation coefficient, which we will simply refer to as correlation in what follows. On a given non-atomic probability space, let X and Y be two square integrable random variables (i.e., having finite second moments: $\mathbb{E}[X^2]$, $\mathbb{E}[Y^2] < \infty$). Pearson's correlation coefficient is defined by

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}},$$

where $\text{cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$ is the *covariance* between X and Y, $\text{var}(X) = \text{cov}(X, X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$ is the *variance* and $\mathbb{E}[X]$ the *expected value* (mean) of X (analogously for Y). In the following, all equalities involving random variables are intended with probability one.

Another notion that we need is that of the *quantile* F^{-1} of a distribution (function) F, defined as

$$F^{-1}(u) = \inf \left\{ x \in \mathbb{R} : F(x) \geq u \right\}, \ u \in (0, 1],$$

and $F^{-1}(0) = \inf \left\{ x \in \mathbb{R} : F(x) > 0 \right\}$, with the convention that $\inf \emptyset = \infty$. If F is continuous and strictly increasing, F^{-1} coincides with its unique inverse; see Embrechts and Hofert (2013). In finance, F^{-1} is referred to as the *Value-at-Risk* (see Myth 4).

The original scope of Pearson's correlation as pursued for instance in the early studies of Hoeffding (1940) (translated in Hoeffding, 1994) and Fréchet (1951), was to provide a measure of similarity (association) between random variables. It is useful then to see how a higher correlation in fact corresponds to a higher similarity. Suppose that two square integrable random variables X and Y have fixed distributions F and G, respectively (in what follows we denote this by $X \sim F$ and $Y \sim G$). Their covariance can be written as

$$\text{cov}(X, Y) = \mathbb{E}[X\,Y - X\mathbb{E}[Y] - \mathbb{E}[X]Y + \mathbb{E}[X]\mathbb{E}[Y]] = \mathbb{E}[X\,Y] - \mathbb{E}[X]\mathbb{E}[Y]. \tag{1}$$

Since the first moments $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are given by the fixed distributions (as well as $\text{var}(X)$ and $\text{var}(Y)$), the only correlation term that depends on the similarity of the two random variables is $\mathbb{E}[X\,Y]$. Why does the expectation of the product play such a role? This can be explained via the formula

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[X^2] - 2\,\mathbb{E}[X\,Y] + \mathbb{E}[Y^2], \tag{2}$$

which implies that the maximization of the expected product of two random variables with fixed distributions is equivalent to the minimization of the average squared Euclidean distance between them (this holds analogously in multivariate spaces; see Puccetti, 2022). As their expected product increases, the correlation between two random variables rises and their expected distance decreases, motivating the interpretation of correlation as a measure of similarity.

We believe that most misunderstandings about correlation originate from the fact that correlation is always told to assume values between -1 and 1, with these bounds identifying the case of perfect negative and perfect positive linear dependence. Though this statement is formally correct, confusion stems from the fact that these bounds are not always achieved for a fixed pair of marginal distributions and do not represent in general the minimum and maximal attainable values of correlation.

Moreover, it is well known that if X and Y are square integrable and stochastically independent, one has $\mathbb{E}[X\,Y] = \mathbb{E}[X]\mathbb{E}[Y]$ and correlation is zero. However, the converse is not true, and in a strong sense. By this we mean that, in the absence of further assumptions, a value of (sample) correlation close to zero does not have any relevant implication, and might derive from independence as well from perfect positive or negative dependence.

Correlation finds its ideal environment within a family of fully specified statistical models closed under linear combinations, such as the multivariate Normal (Gaussian) or Student's t distributions. In the lack of such (strong) assumptions, mistakes and wrong interpretations of correlation values may arise.

To avoid some of the potential mistakes and misinterpretations that we discuss in what follows, we strongly suggest a straightforward alternative to correlation, which is *rank correlation*. To be safe in real practice, one should not compute correlation but correlation between ranks, something that can always be done at very low extra computational effort. Rank correlation does not fix all problems of correlation, but it overcomes many: it is a quick, viable and better alternative to plain correlation.

Myth 1
"We can always estimate correlation"

Linear correlation has an appealingly simple definition, and in most statistical software it can be computed with the press of a button. However, its definition is based on the assumption that the random variables involved have finite second moments, that is, their variances (and therefore their means) have a finite value. Missing this assumption is one of the most common practical pitfalls. From a finite dataset alone, it is not always statistically obvious whether the underlying distribution truly has a finite variance.

The following pedagogical example works well as an in-class experiment for a quantitative risk management course or, more broadly, for any statistics class. Consider two sets of standardized data $x_i, y_i, 1 \leq i \leq N$, with sample mean zero and sample variance one, generated from two different distributions: a Normal distribution for the x-data and a Student's t distribution for the y-data. In a financial interpretation, these could represent two time series of daily log-losses for two stocks; see, for instance, the statistical analysis in Di Lascio et al. (2018).

```r
set.seed(100) # seed for reproducibility
N=500 # sample size
## Generate samples from X and Y
X=rnorm(N) # Normal sample
Y=rt(N,df=2) # t sample
## Standardize samples
X=(X-mean(X))/sd(X)
Y=(Y-mean(Y))/sd(Y)
## Compute sample mean of X and Y
c(mean(X),mean(Y))
```

```
## [1] 3.952394e-17 -1.332268e-17
```

```r
## Compute sample variance of X and Y
c(var(X),var(Y))
```

```
## [1] 1 1
```

```r
## Compute sample correlation between X and Y
cor(X,Y)
```

```
## [1] -0.03499964
```

As the output above shows, computing a sample correlation in R is immediate. The more subtle point is that the function cor will always return a number, even when the data plausibly come from a distribution for which the variance is not finite.

How can one realize whether the correlation between two random variables $X \sim F$ and $Y \sim G$ actually exists? Suppose we want to estimate $\rho(X, Y)$ based on a number N of independent joint observations

$$(x_i, y_i), \ i = 1, \ldots, N,$$

coming from the random vector $(X, Y)'$.

To assess whether the moments of the underlying distributions exist, a particularly handy diagnostic is the so-called *Mean Excess (ME) function*. When $\mathbb{E}[X]$ is finite, the mean excess function $e(u)$ is defined by

$$e(u) = \mathbb{E}[X - u \,|\, X > u],$$

that is, as the mean of exceedances of a random variable X above a given threshold u. Given the x-data $x_1, \ldots, x_N$, the sample version of the ME function is

$$e_N(u) = \frac{\sum_{i=1}^{N}(x_i - u)I_{\{x_i > u\}}}{\sum_{i=1}^{N} I_{\{x_i > u\}}},$$

namely, the sum of the exceedances above the threshold u divided by the number of observations that exceed u. The scatter plot of the ordered data

$$\{x_{(i)}, e_N(x_{(i)}), \, i = 1, \ldots, N - 1\},$$

is called the *ME-plot*. In short: rather than averaging the observations in a sample, the ME function (and its plot) averages their exceedances over a sequence of thresholds, chosen as the ordered sample values. In practice, one usually omits the largest values of $e_N(x_{(i)})$ since these depend on only a handful of observations and can make the plot look erratic. How many points to drop is somewhat arbitrary. In Figure 2, for example, the five largest observations are omitted.

An ME-plot can reveal a surprising amount of information about the tail behavior of a distribution at a glance. A convenient rule of thumb is suggested by Figure 1, which shows the shape of the ME function for several common distributions. If the ME-plot is decreasing or roughly constant, then it is plausible that all moments of the distribution are finite (as in the Normal and Exponential examples). If, instead, the ME-plot increases or looks unstable, then caution is warranted: before estimating moment-based quantities, one should carry out a more careful statistical investigation; see also Cirillo (2013) for some caveats and more refined techniques.

Figure 2 shows the ME-plot for the x- and y-data (the R code used to produce the figure is provided in the additional materials of this paper). The increasing ME-plot for the y-data is an immediate warning sign: the underlying distribution may not have a finite variance. Indeed, the x-sample is generated from a Normal distribution, for which all moments (not only the first and the second) exist, while the y-sample comes from a Student's t with $\nu = 2$ degrees of freedom, a distribution which does not possess a finite variance. Note that requiring finite variances has an important implication: $\rho(X, Y)$ depends on the marginal distributions of X and Y. This somewhat contradicts the goal of correlation, which is meant to measure association independently of distributions or measurement scale. More on this will follow in Myth 3.

In summary, although the sample evidence might tempt us to report an "almost zero" correlation between X and Y, the quantity $\rho(X, Y)$ actually does not exist. The same warning applies to any statistic built on moments (including something as basic as the mean).

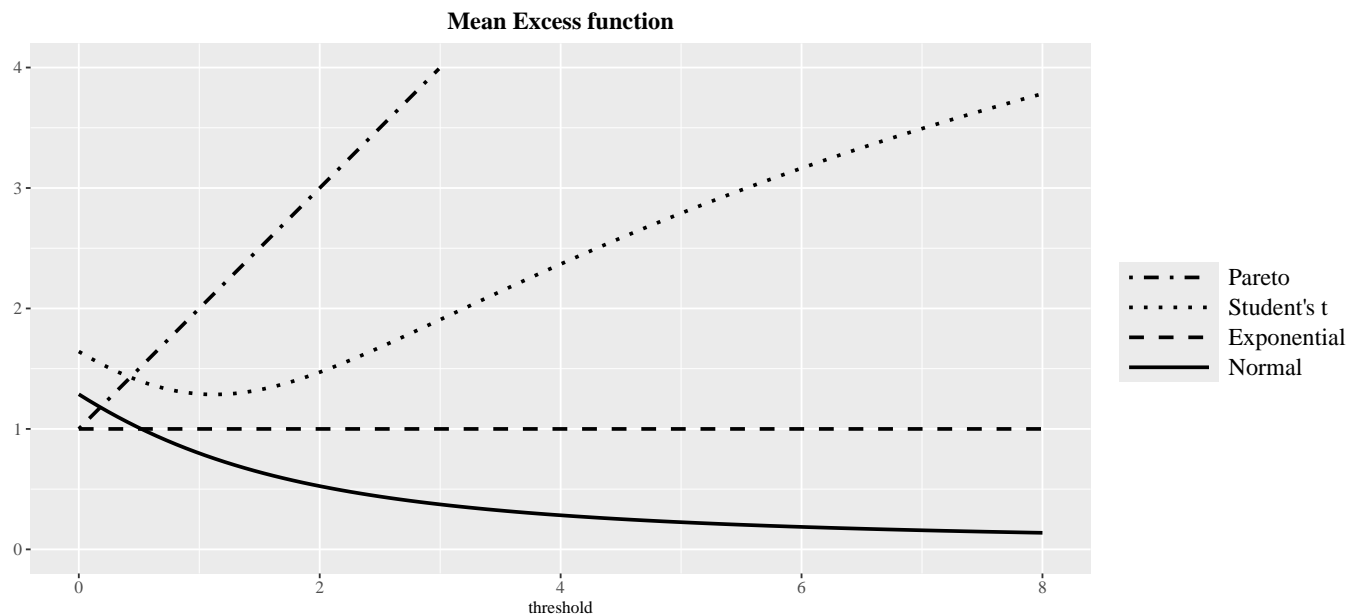> **Key message: Before estimating something, make sure that it exists!**

Figure 1: Mean Excess function for some distributions having mean 1. The t and the Pareto distributions considered here do not possess a finite variance.
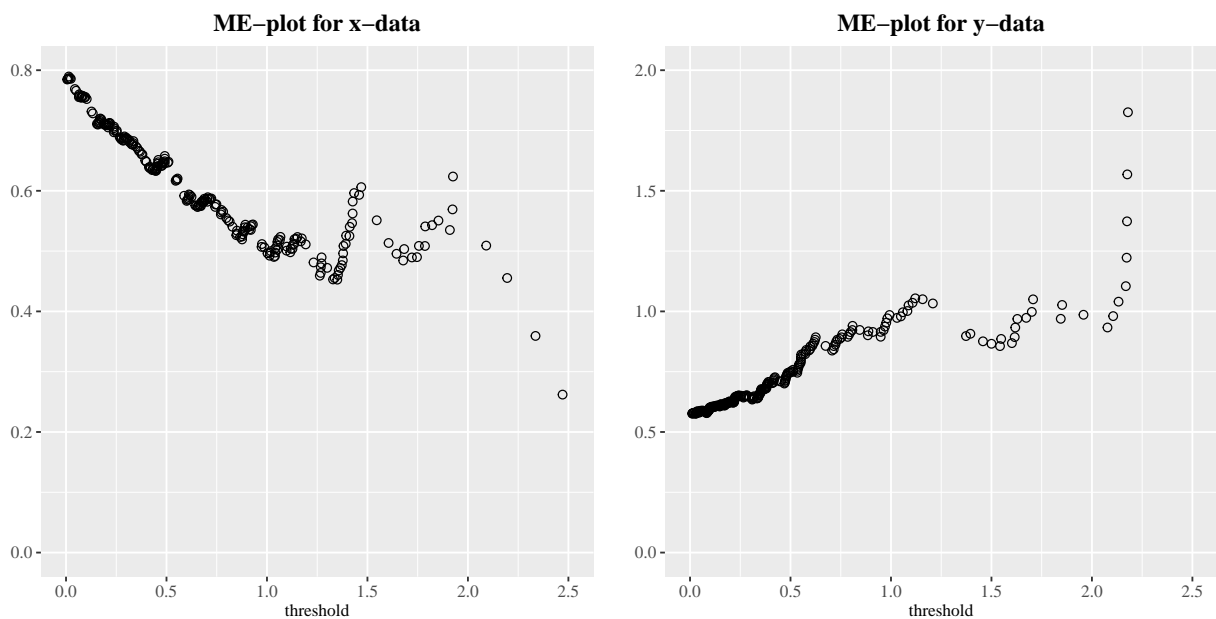

Figure 2: ME-plot for two data samples having the same sample mean and sample variance. The five largest observations are omitted.

Myth 2
"We estimated an almost zero correlation so the risks are independent"

It is well known that if X and Y have finite variances and are independent, then $\rho(X,Y) = 0$. This implication is often a source of confusion, because one naively associates a sample correlation close to zero with almost independence. For example, suppose we obtain a sample correlation of approximately -0.0027. How should this value be interpreted? Setting aside concerns about infinite moments and assuming that both variables have finite variance, does this estimate convey any substantive information?

Figure 6 displays a correlation matrix widely used in risk management for the calculation of the Solvency Capital Requirement (SCR) under the Standard Formula. We will discuss the overall structure of this matrix in Myth 5. For now, let us focus on one specific modeling choice: the so-called life and non-life risk modules for an insurance company are assigned a correlation of zero. What conclusion, if any, is the regulator hoping to draw from this assumption?

The uncomfortable answer is: in general, not much. Uncorrelatedness does not imply independence and, without specifying a full joint model for the random vector $(X, Y)'$, it does not pin down any meaningful form of (in)dependence.

A classical illustration is to let $X \sim N(0,1)$ and $Y = X^2$. The variables are clearly dependent—indeed, X determines Y exactly—yet $\rho(X,Y) = 0$. We illustrate this in the following pedagogical example (where sample correlation is, in fact, -0.0027).

```
set.seed(100) # seed for reproducibility
N=10^6 # sample size
X=rnorm(10^6) # standard Normal sample
Y=X^2 # Y is X squared (so not independent from X)
cor(X,Y) # sample correlation between X and Y
```

```
## [1] -0.002683984
```

More generally, we can state the following proposition, which provides a whole family of counterexamples.

Proposition 1.1. *Let the random variable X be symmetric (about zero) and assume that $\mathbb{E}[X^{2k+1}]$ is finite for some integer k. Then $\rho(X, X^{2k}) = 0$.*

Proof. Being X symmetric, notice that $\mathbb{E}[X^{2i+1}] = 0$, $0 \leq i \leq k$. Using (1), we obtain

$$\text{cov}(X, X^{2k}) = \mathbb{E}[X^{2k+1}] - \mathbb{E}[X]\mathbb{E}[X^{2k}] = \mathbb{E}[X^{2k+1}] = 0. \quad \square$$

In short, independence is sufficient to make correlation vanish, but it is far from necessary. As a result, one should resist drawing quick conclusions from a small (or even zero) value of sample correlation. As we will see in Myth 3, a sample correlation close to zero could arise under independence—but it can also occur under very strong forms of dependence, including perfectly positive or perfectly negative dependence.

> **Key message: Correlation is zero—so what?**

Myth 3

"We estimated a moderate value of correlation so the risks are moderately dependent"

We can push the previous counterexample a bit further. In practice, one would almost never estimate a correlation that is exactly zero even if, in a regulatory matrix, a value of zero may be imposed by design. Much more commonly, one encounters sample correlations that are merely close to zero.

A related misconception is to read a relatively small correlation as evidence of only a modest association between the variables. This intuition is understandable: correlation is well known to take values between -1 and 1 (and, if you ask students, this is often the first fact they will quote!). Unfortunately, this is not always correct. In general, the extreme values correlation can attain depend on the marginal distributions of the variables involved and, somewhat surprisingly, these extremes can be arbitrarily close to zero.

We give a very simple but enlightening example. Consider a random variable $X_1 \sim F$ and a second random variable Y_1 being a linear function of X_1, namely $Y_1 = a X_1 + b$ for some $a > 0$ and $b \in \mathbb{R}$. In this case, the linear correlation between them is given by

$$\rho(X_1, Y_1) = \frac{\text{cov}(X_1, aX_1 + b)}{\sqrt{\text{var}(X_1)}\sqrt{\text{var}(aX_1)}} = \frac{a\,\text{cov}(X_1, X_1)}{\sqrt{\text{var}(X_1)}\sqrt{a^2\,\text{var}(X_1)}} = \frac{a\,\text{var}(X_1)}{|a|\sqrt{\text{var}(X_1)}\sqrt{\text{var}(X_1)}} = 1.$$

From the Cauchy–Schwarz inequality (see for instance Th. 4.5.7 in Casella and Berger, 2002), it follows that $\rho(X, Y) = 1$ if and only if Y is a positive linear function of X (as in the previous case) and $\rho(X, Y) = -1$ if and only if Y is a negative linear function of X (e.g., if above one has $a < 0$). These lower and upper bounds correspond to perfect positive and perfect negative *linear* dependence, and they justify the name *linear correlation*. The following simulation confirms this when X_1 is Normal, but one could choose any square integrable distribution.

```
set.seed(100) # seed for reproducibility
N=10^6 # sample size
X1=rnorm(N) # standard Normal N(0,1) sample
a=3 # standard deviation of Y
Y1=a*X1 # Y1 is a times X1 so has df N(0,a^2)
## Compute sample correlation between X1 and Y1
cor(X1,Y1)
```

```
## [1] 1
```

Correlation is excellent at detecting perfect linear dependence. On the other hand, it can be remarkably short-sighted if a non-linear dependence enters the picture. To see this, consider the random variables $X_2 = \exp(X_1)$ and $Y_2 = \exp(Y_1) = \exp(aX_1)$ (the exponential can be replaced by any strictly increasing non-linear function $f(X_1)$ and $f(Y_1)$). Applying the exponential to X_1 and Y_1 simply means that the same underlying quantities are now being reported on a different (but strictly increasing) scale. One might therefore expect that the correlation between them remains unchanged. Unfortunately, Pearson correlation need not do so: the correlation coefficient can change dramatically and may even become arbitrarily close to 0.

```
X2=exp(X1) # change of scale for X1 (which becomes LogNormal)
Y2=exp(Y1) # change of scale for X2 (which becomes LogNormal)
## Compute sample correlation between X2 and Y2
cor(X2,Y2)
```

[1] 0.2782197

Here the two random variables exhibit a relatively small correlation, yet they are perfectly dependent (indeed, linked by a strictly increasing function!).

As said, the confusion typically comes from a very natural mental shortcut: since correlation is (often said to be) bounded between -1 and 1, a "moderate" value (like the one above) must correspond to a "moderate" degree of dependence. This is not true in general, unless the relationship between the variables is genuinely close to linear. For jointly Normal variables, for instance, correlation carries substantial information about dependence; for lognormal variables it generally does not (apart from when they have exactly the same parameters).

So, what are the minimum and maximal values of linear correlation, and when are they attained? According to the definition (1), the higher the expected product $\mathbb{E}[XY]$, the higher the correlation between the two random variables X and Y having fixed marginal distributions. The question then becomes, for $X \sim F$ and $Y \sim G$, when does $\mathbb{E}[XY]$ attain its minimal and maximal values? The answer goes back to the milestone result given in Theorem 368 in Hardy et al. (1934), where it is proved that the scalar product of two vectors is maximal when the components of the two vectors are similarly ordered (i.e., they are monotonic in the same sense, or comonotonic), and minimal when they are oppositely ordered (i.e., they are countermonotonic).

If F and G are continuous, via approximations of continuous distributions with discrete ones, one obtains that: $\mathbb{E}[XY]$ is maximized when Y is (almost surely) an increasing function of X (we say that X and Y are *comonotonic*); minimized when Y is (a.s.) a decreasing function of X (we say that X and Y are *countermonotonic*); see for instance Puccetti and Wang (2015) for a proof and a history of the following result.

Theorem 1.2. *Assume that $X \sim F$ and $Y \sim G$, with F and G continuous and square integrable. Then we have that:*

 a) *$\mathbb{E}[XY]$ is maximized if and only if $Y = G^{-1}(F(X))$ a.s.;*

 b) *$\mathbb{E}[XY]$ is minimized if and only if $Y = G^{-1}(1 - F(X))$ a.s..*

As a consequence, the range of possible values of $\rho(X,Y)$ forms a closed interval $[\rho_{\min}, \rho_{\max}]$ with $\rho_{\min} < 0 < \rho_{\max}$. Since the distributions F, G, and the corresponding quantile functions F^{-1}, G^{-1}, are increasing, we have that:

 • *the maximum correlation $\rho_{\max}$ is attained if and only if X and Y are comonotonic;*

 • *the minimum correlation $\rho_{\min}$ is attained if and only if X and Y are countermonotonic.*

Based on Theorem 1.2, linear correlation (when it exists) attains its minimum value under countermonotonicity and its maximum value under comonotonicity. The extreme values, however, are equal to -1 and 1 only when the functional relationships $G^{-1}(F)$ and $G^{-1}(1 - F)$ are linear. This is the case, for instance, when F and G belong to the same family of distributions closed under affine transformations, such as when $(X, Y)'$ is assumed to be multivariate Normal.

In the general case when F is not continuous, Theorem 1.2 holds analogously. In general, X and Y are said to be comonotonic if they are both increasing function of a common random factor U and they admit the representation $X = F^{-1}(U)$, $Y = G^{-1}(U)$, a.s., for a uniform random variable $U \sim U(0,1)$. Analogously, X and Y are said to be countermonotonic if and only if X is an increasing, Y a decreasing function of U: $X = F^{-1}(U)$, $Y = G^{-1}(1-U)$, a.s.; see Theorem 2 in Dhaene et al. (2002).

Based on Theorem 1.2, for $X \sim F$ with density f, and $Y \sim G$, we have that the maximum value of $\mathbb{E}[XY]$ is given by

$$\pi_{\max} = \int_{-\infty}^{\infty} x \, G^{-1}(F(x)) \, f(x) dx = \int_0^1 F^{-1}(u) \, G^{-1}(u) \, du,$$

where the second integral comes from the change of variable $u = F(x)$. Analogously, the minimal value of $\mathbb{E}[XY]$ is given by

$$\pi_{\min} = \int_{-\infty}^{\infty} x \, G^{-1}(1 - F(x)) \, f(x) dx = \int_0^1 F^{-1}(u) \, G^{-1}(1-u) \, du.$$

As a consequence, the range of possible values of correlation is given by $[\rho_{\min}, \rho_{\max}]$, where

$$\rho_{\min} = \frac{\pi_{\min} - \mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}, \quad \rho_{\max} = \frac{\pi_{\max} - \mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}. \tag{3}$$

As it is illustrated in Figure 3 for the case of a Normal and a Student's t distribution, correlation never attains the values -1 and 1, and both maximal and minimal correlation goes to zero when the number of degrees of freedom of the t distribution approaches 2 from above.
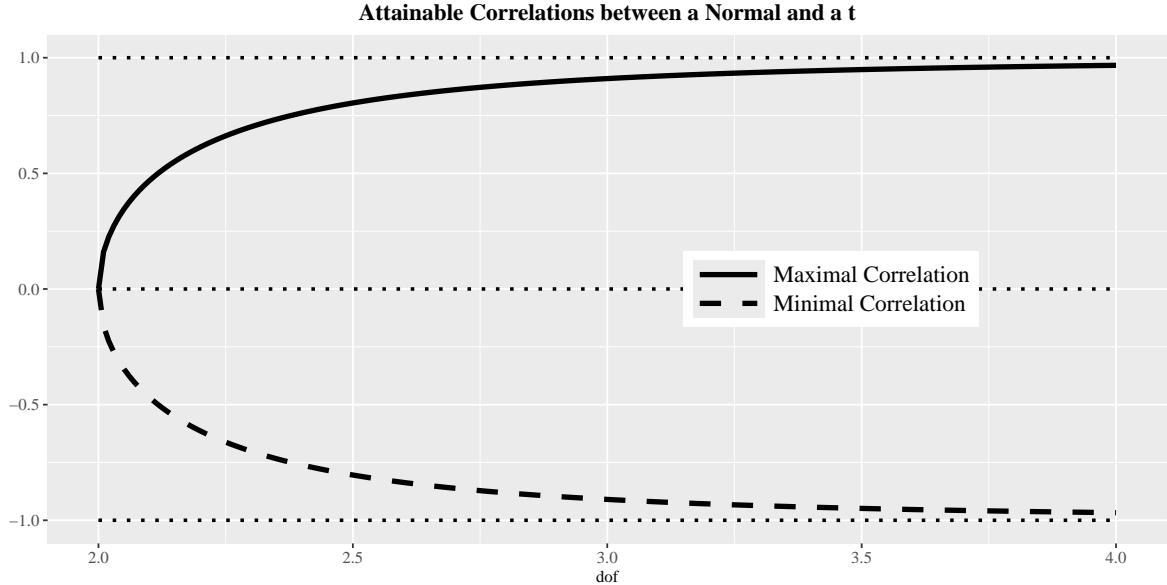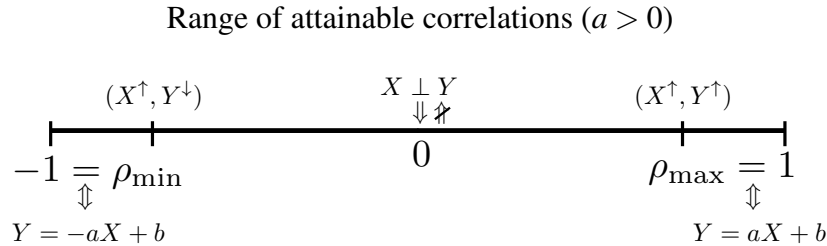

Figure 3: Maximal (solid curve) and minimal (dashed curve) values attainable by correlation between a standard Normal and a standard t distribution, computed versus the number of degrees of freedom (dof) of the t distribution.

Summarizing the above results: the maximum value of correlation is attained in the case of perfect positive dependence (comonotonicity), and the minimum value of correlation is attained in the case of perfect negative dependence (countermonotonicity). However, the corresponding attained values are -1 and 1 if and only if these dependencies are linear, otherwise the minimal value will be strictly larger than -1 and the maximal value strictly smaller than 1.

<div align="center">

Range of attainable correlations ($a > 0$)

</div>

$$(X^\uparrow, Y^\downarrow) \qquad\qquad \begin{array}{c} X \perp Y \\ \Downarrow \nRightarrow \end{array} \qquad\qquad (X^\uparrow, Y^\uparrow)$$

$$\underset{\substack{\Updownarrow \\ Y = -aX + b}}{-1 = \rho_{\min}} \qquad\qquad 0 \qquad\qquad \underset{\substack{\Updownarrow \\ Y = aX + b}}{\rho_{\max} = 1}$$

Fortunately, there is a quick remedy if one wants an association measure that is invariant under increasing transformations of the variables: compute correlation between ranks. For continuous marginals, this corresponds to estimating the *Spearman rank correlation coefficient* $\rho_S(X, Y)$, defined, for $X \sim F$ and $Y \sim G$, by

$$\rho_S(X, Y) = \rho(F(X), G(Y)),$$

see also Hofert et al. (2018) for more details. Rank correlation ρ_S always exists and takes values between -1 and 1; moreover, for any pair of continuous marginals F and G, both bounds are attainable, corresponding to countermonotonicity and comonotonicity, respectively (see Theorem 3 in Embrechts et al., 2002). In the previous example, encountering a moderate value of rank correlation at least allows one to conclude that the variables are not perfectly associated—a conclusion that plain correlation would not support.

In a sense, the original sin of correlation lies in its denominator, which standardizes the covariance only with respect to the value attained under perfect linear dependence. When perfect linear dependence cannot be achieved for the given marginals, the scaling becomes misleading. One alternative is to standardize $\mathrm{cov}(X, Y)$ using the extreme values (3) attained under comonotonicity and countermonotonicity (though these typically require integration and are not symmetric): this is the approach taken in Ai (2024). Also in multivariate settings, computing correlation between ranks remains a robust and practically safe alternative; see Puccetti (2022).

```
## Compute correlation between ranks
Y1<-rank(Y1) # ranks of Y1
Y2<-rank(Y2) # ranks of Y2
# Compute sample correlation between ranks (estimate Spearman's rank correlation)
cor(Y1,Y2)
```

```
## [1] 1
```

> **Key message: No linearity, no correlation: use rank correlation instead.**

Myth 4
"The VaR for the sum of Normal risks is maximal under maximal correlation"

Value-at-Risk (VaR) is one of the two risk measures used in banking and insurance regulation (the other is Expected Shortfall; see for instance section MAR33 in Basel Committee on Banking Supervision, 2026).

The *Value-at-Risk* of a loss random variable (risk) $X \sim F$, computed at the probability level $\alpha \in (0, 1)$, is the α-quantile of its distribution F, that is,

$$\mathrm{VaR}_\alpha(X) = F^{-1}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

If F is continuous and strictly increasing, then $q = \mathrm{VaR}_\alpha(X)$ is the unique real number such that $F(q) = \alpha$, so $\mathrm{VaR}_\alpha(X)$ is usually interpreted as the minimum amount of money that will be sufficient (under the prescribed model F) to cover the losses coming from X with probability α. In line with this interpretation, in risk management regulation the probability level α is typically chosen close to 1 (e.g., $\alpha = 0.95, 0.975, 0.99, \ldots$), depending on the typology of the underlying risk X.

In practice, one is often interested in the VaR of a portfolio, that is, of a sum of risks. This can be computed only after specifying (or at least modeling) how the risks depend on each other. In the special case where the risks are comonotonic, the VaR of a sum becomes particularly easy to calculate. For the proof of the following theorem (stated for simplicity for two risks, but holding analogously for the sum of an arbitrary number of random variables), we refer to Proposition 7.20 in McNeil et al. (2015).

Theorem 1.3. *Assume that $X \sim F$ and $Y \sim G$ are comonotonic. Then*

$$\mathrm{VaR}_\alpha(X + Y) = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y) = F^{-1}(\alpha) + G^{-1}(\alpha). \tag{4}$$

Since comonotonicity corresponds to the strongest form of positive dependence (and, when correlation exists, to the maximal attainable correlation), a common (and rather naive) conclusion is that (4) should provide the largest possible VaR for a sum of risks. Unfortunately, this is not true in general; see for instance the various illustrations given in Embrechts et al. (2014).

A closely related pitfall concerns the set-up of a risk management framework. Very often, the assumption that two risks are individually Normal, say $X \sim N(0, 1)$ and $Y \sim N(0, 1)$ ("we have Normal risks"), is confused with the much stronger assumption that the random vector $(X, Y)'$ is Normal ("we have a Normal vector"). Compared to the first assumption, the second one is stronger because it also specifies the type of dependence (namely, Gaussian dependence) between the risks.

If one adopts only the weaker assumption $X \sim N(0, 1)$ and $Y \sim N(0, 1)$, the maximum attainable VaR and at least one joint distribution attaining it can be computed analytically (see Makarov (1981), Rüschendorf (1982), and Wang and Wang (2011) for more than two random variables) or numerically via the *Rearrangement Algorithm* (Embrechts et al., 2013, Puccetti and Rüschendorf, 2012). Denoting by Φ the standard Normal distribution, one obtains that the worst VaR for the sum of two standard Normal random variables is equal to

$$\mathrm{VaR}_\alpha(X^* + Y^*) = 2\Phi^{-1}\left(\frac{\alpha + 1}{2}\right),$$

while the comonotonic VaR, according to (4), is smaller and given by

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

In Figure 4, left, we represent a bivariate distribution for $(X, Y)'$ which maximizes $\text{VaR}_\alpha(X + Y)$ at the level $\alpha = 0.90$. To maximize the α-quantile of a sum, only the largest $(1 - \alpha)$ part of the marginal probability mass matters. For this reason, it is not relevant how the smallest α probability mass is arranged on the left part of the figure, as long as the assumed marginals are respected. By contrast, the relevant $(1 - \alpha)$ probability mass (the one on the right part of the figure) is arranged so as to make the sum $X + Y$ as close as possible to a constant. This makes intuitive sense: the probability of exceeding the α-quantile is $(1 - \alpha)$. Concentrating more probability mass around (or exactly at) a single high-loss level helps pull the α-quantile further to the right.

As a result, the optimizer in Figure 4 combines positive dependence (on the left) with negative dependence (on the right). It is therefore not surprising that the implied correlation between X and Y is not maximal, i.e., it is smaller than one (and in this particular case of identical marginals, the maximal attainable correlation is indeed 1). Moreover, since the smallest α probability mass on the left can be spread in many different ways, one can also obtain even smaller values of correlation while keeping the same worst-case VaR value.
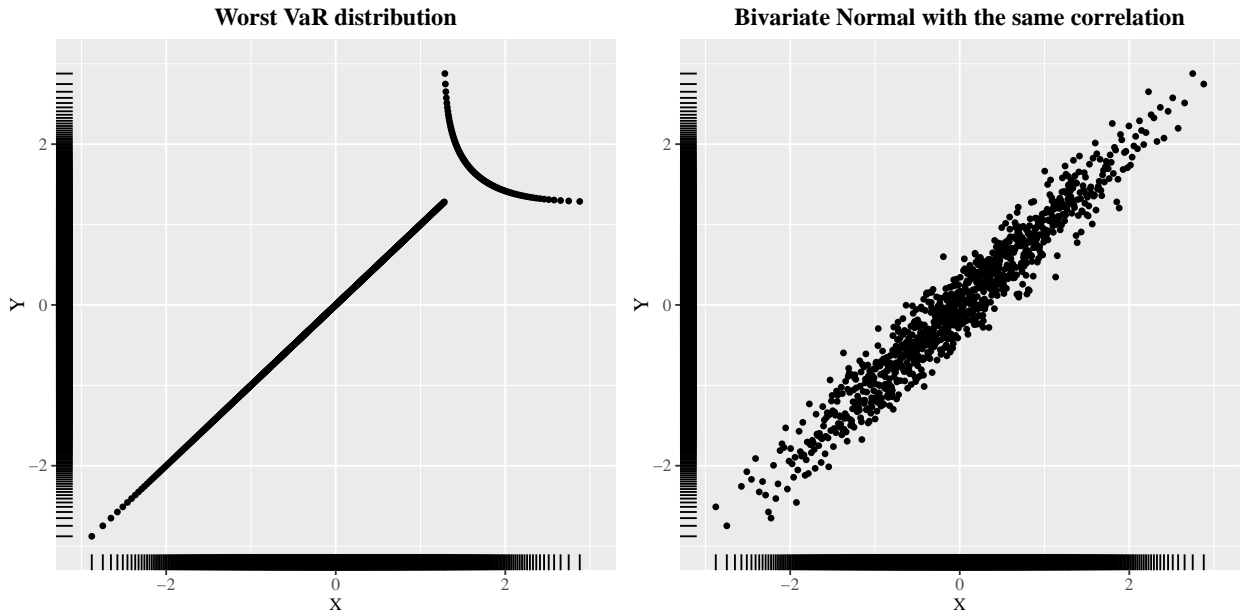

Figure 4: 2000 simulated points from two different bivariate distributions with the same standard Normal marginals and the same value of correlation. The left picture shows data from a distribution attaining the maximum VaR for the sum of its components at the probability level $\alpha = 0.90$; the right picture shows a bivariate Gaussian distribution with the same marginals and the same correlation between its components.

When presented at international conferences, the solution in Figure 4 is often criticized by practitioners as "too weird to be true": why should two risks start to move in opposite directions just beyond a high-loss threshold? From a mathematical point of view, however, maximizing VaR is an optimization problem, and Figure 4 simply displays a solution. To avoid such worst-case scenarios, one must impose stronger structural assumptions on dependence; see Puccetti et al. (2017).

The fact that Value at Risk (VaR) can be maximized by a combination of positive and negative dependence highlights one of its crucial weaknesses as a risk measure. Specifically, this behavior is related to VaR's lack of subadditivity, which means it is not a coherent risk measure—unlike Expected Shortfall; see, for instance, the classical references Artzner et al. (1999) and Klugman et al. (2019).

Figure 4 also illustrates another key point: specifying the marginals of X and Y and fixing the value of correlation $\rho(X, Y)$ is not sufficient to uniquely identify a joint distribution. This is exactly why it is so important to clearly state the dependence assumptions used in a risk management framework. Assuming that the vector $(X, Y)'$ (notice the brackets) has a bivariate Normal distribution (Figure 4, right) would immediately rule out the worst-case solution (Figure 4, left). However, even under this stronger assumption, an extra condition is needed to guarantee that the VaR of a sum is maximized under comonotonicity: the probability level α has to be larger than 0.5, as the following example shows.

Example 1.4. Suppose that the vector $(X, Y)'$ has a bivariate Normal distribution with standard Normal marginals Φ. If X and Y are assumed to be comonotonic, then correlation is maximal and equal to 1. In this case, using (4), it follows that

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

If X and Y are assumed to be independent, then correlation is null and $X + Y \sim N(0, 2)$ (under independence the variance of a sum is the sum of variances). Under independence then

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

Since $\Phi^{-1}(\alpha)$ is negative when $\alpha < 0.5$, one has that

$$\text{VaR}_\alpha(X^\perp + Y^\perp) > \text{VaR}_\alpha(X^C + Y^C) \ \text{ for all } \ \alpha < 0.5,$$

i.e., the VaR of a sum under independence is bigger than the VaR under comonotonicity.

There are other examples of "regular" assumptions for which the sum of the marginal VaRs (the value attained under comonotonicity) is not the worst possible VaR. Figure 5 shows that for the sum of two Pareto distributed risks with infinite mean, even independence is more dangerous than comonotonicity, for any value of $\alpha \in (0, 1)$; see Chen et al. (2025) for a formal proof.

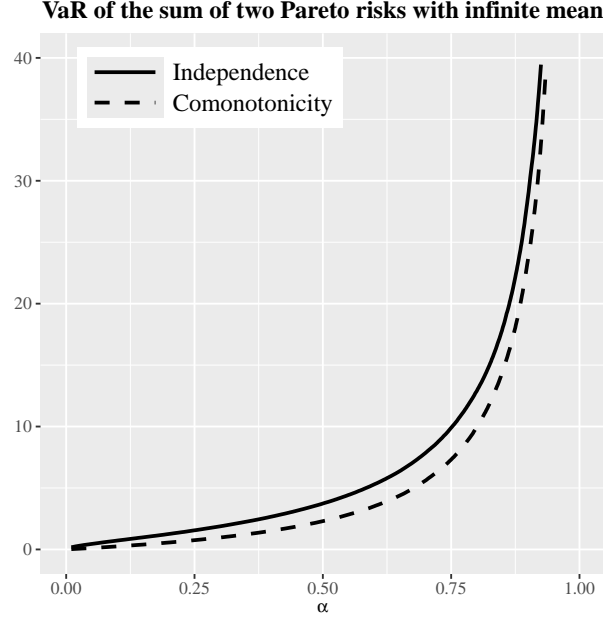> **Key message: In general, a higher correlation does not imply a higher VaR.**

VaR of the sum of two Pareto risks with infinite mean

Figure 5: VaR of the sum of two risks identically distributed like a Pareto(θ), i.e. $F(x) = 1 - (1 + x)^{-\theta}$, $x \geq 0$, with tail parameter $\theta = 0.9$ (an infinite-mean model), under the assumptions of independence (solid curve) and comonotonicity (dashed curve).

Myth 5
"We can freely stress-test correlation matrices"

For a multivariate random vector $(X_1, \ldots, X_d)'$, the *covariance matrix* $\Sigma = \sigma_{ij}$ is a square matrix collecting the covariances $\sigma_{ij} = \mathrm{cov}(X_i, X_j)$, $1 \leq i, j \leq d$, between every pair of components:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \ldots & \sigma_{dd} \end{pmatrix}.$$

The diagonal entries $\sigma_{ii} = \mathrm{cov}(X_i, X_i)$, $1 \leq i \leq d$, are simply the d marginal variances. When these variances are standardized to 1, Σ becomes a *correlation matrix* (and in this case $\sigma_{ij} = \rho(X_i, X_j)$).

It follows immediately from the definition that a covariance (and hence correlation) matrix must be symmetric. What is less obvious—and crucial in practice—is that not every symmetric matrix can arise as a covariance (or correlation) matrix. To be admissible, the matrix must satisfy an additional algebraic condition: it has to be positive semidefinite. For further details, including proofs of the results stated below, we refer to the classical reference Bhatia (2007).

Theorem 1.5. *A $d \times d$ symmetric matrix Σ is a covariance matrix if and only if it is positive semidefinite, that is, if and only if*

$$a' \Sigma a \geq 0, \quad for \ all \ a' = (a_1, \ldots, a_d) \in \mathbb{R}^d. \tag{5}$$

The meaning of this requirement is illustrated by the following example.

Example 1.6. Consider a random vector $X = (X_1, \ldots, X_d)'$ having finite covariance matrix Σ. Let the random vector $Y = a'X$ be a linear combination of X, with weights $a \in \mathbb{R}^d$. For example, if X represents daily joint log-returns of a portfolio of assets at a given trading desk, then

$$Y = a_1 X_1 + \cdots + a_d X_d$$

may be interpreted as a linearized profit/loss operator for that portfolio. The variance of Y can be computed as

$$\mathrm{var}(Y) = \mathrm{cov}(Y, Y) = a' \Sigma a.$$

Evidently, this variance must be non-negative for any choice of weights. Therefore,

$$a' \Sigma a \geq 0, \ \text{ for all } \ a' = (a_1, \ldots, a_d) \in \mathbb{R}^d,$$

which is precisely the requirement in (5).

As a consequence, care is needed whenever we manipulate covariance (or correlation) matrices. If the matrix associated with X fails to be positive semidefinite, then there exist weights $a_1^*, \ldots, a_d^*$ such that the variance of the linear combination $Y = a_1^* X_1 + \cdots + a_d^* X_d$ becomes negative: a mathematical impossibility. Any probabilistic model built on a matrix Σ that violates (5) is therefore internally inconsistent. Fortunately, there is a simple computational way to check whether (5) holds.

Theorem 1.7. *A symmetric matrix Σ with real entries is positive semidefinite if and only if all its eigenvalues are non-negative.*

Figure 6 shows a correlation matrix widely used in risk management, as given in the Annex IV, point (1), of the *Solvency II* directive for the calculation of the Solvency Capital Requirement (SCR) Standard Formula. Here, the random vector consists of $d = 5$ risks $(X_1, \ldots, X_5)'$, namely the market, counterparty default, life underwriting, health underwriting, and non-life underwriting risk modules.

One might be tempted to read the zero entries in the matrix as statements of independence (for instance, concluding that non-life and life risks are independent). Regardless of whether this was ever the regulator's intention, we stress again that, without further assumptions on the joint distribution of $(X_1, \ldots, X_5)'$, the only immediate conclusion is uncorrelatedness (leaving aside, as before, the issue of infinite moments). This is another reminder that a correlation matrix is a rather blunt tool outside the context of a well-specified probabilistic model, especially when dependence is not (assumed to be) approximately linear.

The Solvency Capital Requirement is meant to provide a 99.5% one-year probability level, i.e., roughly a "one-in-200-year" protection level under the prescribed model. In this sense, it can be interpreted as the VaR_α of the aggregate loss $X_1 + \cdots + X_5$ computed at the probability level $\alpha = 0.995$.

A common practice in risk management departments, often motivated by the desire to obtain more conservative capital requirements, is to "stress" the Solvency correlation matrix by increasing selected correlation values $\rho(X_i, X_j)$. The tendency to increase correlations is sometimes accompanied by a further misconception: that raising a few entries of the correlation matrix will necessarily increase the VaR of the aggregated position. As we saw in Myth 4, this implication does not generally hold.

Crucially, correlation matrices cannot be modified arbitrarily. Even if we preserve symmetry and keep all entries in a seemingly reasonable range, small adjustments may violate positive semidefiniteness, as the following example shows.

	Market	Default	Life	Health	Non–Life
Market	**1.00**	**0.25**	**0.25**	**0.25**	**0.25**
Default	**0.25**	**1.00**	**0.25**	**0.25**	**0.50**
Life	**0.25**	**0.25**	**1.00**	**0.25**	**0.00**
Health	**0.25**	**0.25**	**0.25**	**1.00**	**0.00**
Non–Life	**0.25**	**0.50**	**0.00**	**0.00**	**1.00**

Figure 6: Solvency correlation matrix for the calculation of the Solvency Capital Requirement (SCR) Standard Formula, which displays in each cell the correlation between two variables of the Solvency II regulation risk module vector composed by: market risk, counterparty default risk, life underwriting risk, health underwriting risk, and non-life underwriting risk modules.

Example 1.8. Suppose that, for stress-testing purposes, an insurance company changes the correlation between the health underwriting risk module and the market, default, and life underwriting risk modules from 0.25 to 0.7, with the naive aim of obtaining stricter Solvency Capital Requirements (whether this happens depends on the full model used). Before computing any new regulatory capital under the stressed specification, the company should first check whether the new matrix is positive semidefinite, for instance by computing its eigenvalues (e.g., via the R function eigen). If all eigenvalues are non-negative, then the matrix is positive semidefinite; see Theorem 1.7. If even one eigenvalue is negative, the model (and any implied regulatory capital) should be rejected as inconsistent.

	Market	Default	Life	Health	Non–Life
Market	**1.00**	**0.25**	**0.25**	**0.70**	**0.25**
Default	**0.25**	**1.00**	**0.25**	**0.70**	**0.50**
Life	**0.25**	**0.25**	**1.00**	**0.70**	**0.00**
Health	**0.70**	**0.70**	**0.70**	**1.00**	**0.00**
Non–Life	**0.25**	**0.50**	**0.00**	**0.00**	**1.00**

Figure 7: Solvency Capital Requirement correlation matrix - stress test 1.

Table 1: Eigenvalues - stress test 1.

Eigenvalue 1	Eigenvalue 2	Eigenvalue 3	Eigenvalue 4	Eigenvalue 5
2.564775	1.231365	0.75	0.5270761	-0.0732159

One eigenvalue of the correlation matrix in Figure 7 is negative. Consequently, the stressed matrix is no longer positive semidefinite. In Figure 8 we show additional stress tests for which the resulting matrices are not positive semidefinite: any probabilistic models built on such matrices are ill-specified.

The bad habit of not checking whether a matrix is positive semidefinite (that is, whether it has non-negative eigenvalues) likely comes from the fact that in a 2×2 correlation matrix the off-diagonal entry $\rho_{12} = \rho_{21}$ can indeed be any number between -1 and 1, and the matrix will remain positive semidefinite. Once we move to the 3×3 case, this is no longer guaranteed; the matrix

$$\begin{pmatrix} 1 & 0.17 & 0.62 \\ 0.17 & 1 & 0.88 \\ 0.62 & 0.88 & 1 \end{pmatrix}$$

look perfectly "reasonable" as a correlation matrix, but it is not positive semidefinite.

	Market	Default	Life	Health	Non–Life		Market	Default	Life	Health	Non–Life
Market	1.00	0.25	0.25	0.70	0.75	Market	1.00	0.80	0.25	0.80	0.75
Default	0.25	1.00	0.25	0.70	0.75	Default	0.80	1.00	0.25	0.70	0.75
Life	0.25	0.25	1.00	0.70	0.00	Life	0.25	0.25	1.00	0.70	0.00
Health	0.70	0.70	0.70	1.00	0.00	Health	0.80	0.70	0.70	1.00	0.00
Non–Life	0.75	0.75	0.00	0.00	1.00	Non–Life	0.75	0.75	0.00	0.00	1.00

Figure 8: Two stress-tested false correlation matrices.

There is a further subtlety. In many regulatory settings (including SCR), the marginal distributions are specified or strongly constrained. In that case, positive semidefiniteness is generally no longer sufficient to guarantee that a given matrix can actually be realized as a covariance (or correlation) matrix together with those specific marginals—unless one works within a special family (such as the multivariate Normal) where the covariance matrix characterizes the dependence structure. This is a notoriously difficult problem. Even for standard uniform marginals (i.e., for rank correlation matrices), feasibility appears to differ from positive semidefiniteness in dimension 12 or higher; see Wang et al. (2019). The situation is easier (and essentially solved) for Kendall's tau, another popular rank-based measure; see McNeil et al. (2022). In particular, switching from linear correlation to rank correlation does not make the matrix-feasibility issue disappear: one still has to enforce the correct structural constraints.

> **Key message: Check the eigenvalues of your correlation matrix before use!**

Final discussion

This paper revisits a number of myths and common misconceptions about dependence—and does so in what we hope is a somewhat fresh way: without invoking copulas, and relying instead on elementary statistical ideas. Our main focus is a topic of enduring interest in the financial risk management community: how to think about, interpret, and use correlation.

In practice, there are often sizable gaps between the formal definition of correlation and the way "correlated risks" are discussed in day-to-day work. We therefore aim for a presentation that is accessible to a broad readership, while still offering concrete takeaways even for readers who are not specialized in financial risk management.

A major source of confusion is that the classical bounds -1 and 1 are not always attainable once the marginal distributions of the risks are fixed. In other words, even when correlation exists, its effective range may be a strict subinterval of $[-1, 1]$, and in extreme cases that range can shrink dramatically.

Ideally, one would like a dependence measure that equals -1 if and only if the random variables are countermonotonic, equals 1 if and only if they are comonotonic, and equals 0 if and only if they are independent (note that even rank correlation fails to satisfy the last "only if"). Unfortunately, as shown in Scarsini (1984) and in Section 4.2 of Embrechts et al. (2002), no senseful dependence measure can satisfy all these desiderata simultaneously.

Rank correlation offers a partial remedy for Pearson correlation's scaling problem, because it restores the full range $[-1, 1]$. In particular, rank correlation equal to 1 (-1, respectively) corresponds to perfect positive (negative) dependence in the sense of comonotonicity (countermonotonicity). At least, a rank correlation strictly smaller than 1 (strictly larger than -1) tells us that the variables are not perfectly positively (negatively) dependent. This should not be misunderstood as a panacea. For instance, a rank correlation of zero does not imply independence (one can already see this in simple examples; for instance, try computing rank correlation in the example given in Myth 2).

As a final takeaway, the use of Pearson correlation should be restricted to settings where it is genuinely informative—most notably, within families of distributions that are closed under affine transformations, such as the multivariate Gaussian or multivariate t models. Outside such strong modeling assumptions, it is generally safer to work with rank-based measures of association.

References

Ai, X. (2024). Adjust Pearson's r to measure arbitrary monotone dependence. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, 9(3):203–228.

Basel Committee on Banking Supervision (2004). *International Convergence of Capital Measurement and Capital Standards*. Bank for International Settlements, Basel. Available at https://www.bis.org/publ/bcbs107.pdf.

Basel Committee on Banking Supervision (2026). *The Basel Framework*. Bank for International Settlements, Basel. Available at https://www.bis.org/basel_framework/.

Bhatia, R. (2007). *Positive Definite Matrices*. Princeton University Press.

Casella, G. and Berger, R. L. (2002). *Statistical inference*. Duxbury Pacific Grove, CA, second edition.

Chen, Y., Embrechts, P., and Wang, R. (2025). Technical note–An unexpected stochastic dominance: Pareto distributions, dependence, and diversification. Available at `https://doi.org/10.1287/opre.2022.0505`.

Cirillo, P. (2013). Are your data really Pareto distributed? *Physica A: Statistical Mechanics and its Applications*, 392(23):5947–5962.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., and Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance Math. Econom.*, 31(1):3–33.

Di Lascio, F., Giammusso, D., and Puccetti, G. (2018). A clustering approach and a rule of thumb for risk aggregation. *Journal of Banking & Finance*, 96:236–248.

Embrechts, P. and Hofert, M. (2013). A note on generalized inverses. *Math. Methods Oper. Res.*, 77(3):423–432.

Embrechts, P., McNeil, A. J., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk management: Value at Risk and Beyond*, pages 176–223. Cambridge Univ. Press, Cambridge.

Embrechts, P., Puccetti, G., and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *J. Bank. Financ.*, 37(8):2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R., and Beleraj, A. (2014). An academic response to Basel 3.5. *Risks*, 2(1):25–48.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon. Sect. A. (3)*, 14:53–77.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1934). *Inequalities*. Cambridge, at the University Press.

Hoeffding, W. (1940). Massstabinvariante Korrelationtheorie. *Schriften Math. Inst. Univ. Berlin*, 5(5):181–233.

Hoeffding, W. (1994). Scale–invariant correlation theory. In Fisher, N. and Sen, P., editors, *The Collected Works of Wassily Hoeffding*, pages 57–107. Springer New York. This is the English translation of Hoeffding (1940).

Hofert, M., Kojadinovic, I., Mächler, M., and Yan, J. (2018). *Elements of Copula Modeling with R*. Springer, Cham.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2019). *Loss Models: From Data to Decisions*. John Wiley & Sons, Hoboken NJ, fifth edition.

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

Makarov, G. D. (1981). Estimates for the distribution function of the sum of two random variables with given marginal distributions. *Theory Probab. Appl.*, 26:803–806.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management. Concepts, Techniques and Tools. Revised edition.* Princeton University Press.

McNeil, A. J., Nešlehová, J. G., and Smith, A. D. (2022). On attainability of Kendall's tau matrices and concordance signatures. *Journal of Multivariate Analysis*, 191:105033.

Puccetti, G. (2022). Measuring linear correlation between random vectors. *Information Sciences*, 607:1328–1347.

Puccetti, G. and Rüschendorf, L. (2012). Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.*, 236(7):1833–1840.

Puccetti, G., Rüschendorf, L., Small, D., and Vanduffel, S. (2017). Reduction of value-at-risk bounds via independence and variance information. *Scandinavian Actuarial Journal*, 2017(3):245–266.

Puccetti, G. and Scherer, M. (2018). Copulas, credit portfolios, and the broken heart syndrome. *Dependence Modeling*, 6(1):114–130.

Puccetti, G. and Wang, R. (2015). Extremal dependence concepts. *Stat. Science*, 30(4):485–517.

Rüschendorf, L. (1982). Random variables with maximum sums. *Adv. in Appl. Probab.*, 14(3):623–632.

Salmon, F. (2012). The formula that killed Wall Street. *Significance*, 9(1):16–20.

Scarsini, M. (1984). On measures of concordance. *Stochastica*, 8(3):201–218.

Wang, B. and Wang, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.*, 102(10):1344–1360.

Wang, B., Wang, R., and Wang, Y. (2019). Compatible matrices of Spearman's rank correlation. *Statistics & Probability Letters*, 151:67–72.