Risk Measure and Allocation Terminology

Gary G. Venter and John A. Major February 2009

Notation

Y is a random variable representing some financial metric for a company (say, insured losses) with cumulative distribution F(y) and with $Y = \Sigma X_j$ being the sum of similar financials for the business units (which could even be individual policies). $\rho(Y)$ is a risk measure on Y and r is the allocation, i.e., $\rho(Y) = \Sigma r_j(X_j)$.

Terminology

Allocation Terminology

Proportional allocation: allocating a risk measure by calculating the risk measure on the company and each business unit, and allocating by the ratio of the unit risk to the company risk: $r(X_j) = \rho(Y)\rho(X_j)/\Sigma\rho(X_j)$

Marginal allocation: allocating in proportion to the impact of the business unit on the company risk measure. This can be: **last-in marginal allocation** where the impact of the business unit is $\rho(Y)\text{-}\rho(Y\text{-}X_j)$, the company risk measure with and without the unit; **Aumann allocation**, where the impact is averaged over every coalition of business units that unit can be in, $\rho(Y)-\rho(Y-(X_{k1}+\ldots+X_j))$; or **incremental marginal allocation** where the impact is $[\rho(Y)-\rho(Y-\varepsilon X_j)]/\varepsilon$, the change in the company risk measure from eliminating a small proportional part of the unit, grossed up to the size of the whole unit. In the limit $\varepsilon\to0$ this is the derivative of the company risk measure with respect to the volume of the business unit.

Marginal decomposition: when the incremental marginal impacts add up to the whole risk measure (i.e. the required proportionality constant is one), the allocation is called a marginal decomposition of the company risk measure. By Euler's Theorem, this happens when the risk measure is homogenous degree 1: for a positive constant k, $\rho(kY) = k\rho(Y)$. Marginal decomposition is also called Euler allocation. In all known examples, the decomposition can be expressed as a co-measure (see below). However not all co-measures produce marginal decompositions.

Co-TVaR and Myers-Read (below) are examples of marginal decomposition. The standard deviation has a marginal decomposition equal to the covariance of the unit with the company divided by the standard deviation of the company. This can be seen by taking the derivative of the company standard deviation with respect to the volume of the business units, using l'Hôpital's rule for the limit.

Suitable: under a suitable allocation, if you allocate capital in proportion to the allocation of a risk measure, and compute the expected return on allocated capital, then increasing the size of a business unit that has a higher-than-average return on capital will increase the return on capital for the firm. Marginal decomposition always produces a suitable allocation, and is the only method that does.

Co-measure: defined if $\rho(Y)$ can be expressed as:

$$\rho(Y) = \Sigma_i\{E[h_i(Y)L_i(Y) \mid i^{th} \text{ condition on } Y]\},$$

where h_i is a scalar function which is additive, i.e., $h(V+W) = h(V)+h(W)$, $L_i(Y)$ is a random variable whose value, given the value of Y, is deterministic,[1] and the only other restriction is that the conditional expected value exists. Then the co-measure is defined by:

$$r_j(X_j) = r(X_j) = \Sigma_i\{E[h_i(X_j)L_i(Y) \mid i^{th} \text{ condition on } Y]\}$$

By the additivity of the h's, the co-measures add up over the business units to the whole risk measure. For instance if there is only one h and L, with the condition $Y > F^{-1}(0.99)$, and $L(Z) = 1$ and h is $h(Z) = Z$, then:

$\rho(Y) = E[Y \mid Y > F^{-1}(0.99)]$ is TVaR at 99% and $r(X_j) = E[X_j \mid Y > F^{-1}(0.99)]$ is the j^{th} co-TVaR. The ability to take a sum of several functions (indexed by i) allows risk measures like the sum of TVaR at different probability levels to be represented by co-measures also.

A less trivial example is risk-adjusted TVaR, or RTVaR. This has two sets of h's and L's. The condition on Y is $Y > F^{-1}(\alpha)$ for both sets, with $h_1(Z) = h_2(Z) = Z$, $L_1(Y) = 1$, and $L_2(Y) = c(Y - E[Y \mid F(Y)>\alpha])/\text{Stdev}(Y \mid F(Y)>\alpha)$ for some constant c, which will usually be between zero and one. Then:

$$\rho(Y) = E[Y \mid Y > F^{-1}(\alpha)] + c\,\text{Cov}(Y,Y \mid F(Y)>\alpha)/\text{Stdev}[Y \mid F(Y)>\alpha]$$

$$= \text{TVaR}_\alpha + c\,\text{Stdev}[Y \mid F(Y) >\alpha].$$

The co-measure, co-RTVaR, which is marginal (see definition below), is:

---

[1] This is almost, but not quite, the same as saying L_i is a scalar function. The difference is that functional aspects, like subtracting the expected value of Y, might be incorporated into the definition of $L_i(Y)$. This distinction is important when computing derivatives of the risk measure.

$$r(X_j) = \text{co-TVaR}_\alpha(X_j) + c\,\text{Cov}(X_j, Y \mid F(Y)>\alpha)\,/\text{Stdev}[Y \mid F(Y)>\alpha].$$

Riskiness-leverage function: a function L(y) to express the risk-adjusted value of a random variable ρ(Y) as E[(Y-aEY)L(Y)]. It is from the Kreps paper "Riskiness Leverage Models," PCAS XCII, 2005. In this case the co-measure is $r(X_j) = E[(X_j\text{-}aEX_j)L(Y)]$. For instance if L(Y) is the indicator function for F(Y) > 0.99, and a = 0, this is TVaR at 99%.

RMK algorithm: a method for determining a company's riskiness leverage function L(y) to express its risk preferences, creating a risk measure ρ(Y) from it, and allocating in accord with the method of co-measures. The name derives from the Kreps paper on riskiness leverage functions and the Mango-Ruhm paper, "A Risk Charge Calculation Based on Conditional Probability" 2003 ASTIN Colloquium, which worked out the result in the case a=0, which includes TVaR. Both papers were originally in the risk-pricing context, but have been applied to capital allocation as well.

Myers-Read allocation: an additive marginal allocation method that requires that the value of the default put option be the same fraction of expected loss for each business unit. The risk measure is required capital itself, and the incremental marginal change in required capital from a small proportional reduction to a business unit is the amount by which capital can be reduced and still keep the same company-wide ratio of default put value to expected losses. This incremental marginal change in capital is grossed up to the volume of the whole business unit to give the capital allocated to the unit. It turns out that these allocations add up to the overall capital. Thus the Myers-Read method is one of many marginal decompositions.

Allocation by layer: an allocation method introduced by Niel Bodoff, 2008 "Capital Allocation by Percentile Layer," CAS Forum. It is easiest to describe in the context of a simulation model. Let $X_{i,k}$ be the loss to unit i in the k^{th} simulation, which has total company losses Y_k. Assume all the simulations are equally likely. In Bodoff's original paper, the risk measure $\rho(Y) = VaR_\alpha = F^{-1}(\alpha)$, estimated by $Y_{N\alpha}$, is to be allocated to unit. However, it turns out the same allocation applies to a particular capital amount C, regardless of the risk measure used to compute it.

Layers of losses up to C are needed for the allocation. In Bodoff's paper, he takes layers equal to the intervals between the sorted simulated losses. However, the method can be made to work for any definition of layers. To take a fairly extreme case, assume that layer z is the layer from (z-1)US¢ to z¢, and that all simulations have been rounded to whole cents.

(Yen would be approximately the same.) Define n_z as the number of simulations Y_k that are z or greater. The allocation of layer z to unit i is $\frac{1}{n_z}\sum_{k\ with\ Y_k\geq z}\frac{X_{i,k}}{Y_k}$. The allocation of C to unit i is $r(X_i)=\sum_{z=1}^{C}\frac{1}{n_z}\sum_{k\ with\ Y_k\geq z}\frac{X_{i,k}}{Y_k}$. As a check, summing the allocation over the units (i's) gives $\sum_{z=1}^{C}\frac{n_z}{n_z}=C$.

This allocation has some good properties. All layers contribute to the allocation, so it does not ignore smaller but potentially painful losses. Also the larger simulations get into the allocation for all lower layers, so they accumulate a greater allocation overall. Thus the units that generate large losses get a bigger allocation. The main weakness of the method seems to be that it is not marginal for any known risk measures; in particular, it is not a marginal allocation of *VaR*. Also, it is not clear how to apply the method to financial metrics that have both upside and downside, like net profit.

Aumann-Shapley method: an accounting method of cost allocation for the cost of production facilities that are used to produce a variety of products. For risk measures or cost functions that are homogeneous of degree 1, this is the same as the Euler method, or marginal decomposition. For other functions it is a similar average over all production levels from zero to full capacity, but is not really applicable to insurance lines of business.

Shared-asset (also called **capital consumption** or **Merton-Perold):** not really an allocation method but a way of computing the cost of capital for each business unit. Not even an *allocation* of the cost of capital, in that the total cost of capital is not used in the calculation and the unit capital costs might not add up to the company capital cost calculated from other methods. The cost of capital for a unit is the value of its right to use the capital of the firm if it runs out of its own funds, and so is the value of a put option. Merton and Perold make some strong assumptions about the random variables in order to use the Black-Scholes formula to evaluate the option values, so they provide only one example of the shared-asset method. The value added by a unit is its profit minus its cost of capital. The profit measure is also an option (a call), in that the company takes all the profit if there is any, and none otherwise. The values added by the units can be summed to get a value for the firm, so capital consumption can be considered as an allocation of firm value. Don Mango, 2005 "Insurance Capital as a Shared Asset," ASTIN Bulletin explores more realistic insurance-oriented distributional and pricing assumptions for calculating the put-option value.

Risk Measure Terminology

Coherent: a risk measure meeting a few mathematical requirements, the most controversial and most often failing being *subadditivity*: the risk measure of a sum of random variables $\rho(X_1+\ldots+X_n)$ should not be greater than the sum of their risk measures $\Sigma\rho(X_j)$. This is a useful criterion if the question you are addressing is measuring the diversification benefit from combining business units, and you want to guarantee in advance that the answer will not be negative. Otherwise it is not really a necessary requirement. Since marginal allocation does not look at the risk of individual units, but rather their contribution to the risk of the whole, subadditivity is usually not relevant.

Homogenous degree n: for any positive constant k, $\rho(kY) = k^n\rho(Y)$. For example, variance is homogenous degree 2.

Spectral: a risk measure of the form $E[Y\eta(F(Y))]$ for some non-negative scalar weighting function η on the unit interval. For example, $TVaR_\alpha(Y) = E[Y|F(Y) > \alpha] = E[Y\eta(F(Y))]$ where $\eta(p) = 0$ if $p < \alpha$ and $\eta(p) = 1/(1-\alpha)$ if $p > \alpha$. Another example is:

$$\eta(p)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{p-\alpha}{\sigma}\right)^2\right)$$

which makes the weight a normal distribution centered at the α percentile. This gives a blurred VaR. You could also define a blurred VaR using a uniform distribution centered at α. Usually the allocation of VaR by numerical computation is actually the allocation of a blurred VaR of some kind. Some authors used to limit spectral measures to coherent spectral measures.

Distortion: defined by a distribution function g(x) on the unit interval: $\rho(Y) = \int_0^\infty g[S(y)]dy$, where $S(y) = 1 - F(y)$ is the survival function. Here the role of g is to transform the probabilities of Y, and in fact a distortion risk measure is a transformed mean. The marginal decomposition of a transformed mean is the transformed mean of the unit, where the transform uses the transformed probabilities of the aggregate firm variable. Using the same *g* on the survival functions of the units will not always give the same result.

Famous examples of distortion measures are $g(p) = p^a$ (proportional hazards transform) and the Wang transform $g(p) = 1 - T_a[\Phi^{-1}(1-p) - b]$, where T_a is the t-distribution function with *a* degrees of freedom, and Φ is the standard normal distribution.

However $VaR_{0.99}$ and $TVaR_{0.99}$ are also distortion measures. They both have $g(p) = 1$ if $p > 0.01$. Note that $g[S(y)]$ then is 1 when $F(y) < 0.99$, so the portion of the integral from 0 to $F^{-1}(0.99)$ is $F^{-1}(0.99)$. VaR has $g(p) = 0$ otherwise, whereas TVaR has $g(p) = p/0.01$ otherwise.

Since the transform depends only on the probabilities, it might be suspected that distortion measures are spectral measures. In fact they are. A bit of calculus can show that taking $\eta(p) = g'(1 - p)$ will put any distortion measure into the spectral form.

Complete: the risk measure uses the entire probability distribution of Y in a non-trivial way. This can be formally defined for distortion risk measure by requiring that g(p) is not constant on any interval, and so is an increasing function on the unit interval. No tail measures would satisfy this definition. The motivation is that if the risk measure is to be used to express a preference among random variables, this cannot be done using the tail alone. This is especially the case for variables that are the same in the tail but differ elsewhere.

Adapted: if a risk measure is going to be used in pricing, typically you would not want it to be less than the mean of Y. For a distortion measure this requires that $g(p) \geq p$. However another typical requirement is that in the tail the relative risk load is unbounded. This would be needed, for instance, to get a minimum rate on line. For distortion measures this can be expressed as $g'' < 0$ and g' goes to infinity at $p = 0$. An adapted risk measure is one that meets both criteria. The Wang transform is an example. If the unlimited variance of Y is infinite, then the standard deviation loading of higher layers would also increase without bound, and so would meet this criterion. Usually minimum rates on line are used only with heavy-tailed distributions, so this is a realistic example.

Transformed distributions: not every transformed distribution is a distortion measure. Consider for instance the Esscher transform $f^*(y) = f(y)e^{y/c}/E[e^{Y/c}]$. This does not exist for many heavy-tailed distributions, but in practice losses will be capped by policy limits which will make the transform finite. It has a free parameter *c* that determines the change in level. The change in probability depends on the value *y* of the loss. This does not happen with distortion measures since they are spectral measures, that is the transform is a function of the probability but not the value of the loss. The Esscher transform is thus not a spectral measure.