EDUCATION COMMITTEE

OF THE

SOCIETY OF ACTUARIES

ENTERPRISE RISK MANAGEMENT STUDY NOTE

DEVELOPMENTS IN MODELLING RISK AGGREGATION

Joint Forum

Basel Committee on Banking Supervision

ERM-103-12 Printed in the USA

Annex G

Technical underpinnings of aggregation methods

This annex discusses the technical underpinnings of the most popular aggregation techniques used for the purposes of capital adequacy or solvency assessment, capital allocation and risk pricing. Aggregation methods have typically coalesced around a standard tool box. The popularity of these tools and techniques can be explained in part by their distinct advantages, which include, in different cases, computational convenience, flexibility, and ease of interpretation. Once they obtain a certain acceptance and use in financial practice, the popularity of these tools and techniques may be self-reinforcing and persistent.

A potential pitfall of popularity is that the techniques may develop some measure of authority mainly on the grounds of their widespread use and familiarity. Many users may question insufficiently their appropriateness for a given application, or may not thoroughly consider their limitations. This section will summarise the technical concepts involved in these methods and provide background into the implications, requirements, and limitations of using them. The three techniques discussed are the variance-covariance approach (and its simpler variants such as simple summation), copula-based simulation, and scenario-based simulation.

G.1 VarCovar approach

The variance-covariance (VarCovar) approach is a convenient and commonly used analytical technique that allows managers to combine marginal (ie, "standalone") distributions of losses, or distinct tail losses[48] directly (for capital requirements), into a single aggregate loss distribution or tail loss estimate. The sole requirement is to characterise the level of interdependence of standalone losses, which is typically accomplished with a matrix of linear correlations. Some organisations apply VarCovar at lower levels of risk aggregation, eg, to aggregate market price risk in the trading book. A common use, and one that will be the focus of this section, is to apply VarCovar at the top level of risk aggregation, where fundamental drivers used to model lower-level risks often cannot easily be combined. At a banking organisation, for example, VarCovar may be used to aggregate losses from "trading book market risk," "banking book credit risk," "operational risk," and so on. The main advantages of VarCovar are that it uses a limited number of inputs, can be evaluated formulaically, and does not require fundamental information about lower-level risks.

Statistical foundation of VarCovar

Some of the simplest top-level risk aggregation practices observed are special cases of VarCovar, although their VarCovar foundations may not always be acknowledged. These include the practice of estimating total capital requirements as the sum of lower-level capital requirements, as well as the practice of taking the square root of the sum of squared lower-

[48] In this discussion, "tail risk" is used as a generic term for any specific tail risk measure – eg, 99% VaR, 95% VaR, expected shortfall at the 99% tail – that is assumed to be evaluated in a consistent way at a fixed confidence level for both the aggregate risk and individual lower-level risks. Most of this discussion does not depend on the specific measure used.

level capital requirements (implying an assumption of independent lower-level risks). Any use of VarCovar presumes certain characteristics about the underlying loss distributions. Whether these accurately describe the actual loss distribution in itself underpins the validity and implications of applying the approach. For a practical example, if these conditions do not hold, then calculating aggregate risk level as the sum of lower-level risks, commonly interpreted as a ceiling on aggregated risk, does not mean that one fully considers the potential interaction between the lower-level risks or that the aggregate risk is quantified at the chosen theoretical confidence level. While VarCovar is a simple and highly tractable approach to risk aggregation, the cost to the unwary user is that it effectively fills in unspecified details about the nature of the loss distributions, which may or may not be accurate or intended.[49]

An expression[50] for aggregate risk under VarCovar is as follows:

$$R = \lambda \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \operatorname{cov}(i,j)} = \sqrt{\sum_{i=1}^{N} w_i^2 r_i^2 + 2\sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} w_i w_j r_i r_j corr(i,j)}$$

where *R* is the aggregate risk or capital requirement, *r* are the lower-level risks which compose the aggregate risk (evaluated at a fixed confidence level), cov(*i,j*) is the covariance between variables *i* and *j*, corr(*i,j*) is their correlation, and w(*i*) are concentration weights for the lower-level risk sources (equal to 1 if lower-level risk is already scaled in the end units). The role of λ will be detailed below. The covariance between two variables is equivalent to the product of the correlation coefficient, the standard deviation of the first variable in the pair, and the standard deviation of the second. The variance of each variable (the square of the standard deviation) is found on the diagonal on the covariance matrix and is equivalent to the covariance of the variable with itself. Assuming the weights are 1, the formula for *R* is expressed in matrix notation as $\sqrt{(r'Cr)}$, where r is the vector of lower-level risks and C is the correlation matrix. R can be any tail risk measure consistent with $\alpha.f(g,h,..;C) = f(\alpha.g(..),\alpha.h(..),..;C)$, where f(..) is the aggregate tail risk corresponding to lower-level tail risks g(..), h(..), and so on; and correlation matrix C (this property corresponds to the Positive Homogeneity for a coherent risk measure (see Box A - Coherent Risk Measures)). An increase in all lower-level tail risks (eg, 99% VaR) by a fixed proportion must increase aggregate VaR by the same proportion. A canonical example of where this holds is where lower-level risks are assumed to be normally distributed. Empirical distributions of aggregate and lower-level losses may be strained to meet this requirement.

A statistical foundation of the variance-covariance approach is that the mean and variance of a real variable are known if the variable can be expressed as a linear combination of other variables whose means, variances, and covariances are defined and known. Direct substituting these relationships into the VarCovar formulation, R divided by λ can be assumed to represent the standard deviation of an aggregate loss distribution so long as each $r(i)/\lambda$ represents the standard deviation of the "i^{th}" lower-level loss distribution, the correlation matrix contains the true linear correlation coefficients (more formally, the Pearson product moment correlations) between any two lower-level losses, and (as customary in calculating capital requirements) the expected loss in each distribution is assumed to be zero. λ is the ratio of the tail risk value to the standard deviation; this is specific to the shape

---

[49] J McNeil, R Frey, and P Embrechts, Quantitative Risk Management by A Princeton, 2005.

[50] Rosenberg, Joshua V and Schuermann, Til, A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risk. Journal of Financial Economics, FRB of New York Staff Report No. 185, SSRN - http://ssrn.com/abstract=880422.

of the loss distribution and the choice of risk measure (eg, 99% VaR), but must be jointly applicable to both lower-level and aggregate risks. This is again consistent with the scaling property described earlier, but has no guarantee of being an empirical reality.

Perfect linear dependence, independence and diversification

The assumed correlation matrix in effect controls the level of diversification recognised by the enterprise across the lower-level risks using VarCovar. The lower the correlations on the non-diagonal elements of the matrix (diagonal elements must be equal to 1), the greater the level of diversification that can be realised with incremental (long) exposure to a risk component. In addition, simply for VarCovar to be evaluated, the matrix must satisfy numerical constraints which are explained in Box C discussing the Positive Semi-Definiteness Assumption. The simple cases of VarCovar that were described earlier are a consequence of different assumptions for the correlation matrix: for example, assuming a matrix of 1s is identical to simply summing the lower-level risks to produce aggregate risk, while applying the identity matrix (1s on the diagonal, 0s elsewhere) is equivalent to calculating aggregate risk as the square root of the sum of squared lower-level risks; the former represents an assumption of perfect linear correlation and the latter an assumption of linear independence. In both cases, however, if the aggregation distribution does not scale as inherently assumed, neither will represent what is intended.

The correlations within the VarCovar

Practitioners usually interpret the elements of the correlation matrix used in VarCovar as the linear correlations between any given pair of variables, although strictly speaking, the factors applied in the VarCovar need not represent these as long as they follow the usual numerical constraints on correlations. In any case, for most multi-variable distributions, the correlation matrix (containing a single number for each distinct pair of variables) is not sufficient to determine all the ways that two variables can interact. That appears to be the case only for members of the family of so-called elliptical distributions, which includes the normal or Gaussian. Otherwise, it provides only partial information about dependence, and more information would be needed to describe the full dependence structure (see Box B – Correlations vs Dependencies). Since capital requirements are concerned with improbable outcomes, joint behaviour when losses are significant is more important than the correlation coefficient measured over the entire range of outcomes, good and bad, for those variables, which represents at best a sort of average of conditional linear dependencies.

Risk managers may try to overcome weaknesses in applying correlations by substituting a pseudo-correlation matrix, such as a "stressed" or "tail correlation" measure, which may be derived independently or as an adjustment on historical correlations, as well as by taking ad hoc adjustments to the risk measure under the previously mentioned simple-sum ("perfect correlation") assumption. One drawback of using a subjective or judgment-based assessment is that it may be calibrated to match a desired overall outcome rather than receiving an appropriate level of independent justification. More mechanical means of estimating tail dependence between risk sources, including the use of local approximation for tail correlation matrices that can be justified under certain hypothetical circumstances, reduce the level of subjective judgment required.

An alternative to linear correlation matrices involves the use of rank correlation measures independently of assumed marginal distributions, possibly to accommodate a more conservative joint distribution or tail correlation matrix. A given rank correlation matrix can be applied to an unlimited choice of specific standalone distributions for the underlying variables. In particular, any system of "fat tailed" marginal distributions can be combined as a jointly fat-tailed multivariable distribution using their rank correlations, including those derived from limited or thinner-tailed data. Such a distribution may be capable of producing more severe and realistic examples of joint behaviour under stress than one could produce by entering their linear correlations into VarCovar. Rank variables, which are used to calculate Spearman's rank correlation, have much of the same properties as the real variables from which they are derived and can be manipulated similarly. However, since rank transformations do not preserve the assumptions required of VarCovar (such as uniform risk scaling), the calculation of an aggregate risk measure using rank information is better suited to simulation via copula functions, which will be discussed in the next section.

Still another possibility, particularly to overcome lack of data, is the use of factor decomposition of lower-level risks to determine the correlation between them[51]. Factor models estimate potential changes in the value of a risky asset based on its factor sensitivities to available risk factors and an idiosyncratic (residual) component. In a pure factor model, the risk factors are orthogonal, and the idiosyncratic component consists of independent, Gaussian draws. The covariance of returns across any two assets is determined by their individual sensitivities to the common factors and their common factor variances (their correlations are dependent on those things as well as the variances of their residuals). Factor structures can be estimated using regression or other numerical techniques, and adjusted in specific ways to engineer a suitable correlation matrix.

VarCovar and other top-down aggregation tools (including copulas) also face difficulty in dealing with circumstances in which "standalone" risks are not actually exclusive but are believed to be integrated. This is, for instance, the case for banking market risk and credit risk, which, while often calculated separately, may originate from the same portfolios, same underlying events, or the very same entities. Integrating such risks is still a frontier issue in risk aggregation, and involves interplay of continuous and somewhat discontinuous risk factors that may not lend themselves to the smooth assumptions of top-down approaches, particularly constant linear correlations.

Conclusions

It is important to note, in conclusion, that in nearly all cases where it is applied for risk management, the VarCovar is an imposition of simple dependency structure on what is believed to be a more complex web of dependencies. Almost all empirical dependencies involve a huge amount of information and are not readily reduced to a single number per distinct pair of variables. Copulas, by contrast, are capable of specifying a full dependence structure, with minimal requirements on what the distributions must actually look like. While copulas can be made as flexible as the user requires, the results of VarCovar are most akin to those of copulas simulations on the joint behaviour of known elliptical multivariable distributions such as the normal/Gaussian. Similarly, pure factor-models, whose correlations may feed VarCovar formulations, are Gaussian in foundation, though they can be extended. These limits of the VarCovar, which are inextricably linked to aforementioned VarCovar constraints on distributions of the standalone risks, can lead to deeply misleading results if those inherent assumptions do not coincide with the experience or intention of the risk manager.

---

[51] Meucci, Attilio, *Risk and Asset Allocation,* 2007, Chapter 3.4.

G.2 Distribution-based aggregation

In contrast to the VarCovar approach, the copula-based methods described in this section use entire loss distributions as inputs to the aggregation process, as opposed to single statistics or risk measures. These allow direct control over the distributional and dependency assumptions used, and make it possible to impose a wide variety of dependency structures on the aggregated distributions. Most of the methods in this category are analytically complex, and do not lend themselves to implementation with closed-form formulas. As a consequence, these methods almost always involve simulation when used in applications.

Definition

A copula, in simplest terms, can be viewed as a random vector (ie a multivariate distribution) whose individual components (ie marginal distributions) are all random variables that are uniformly distributed on the interval [0,1]. The copula-based approach can be used to describe any multivariate distribution as a set of marginal distributions together with a copula. The copula specifies the dependency structure among the individual random variables, and is used to join the marginal distributions together. Sklar's Theorem (1959) states that any multivariate distribution is uniquely determined by its marginal distributions and a copula, and that any combination of marginal distributions with a copula gives rise to a valid multivariate distribution.

This decomposition of multivariate distributions into marginal distributions and a copula allows practitioners to match any set of individual distributions to a specified dependence structure using a bottom-up approach. For a given set of random variables, different dependency structures can be imposed on the variables by specifying different copulas. Conversely, given a specific copula, random variables having various types of distributions can be joined together using the copula to produce multivariate distributions having different marginal distributions but similar dependency structures.

How copulas are used for risk aggregation

Copula techniques depend on the following property that relates a random variable to its distribution function: If X is any continuous random variable and F_X is the distribution function of X, then $F_X(X)$ is distributed uniformly on the interval [0,1]. One consequence of this is that if U is a random variable that is uniformly distributed on the interval [0,1], the random variable $F_X^{-1}(U)$ (this is simply the U-th percentile of the random variable X) has the same distribution as X. This property can be used to simulate X by drawing random samples from a uniform [0,1] distribution and then evaluating the corresponding percentiles of X, given by the function F_X^{-1}, at the sampled points.

In practice, an entity will have several loss distributions, corresponding to different types of losses, that it wishes to aggregate. If we assume that $X_1,\ldots,X_n$ are random variables (not necessarily identically distributed) for n different loss types, whose distribution functions are $F_{X_q},\ldots,F_{X_n}$ respectively, then the procedure for sampling from the aggregate loss distribution using a copula is as follows:

1. Draw a joint sample of uniform random variables $(\tilde{u}_1,\ldots,\tilde{u}_n)$ from the distribution specified by the copula.

2. Translate the sample from the copula distribution into a sample from the conjoined loss distribution by calculating the $\tilde{u}_1$-th percentile of X_1, the $\tilde{u}_2$-th percentile of X_2, etc. (in vector form, this is $\left(F_{X_q}^{-1}(\tilde{u}_1),\ldots,F_{X_n}^{-1}(\tilde{u}_n)\right)$)

3. Calculate the realised sample for the aggregate loss as the sum of the percentiles drawn from each distribution (ie $F_{X_q}^{-1}(\tilde{u}_1)+\cdots+F_{X_n}^{-1}(\tilde{u}_n)$).

4. Drawing many samples for the aggregate loss distribution will produce a simulated distribution. Any measure of risk (such as VaR or expected shortfall) can be computed from this simulated distribution.

Step 1) of the above process involves simulating a multivariate distribution, while step 2) only involves simulating single-variable distributions successively. Thus, step 1) is usually the hardest part of the process, although many copulas can be simulated without much difficulty.

Example: A company wishes to aggregate two loss distributions: a lognormal distribution that is the exponent of a normal distribution with mean 2 and standard deviation 1, and an exponential distribution having mean 12. The company uses a two-dimensional copula that generates the joint samples shown in columns 2 and 3 of the table below. The uniform sample in column 2 is translated as a percentile into a sample loss from the lognormal distribution in column 4, while the uniform sample in column 3 is translated into a sample loss from the exponential distribution in column 5. The corresponding samples from the aggregate loss distribution are shown in column 6.

Sample Number (1)	Copula Sample (first component) (2)	Copula Sample (second component) (3)	Lognormal Distribution Sample (4)	Exponential Distribution Sample (5)	Aggregate Loss Sample (6) = (5) + (4)
1	82.3%	40.6%	-2.9	-10.8	-13.7
2	50.3%	79.8%	-7.3	-2.7	-10.0
3	9.3%	18.5%	-27.7	-20.3	-48.0
4	66.6%	25.5%	-4.8	-16.4	-21.2
5	28.4%	61.7%	-13.1	-5.8	-18.9
6	42.1%	44.4%	-9.0	-9.7	-18.8
7	60.9%	98.6%	-5.6	-0.2	-5.8
8	30.6%	10.2%	-12.3	-27.3	-39.6
9	97.6%	56.5%	-1.0	-6.9	-7.9
10	42.4%	54.0%	-8.9	-7.4	-16.3
11	41.0%	2.9%	-9.3	-42.3	-51.6
12	14.6%	22.8%	-21.2	-17.8	-38.9
13	91.5%	40.1%	-1.9	-11.0	-12.8
14	38.4%	93.6%	-9.9	-0.8	-10.7
15	55.0%	70.4%	-6.5	-4.2	-10.7
16	6.4%	27.2%	-33.8	-15.6	-49.5
17	63.1%	70.7%	-5.3	-4.2	-9.4
18	8.0%	72.9%	-30.2	-3.8	-34.0
19	32.1%	20.6%	-11.8	-18.9	-30.7
20	21.8%	55.1%	-16.1	-7.2	-23.3

The next subsection gives a short overview of some of the copula functions used by the financial industry to aggregate risks.

Distribution functions of copulas

A copula, being a multivariate distribution, can be specified completely by its distribution function, and copulas are most often analysed in terms of their distribution functions. Since all of the components of a copula range over the interval [0,1], a copula can be described as a function *C* mapping the Euclidean cube $[0,1]^n$ to the interval [0,1]. This function must satisfy all of the conditions that a multivariate distribution function must satisfy (ie non-decreasing in each component, right continuity, limits of 0 and 1, rectangle inequality). In addition, since all of the marginal distributions must be uniform, *C* must satisfy the condition that, for all arguments of the function and all *u* in the interval [0,1]:

$$C(1,\ldots,1,u,1,\ldots,1) = u$$

Any function meeting all of these conditions corresponds to a unique copula[52].

Copulas from known distributions

One simple way to generate copula distribution functions is from known multivariate distributions. Given any multivariate distribution function **F** having marginal distribution functions $F_1,\ldots,F_n$, the function:

$$C(u_1,\ldots,u_n) = \mathbf{F}\left(F_1^{-1}(u_1),\ldots,F_n^{-1}(u_n)\right)$$

defines a copula. The widely-used Gaussian copula is defined in this manner: if $\boldsymbol{\Sigma}$ is a positive semi-definite correlation matrix and $\boldsymbol{\Phi}_{\boldsymbol{\Sigma}}$ is the standardised multivariate normal distribution function having correlation matrix $\boldsymbol{\Sigma}$, the distribution function for the Gaussian copula is given by:

$$C(u_1,\ldots,u_n) = \boldsymbol{\Phi}_\Sigma\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_n)\right)$$

where Φ is the standardised (univariate) normal distribution function. This copula is easy to simulate because the underlying multivariate normal distribution with correlation matrix $\boldsymbol{\Sigma}$ is easy to simulate: if $(x_1,\ldots,x_n)$ is a sample from the correlated multivariate normal distribution, then $(\Phi(x_1),\ldots,\Phi(x_n))$ is a sample from the corresponding Gaussian copula. This technique is often used with more general multivariate distributions for which the correlation matrix is a key parameter, such as the class of elliptic distributions.

Archimedean copulas

Another technique for generating copulas is to directly construct functions that meet all of the requirements to be a distribution function for a multivariate random variable. One example of such a construction is the class of Archimedean copulas, defined by:

$$C(u_1,\ldots,u_n) = \varphi^{-1}(\varphi(u_1)+\cdots+\varphi(u_n))$$

---

[52] We refer to Nelson, Roger, *An introduction to Copulas*, Springer, 2006 for a more detailed discussion on distribution functions of copulas

where $\varphi:[0,1]\to(0,\infty)$ is a strictly decreasing, surjective, infinitely differentiable convex function. Examples of Archimedean copulas include the Gumbel copula, generated by the function $\varphi(x)=(-\ln x)^{\alpha}$ for $\alpha\geq 1$, the Clayton copula, generated by the function $\varphi(x)=(x^{-\theta}-1)/\theta$ for $\theta>0$, and the Frank copula, generated by $\varphi(x)=\ln\left((e^{\alpha x}-1)/(e^{\alpha}-1)\right)$. In contrast to the Gaussian copula, the Archimedean copulas have distribution functions that can be simply described in closed form. However, unlike the Gaussian copula, it is often necessary to use advanced techniques (such as Laplace transforms) in order to simulate Archimedean copulas.

While it is possible to generate a large range of Archimedean copulas through various choices of generator φ, all of the Archimedean copulas created using the above formula have the disadvantage of being highly symmetric. Specifically, if one exchanges any two of the arguments in the distribution function, the function will remain unchanged. This symmetry limits the use of these copulas to aggregating risks that are uniform and interact in the same manner, such as credit portfolios of homogeneous risks. They cannot be used to model asymmetric behaviour, which is quite commonly observed within risks (in bad times, there are more adverse risk outcomes observed than there are beneficial outcomes observed during good times). By contrast, the Gaussian copula will not have this symmetry property unless all off-diagonal elements of the correlation matrix are the same. There have been many successful attempts in the research literature to generalise the class of Archimedean copulas to include copulas that are asymmetric.

Measures of dependence for copulas

When aggregating risk exposures, the issue of dependence is extremely important. For example, a practitioner may wish to have a model that reproduces the phenomenon observed in the real world that, during stress periods, risks tend to materialise at the same time. This dependence is a crucial determinant of the shape of the distribution and the computed risks. Under the copula approach, the entire dependence structure between a set of random variables is encapsulated in the choice of copula. Thus, any desired dependence structure can be specified through the choice of copula, which can then be used to aggregate any set of marginal loss distributions.

Relation between size and dependence

The magnitude of a copula distribution function, or alternatively, the rate at which it increases from 0 to 1, serves as an indicator of the level of dependence it imposes among the distributions that it aggregates. For a copula in two variables, the lowest possible value of a distribution function is given by the Fréchet lower bound:

$$C(u_1,u_2)=\max(0,u_1+u_2-1)$$

This copula implies perfect negative dependence between the aggregated random variables, ie that one variable is a decreasing function of the other. Under perfect negative dependence, when one of the aggregated variables is at a high percentile in its range, the other variable will be at a correspondingly low percentile.

The highest possible value of the distribution function is given by the Fréchet upper bound:

$$C(u_1,u_2)=\min(u_1,u_2)$$

and implies perfect positive dependence, namely that one variable is an increasing function of the other. Under perfect positive dependence, when one of the aggregated variables is at a high percentile, the other variable will be at a similarly high percentile.

In between these lower and upper bounds lies the product copula given by:

$$C(u_1, u_2) = u_1 u_2$$

which implies that the two aggregated variables will be independent. Similar concepts of positive dependence and independence can be extended to copulas in dimensions higher than 2.

Correlations

When a set of distributions is joined by a copula, the standard (Pearson) correlation matrix of the resulting multivariate distribution will vary with the marginal distributions that are input to the copula. Consequently, it is difficult to use the standard correlation measure when working with copulas, as the effect of the marginal distributions will be confounded with the properties of the copula. If copula parameters are fit based on the standard correlations observed for a particular set of marginal distributions, the parameters are likely to lead to invalid results when the copula is used to aggregate a different set of marginal distributions. This type of error is often made in practice, and may severely reduce the reliability of the aggregation measure if not corrected.

In order to avoid having a model's results rendered invalid by the effect of the marginal distributions on the standard correlation, it is necessary to use measures of correlation that depend only on the copula itself. This need is met by measures of rank correlation, specifically the Spearman rho and Kendall tau correlation coefficients [see Box B – Correlations vs Dependencies]. These have the property that they are invariant under increasing functions, because they depend only on the relative rank of an observation within a data set rather than the actual value of the observation. This implies that the measures will be the same for all multivariate distributions having the same copula, and that they can be calculated directly from the copula distribution function. As an example, the matrix of Spearman rho correlations for a copula is simply the Pearson correlation matrix of the copula's uniform marginal distributions.

These measures find their greatest use in the calibration of copula parameters, because the rank correlation measures from observed data can be used to calibrate a copula directly, without having to make any a priori assumptions about the marginal distributions of the observations.

Tail dependence

It has been observed that large losses, either from different risk types or within the same risk type, tend to strike simultaneously during stress situations, and practitioners often wish to capture this phenomenon in their copula models. This concept can be formalised through the definition of tail dependence.

Given a copula *C* in two variables (or alternatively, a two-variable marginal copula taken from a higher-dimensional copula), a stress situation will correspond to one or both of the variables taking values close to zero; this will translate into large losses when the copula values are fed into the inverse distribution functions of the loss random variables. If we know that one of the copula variables has taken on a small value, this indicates that a stress scenario may be underway, and that the other copula variable is more likely to take on a small value than it usually would be. Mathematically, this means that if U_1 and U_2 are the two uniform copula variables and v is a value close to zero, the conditional probability:

$$\Pr(U_1 \le v \mid U_2 \le v)$$

will be higher than v, which is the unconditional probability. Since this conditional probability can be expressed as:

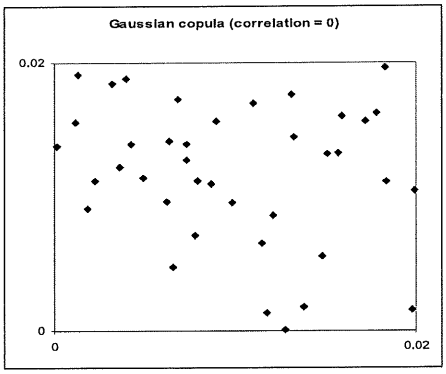
$$\frac{\Pr(U_1 \le v \text{ and } U_2 \le v)}{\Pr(U_2 \le v)} = \frac{C(v,v)}{v}$$

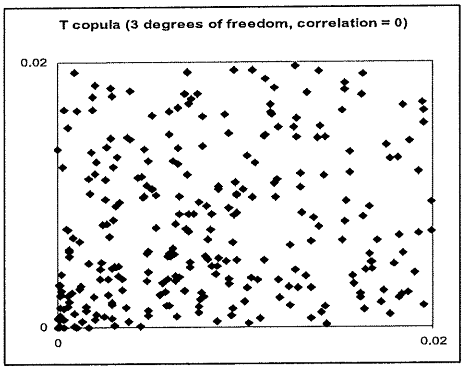
the coefficient of lower tail dependence for the copula is defined to be:

$$\lim_{v \to 0} \frac{C(v,v)}{v}$$

and the copula is said to exhibit lower tail dependence if this limit is greater than zero.

The Gaussian copula does not exhibit any tail dependence between pairs of its variables, even if the correlation matrix used in the copula is different from the identity matrix. This is considered to be a major drawback of the Gaussian copula that limits its use in applications. However, the more general class of elliptic copulas does contain copulas having tail dependence. In particular, copulas derived from the multivariate *t* distribution are often chosen in order to incorporate an explicit degree of tail dependence into the aggregate risk distribution.





The two charts represent the 98% tail regions for a Gaussian copula and for a t copula with 3 degrees of freedom. Both copulas were simulated assuming zero correlation between the two uniform marginal distributions, and both copulas had 100,000 points sampled. Because of tail dependence, the t copula has almost seven times as many observations falling within the tail region, and clusters noticeably at the lower left corner.

Conclusions

The following table summarises the properties of particular copulas that have been mentioned above:

Copula type:	**Gaussian**	**t**	**Archimedean**
Ease of simulation	Easy	Easy	Difficult
Capable of modelling tail dependence?	No	Yes	Yes
Symmetry	Symmetric in 2 dimensions, but generally asymmetric in higher dimensions		Standard construction is symmetric

The copula approach is well suited for use in aggregating financial risks because it works directly with the percentile measures of the loss distributions (this is what the uniform marginals of a copula represent conceptually). Since virtually all risk measures are based on percentiles or levels of confidence, the copula approach allows the practitioner to precisely specify the dependencies in the areas of the loss distributions that are crucial in determining the level of risk. Another advantage of copulas is that they are usually easy to implement from a computational standpoint; one side benefit is that simulated losses can be stored and used for applications beyond aggregate loss modelling.

However, the specification of a copula is very abstract and difficult to interpret, especially when the copula is given in terms of a distribution function rather than being derived from a known multivariate distribution. Furthermore, fitting the parameters of a copula is a difficult statistical problem – the estimators used are often complex and not always robust. These estimators (particularly the correlation matrices used in Gaussian, t, or other elliptical copulas) have the vulnerability that they may change over time or during stress periods. Many firms have discovered that static models based on historical correlations do not prove accurate when market variables undergo stress. For all of these reasons, implementing copulas requires a high level of statistical expertise on the part of the practitioner, and management and other employees who use the output from copula techniques must be sufficiently versed in the technical aspects of this approach to understand the limitations for a given aggregation of the firm's risks.

G.3 Scenario-based aggregation

The previous approaches aggregate by combining statistically derived distributions. In contrast, scenario-based aggregation aggregates risk expressions to common underlying scenarios. A scenario is an expression of the state of the financial institution or its portfolios under certain defined conditions of the external environment. The total profit or loss characterising the financial institution is usually a simple summation of the partial profit and losses of the various positions composing the portfolios.

Scenario-analysis - Determining risk drivers and exposures

Developing relevant scenarios requires profound knowledge of the portfolios of the financial institution to adequately identify and understand the positions taken by the financial institution.[53] In addition, it requires an identification of the risk drivers of these positions; ie the stochastic variables that determine the performance of the respective position. Risk drivers are commonly external to the financial institution, eg financial and economic variables such as interest rates, GDP and unemployment, health conditions, or weather conditions determining the natural catastrophes, or social conditions determining material damages caused by humans (eg car accidents, fire). However, the risk drivers can also be internal or partly internal to the financial institution, for instance drivers of operational risk. In a next step, the risk positions are defined in terms of risk exposures relative to the risk drivers. Risk exposures are summary risk expressions characterising how the economic value of the risk position depends on the risk driver.

These relationships between the risk exposures and the risk drivers are typically linearised to preserve a certain level of simplicity. In this case the sensitivity of the risk position to a risk driver is described. However, practitioners should bear in mind that dependencies are not always linear (eg with interest rates, or in presence of optionalities).

A comprehensive analysis of the risk portfolios of the financial institution to well identify and understand the risk positions, and their drivers, and studying and describing the impact of the risk drivers and the changes in exposures comprise the building blocks of scenario analysis. These building blocks are necessary aspects to adequately perform scenario stress tests. In this respect scenario stress tests can be seen as a particular type of scenario analysis, that focuses on capturing and assessing potential "real-life" extreme events on the economic value of the financial institution.

Some interviewed firms go one step further and simulate a multitude of scenarios, these simulations are based on scenario generators.

Scenario simulation

Scenario simulation and scenario generation is made possible through the high computing power of modern computers. Large series of scenarios are generated by independently drawing large numbers of random variables and processing the random draws through models that describe particular processes or phenomena (eg a natural storm, a pandemic or an economic evolution).

Three types of models or algorithms are distinguished underlying these scenario generators:

1. A first category consists of models that tries to describe and proxy "real physical processes or natural laws". These processes usually rely on dynamic modelling that let risk factors develop through time. Examples are for instance pandemics, for which models describe how the virus can change, propagate between individuals and how it acts on the individual resulting in states of sickness of various severity or in death. Similar approaches are being used for wind storms and earthquakes.

2. A second category would be models that describe processes for which there is no real physical model. The models in this category may rely on a particular theory, which might (partially) fit certain historical observations. These models then underlie the scenarios simulated. Often certain ad-hoc distribution assumptions are used to

[53] Positions on the asset side as well as on the liabilities side are considered by the financial institutions as well as risk positions arising from internal processes of the enterprise (for instance, intra-group transactions, service agreements and commitments)

perform the simulations. This is, for instance, the case for most of the financial processes (eg interest rates, equities prices and indices, exchange rates).

3. Finally the third class of processes combines the first two categories and mixes physical with theoretical descriptive and empirical processes. Examples might for instance be processes that integrated natural catastrophe scenarios with economic and financial scenarios.

Firms may use several different and unrelated event generators to generate particular scenarios for different portfolios. These different types of scenarios may need to be integrated in order to obtain global scenarios that are applied on the different portfolios of the financial institution. When the scenarios are considered to be independent, scenarios can simply be randomly combined. However, when the scenarios are not independent mechanisms have to allow for interactions between the generators. These interactions are introduced through different methods. A simple technical method consists of using a covariance matrix (ie VarCovar approach as described above). Copulas are also often used to combine events from different event generators. In case the methods clearly show limitations and interactions between the scenarios are complex then the modelling of the different processes have to be performed simultaneously, leading to complicated event generators (eg pandemics and financial market risk factors).

As is typically the case when relying on simulation techniques, an issue to address is the number of simulation runs or events required to obtain an adequate level of precision in the estimate. Determining the required precision is not a difficult task but it is important to consider in the risk measurement process.

Conclusions

Scenario-based aggregation is conceptually and intellectually appealing as it eliminates ad-hoc methods of aggregation by aggregating exposures on the basis of common scenarios. As a result, the risk aggregation process avoids the common approach of inserting statistically derived distributions into risk management processes that may not reflect operational or legal business lines.

Proper scenario-based aggregation requires a profound understanding of the risks the firm is exposed to; it forces the firm to make extensive assessments of its portfolio risks and to identify risk drivers and assess the exposures to these risk drivers. Obtaining a clear view on the economic condition of the firm and deriving relevant economic scenarios proves hard and requires a strong reliance on the expert judgment and qualitative insights of the management. It requires experts and managers to develop solid representations of and views in various areas of the economic "reality" of their financial institution. Building this required knowledge and understanding takes time and it is not without its own risks as the potential to overlook exposures or have the profile change during the modelling period is an issue that the experts need to remain cognizant of.

In addition, scenario-based aggregation relies heavily on a range of assumptions (eg regarding the scenarios considered or developed, the scenarios being selected and the expression of risk positions in terms of exposures to the scenarios) which have to be well understood and considered when interpreting the results. In particular, it is not straightforward to reconcile the scenarios with the more traditional parametric descriptions of the risk.

The results of the scenario analysis and scenario simulation can be relatively easily and meaningfully interpreted in an economic and financial context. In addition, a firm can, for example, develop emergency or recovery plans from extreme scenarios. Consequently, although the team performing the modelling usually needs to be relatively highly specialised, the results and interpretation of the results can be understood by a non expert executive manager.

Scenario simulation requires sufficient computing power and solid IT programs and platforms. The modern and well built simulation programmes commonly support changes to the dependencies and the distributions, allowing the user to make prospective studies, study sensitivities and stability of processes and thereby test the robustness of the methods. The programs can allow certain scenarios to be given specific weights and additional scenarios can be (artificially) added to focus on particular aspects of the risk. These methods demonstrate a great deal of flexibility that does not exist in the more simple aggregation methods.

Box A – Coherent risk measures

A reasonable description of a risk situation requires knowledge of a lot of information whereas a risk measure is a single number, leading to a significant reduction of information. On the other hand, a financial actor (eg investor, underwriter or regulator) takes a binary decision (to invest , to subscribe, to authorise). In this respect, a risk measure should as adequately as possible reflect the properties of the risk being considered.

Properties required by a coherent risk measure

At least 15 years of agitated discussions have resulted in the establishment of the essential properties that an adequate risk measure should obey([1], [2]):

- The choice of a reference instrument (usually a one year risk free state bond) is a vital ingredient and can be considered as a yardstick to which the risk will be compared.
- Positive homogeneity: implies that if a position has a risk, doubling the risk position leads to doubling the risk.
- Sub-additivity: means that the risk of the sum of two positions is always smaller or equal to the sum of the risks of the two positions.
- Translation invariance: adding to a portfolio an amount of cash invested in the reference instrument reduces the risk measurement of this portfolio by the same amount.
- Monotonicity: a position that always results in smaller losses than another position always has a smaller risk than the other position.

Examples of coherent and non-coherent risk measures:

Several risks measures are already widely used in practice. They have different properties:

- Total exposure: is a coherent risk measure, in fact the most severe one for a given reference instrument. In practice, the exposure is traditionally "weighted" to provide a more adequate risk measure.
- Standard deviation based risk measures: defined as the standard deviation relative to the expected value of the position. A certain refinement would be the use of the semi variance on the loss side of the distribution. It is not a coherent risk measure as it does not fulfill the monotonicity property.
- Value at Risk: simply reflects the quantile at a particular defined quantile level (α). This is a widely used and intuitive risk measure, however, it is not a coherent risk measure as it does not fulfill the sub-additivity property. This poses a severe draw-back to the use of the VaR measure.
- Expected Shortfall (or Tail-Value-at-Risk): has been mainly developed to cope with the non sub-additivity condition of the Value-at-Risk measure. It is defined as the average value of the losses at quantiles lower than the specified quantile (α).

Coherent measures of risk in practice

A crucial property for a coherent risk measure is the sub-additivity condition. As mentioned previously the VaR measure fails to satisfy this condition although the measure is widely used in practice. This can be explained by the fact that when distributions are normal or close to normal it can be shown that Value at Risk and Expected Shortfall are quite close and behave similarly. In this respect, normal distribution assumptions are quite common to simplify the risk measurement. However, as soon as a risk position is characterised by a long tail behaviour, the similarity between VaR and ES does not hold anymore. Unwarily employing the VaR measure to risk positions characterised by long tail risk may lead to a strong underestimation of the risk. Furthermore mixing risk measures established with different reference instruments and with different currencies can also lead to unexpected behaviours.

Coherent risk measure and use of scenarios

A highly interesting property (known as the representation theorem, see [1]) allows establishing a comprehensive link between coherent risk measures and scenarios.

A scenario is stricto sensu a well precise possible realisation of the future (eg stock prices fall by 20% and interest rates rise by 100 bp). This concept can be generalised by considering a set of such scenarios weighted by some probabilities that are subjective and represent what practitioners (industry and regulators) consider as potential future realisations. This information represents nothing else but a probability density function called a generalised scenario.

The representation theorem states that a coherent risk measure is fully defined by a family of generalised scenarios and vice versa. This property emphasises and favours the use of scenarios by financial institutions to assess their risks as it allows more than the other methods to stay compatible with the coherence of the risk measure which has be shown to be a fundamental property.

References to articles:

- [1] Philippe Artzner, "Application of Coherent Risk Measures to Capital Requirements in Insurance", North American Actuarial Journal, Volume 3, Number 2, April 1999
- [2] Glenn Myers, Coherent measures of risks, an exposition for the Lay actuaries, Casualty Actuarial Society, 2000.

Box B – Correlations vs dependencies

Linear correlations have been quite popular in finance and business applications. The calculation of the linear correlation coefficient (also called Pearson Product Moment Correlation) is straight-forward, and it is easily linked to linear models that are intuitive and readily explainable to a wide variety of users. A given correlation suffices, in a relatively straightforward fashion, to explain the joint or combined probability distribution for two or more normal random variables, thus preserving the general computational advantages and intuition (and real-world weakness) of modelling with the normal distribution. This is highly convenient both for underlying stochastic models involved in pricing and risk analysis as well as models that may be used for top-of-the-house risk aggregation. The following table explains correlation measures such as linear correlation in the context of the more general concept of statistical dependence, with which it may easily be confused.

	Dependence	Correlation
Independence	Event A: Credit-related losses on the consumer loan portfolio will exceed USD 150 million. Event B: Insurance claims will exceed premiums by USD 20 million. In statistics, A and B are *independent* if Pr (A | B) = Pr (A) or, equivalently, Pr (B | A) = Pr (B) Verbally: "Probability of A given B is equal to the Probability of A." "Probability of A is the same whether or not B occurs."	
Dependence versus correlation	Dependence means that the probability distribution of a variable is different depending on the state of the other variable. If the events are dependent, whether one of the events occurs should cause one to change his or her estimate of the probability that the other occurs.	Correlation is a commonly used label for specific measures of dependence between pairs of variables. In qualitative discussions, "correlation" is often not carefully distinguished from "dependence".
Characterisation & scaling	Qualitatively, there are varying degrees of dependence. Variables that are highly dependent may have conditional distributions (eg, probability of A given B) that are very different from their unconditional distributions (probability of A assuming nothing about B). The degree of dependence may vary with the value of the conditioning ("given") variable. If extreme values for that variable are associated with relatively high conditional probabilities for dependent variables, which may signify high tail dependence.	Correlation is a normalised measure of dependence. "Correlation coefficient" most often refers to Pearson Product Moment Correlation. This is a measure of linear relationship and is scaled from -1 to 1. Other correlation coefficients include Spearman's Rank Correlation & Kendall's Tau, also scaled from -1 to 1. Independent random variables have a correlation of zero. The closer to -1 or 1, the stronger the level of dependence.

Limits	Dependence is described in full only by knowing the entire joint probability distribution – ie, the function describing the probability of any possible combination of outcomes.	Correlation measures can be limited in how they represent dependence. A set of variables can have important dependencies (eg, at the tail) that are not represented clearly by a specific measure of correlation. Also, a given correlation might not distinguish between two very different joint distributions. Risk managers hope to capture important dependencies rather than measure correlation well, per se.
	In business and financial time series, correlation measures often prove to be unstable. Particularly during stressful periods, correlation may appear, after the fact, to have increased significantly, leading to greater than anticipated losses. It can be difficult or impossible for risk managers to obtain reliable, time-independent measures of dependence due to potential changes in the overall dependence structure.	