Risk Measurement in Insurance

A Guide To Risk Measurement, Capital Allocation And Related Decision Support Issues

Paul Kaye

Abstract

Risk measurement provides fundamental support to decision making within the insurance industry. In spite of this, the limitations of the common measures are not well appreciated and there is little non-specialist awareness of the more powerful techniques.

The published material on risk measurement is strong and has developed significantly in recent years. However, it is fragmented and is not always in a form that is accessible to many industry practitioners. Also, notwithstanding the theoretical merits or otherwise of different techniques, many practical attempts to measure risk can be compromised by inappropriate use and interpretation.

This paper aims to give an accessible overview of the full range of risk measurement and allocation techniques, critiquing both technical properties and practical considerations. A simple example is used throughout the paper to help illustrate the various measures and methods, with values being calculated using stochastic simulation.

Keywords. Risk Measurement; Capital Allocation; Dynamic Financial Analysis.

1. INTRODUCTION

Risk measurement is fundamental to the insurance industry, from the pricing of individual contracts to the management of insurance and reinsurance companies to the overall regulation of the industry. Putting aside the inherent complexities of risk modelling and quantification, there is a more fundamental issue: are the common risk measurement techniques adopted in the industry appropriate for the purpose for which they are being used? Also, are the more recent developments in risk measurement (and risk/capital allocation) thinking and their potential benefits well understood within the industry?

- Common risk measures, such as the outcome at specific percentiles (e.g. the 1 in 100 loss exceedence is X) and standard deviation, are often misinterpreted and abused – partly due to some fundamental limitations in the ability of these measures to describe 'risk'
- More sophisticated techniques typically require a more complete articulation of the probability distribution of outcomes – such as might be provided from a stochastic model. However, these methods are often technically more robust, are often not as intimidating as they first sound, and have the potential to gain much wider

acceptance and use.

There is a risk that any discussion on risk measurement (and related issues such as capital allocation) can become very academic and technical in nature. This is because, particularly at its more exotic extremes, the topic can involve relatively advanced mathematical thinking with all the associated perils of inaccessibility to non-academic mortals – mathematical notation and jargon, the presumption of prior understanding of background topics, and a limited focus on the perennial 'so what?' challenge (beyond the inevitable 'academic interest').

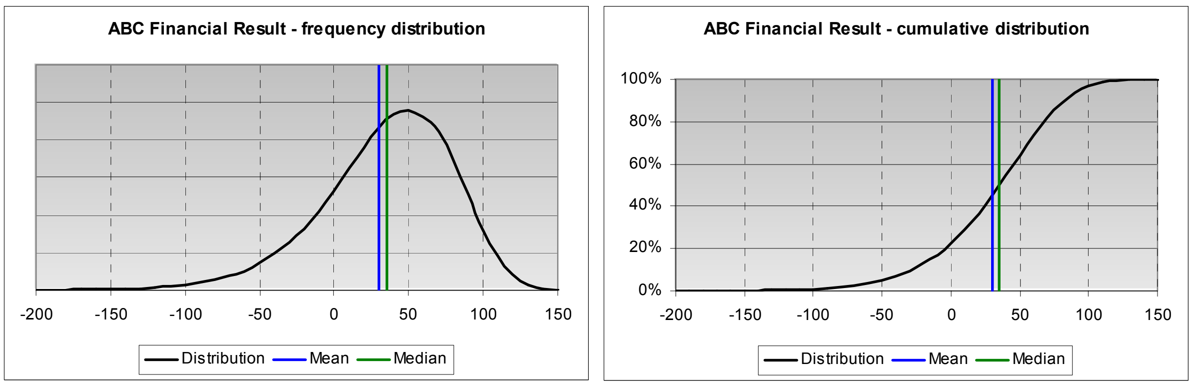
This paper attempts to give a simplified guide through the maze – the commercial need for risk measurement, the methods available and their relative pros and cons, including the con of incomprehensibility, and, finally, a practical illustration using a case study with some real numbers. The extent to which it succeeds is for the reader to judge and any comments to help improve the presentation of ideas would be welcome.

Where possible, attempts are made to give credit to the original thinking behind the methods and concepts covered. However, there is a risk that this in itself can add to the confusion. It is apparent that a number of ideas have their groundings in more than one area, be it traditional statistics, game theory, operational research, financial pricing theory, corporate finance and, of course, insurance-related actuarial science. Many methods have gained names, in some cases more than one (or at least several different names for similar variants). In general the approach taken is to focus primarily on what the methods and concepts are about, occasionally at the inevitable expense of misrepresenting a name and/or the true originators of the thinking. Apologies are given where proper recognition is possibly misstated, or is omitted – efforts will be made to correct any oversights in later versions.

Note that this paper is limited to taking a relatively purist 'economic value' perspective of risk. In practice, when considering topics such as capital adequacy, issues arise over the foibles of the pragmatic compromises made by rating agencies and regulatory authorities. The resultant anomalies form an important and interesting topic in its own right – but one that is too large to cover here adequately.

2. THE CHALLENGE

Consider a company ABC with the following distribution of financial outcomes:



What business questions might we ask regarding this uncertainty?

- Likelihood of going bust? How much capital is required to support this business?
- Likelihood of failing to meet one or more near term performance objectives?
- Contribution of each sub-portfolio to overall ABC risk?
- What minimum profitability margin should be targeted for each risk underwritten in order to deliver an acceptable individual and overall return on capital?
- How might the dynamics of this uncertainty be changed: business strategy, reinsurance strategy, general risk management, asset allocation, capital management etc?

In all cases we are looking for some kind of quantification of risk to provide decision support to the relevant stakeholder (management, investor, underwriter, rating agency etc). Note that it is common to look at risk in organisations in terms of the capital required to support a given exposure. However using the term 'capital' for certain risk measurement problems can be unhelpful. Capital is necessary as a buffer against extreme scenarios – but not all problems/stakeholders are concerned with remote organisation-threatening occurrences. In certain situations it is valuable to look at risk in a more abstract way – as illustrated in Mango's concept of a 'concentration charge'[1] for catastrophe pricing. The use of risk measurement for both capital and other more abstract risk based decision support challenges will be considered as part of the evaluation of the various methods discussed in this paper.

To maximise the decision support provided, the risk quantification will need to satisfy a number of requirements:

- Different stakeholders have different levels of interest in different parts of the distribution – the perspective of the decision-maker is important. Regulators and rating agencies will be focused on the extreme downside where the very existence of the company is in doubt. On the other hand, management and investors will have a greater interest in more near-term scenarios towards the middle of the distribution and will focus on the likelihood of making a profit as well as a loss.
- The approach taken to measure risk needs to be suitable for the purpose for which it is being used. This refers to both the properties of the risk measure selected as well as the risk tolerance(s) selected for a given measure. For example, risk is commonly measured by looking at the result for a specific return period. What are the limitations in using such a measure? In what circumstances will such limitations come to the fore? Which return periods might be considered for the stakeholder/question (e.g. 1 in 10 or 1 in 100 risk or both)?
- Is the risk measure understood by the decision-maker? A detailed technical understanding may not be essential if there is a good appreciation of how the measure should be used and its values interpreted. However, the complexity of more sophisticated measures may create a barrier for a decision-maker when it comes to the critical stage of making softer judgements alongside the hard quantitative results of the risk analysis.

So the challenge…

What alternatives are there for measuring risk?

What alternatives are there for allocating a given risk measurement value across sub-portfolios?

Which methods work well?

Which methods are more appropriate in which circumstances?

> **How can risk measurement be made as comprehensible as possible, particularly for non-technical decision makers?**

The two graphs shown at the beginning of this section come from a simple stochastic simulation model built to illustrate the concepts and methods discussed in this paper. Company ABC is made up of three sub-portfolios, A, B and C. Each has non-variable income of 100 less 25% expenses and losses distributed LogNormally with a mean of 65 and a standard deviation of 20. The losses for A and B are highly correlated while those for C are totally uncorrelated with A or B. The results shown in this paper are based on a stochastic simulation of 100,000 trials.

The concepts and methods covered in this paper can be applied to the simplest and most sophisticated of risk models. The model used for illustration is obviously somewhat simple but that does not mean it could not be readily expanded to include, or be deemed to include, more complex features. For example, 'losses' might be considered to include all aspects of uncertainty including asset and reserving risk etc. In addition investment income, whilst not explicitly recognised here, might be viewed as having been reflected within the income figure and by the losses being parameterised on a present value basis. Similarly, issues over single- and multi-year uncertainties are not specifically considered – the results are notionally for a single year. However, but this does not preclude multi-year thinking being incorporated within the parameterisation.

3. DESIRABLE TECHNICAL PROPERTIES

Aside from the issues of comprehensibility and application, we should consider the basic properties we might expect of something that purports to measure risk. The concept of coherence is introduced both for risk measures (Artzner et al [2]) and risk allocation methodologies (Denault [3]) in 3.1 and 3.2. These represent a basic set of common sense rules, the failure to comply with which must put into question a method's suitability for measuring or allocating risk.

3.1 Risk Measures

A risk measure shall be deemed to be Coherent if it satisfies the following properties:

- SUB-ADDITIVITY

  Combining two portfolios should not create more risk. When portfolios of risk are brought together, there may or may not be some risk diversification benefit. However, it would be counterintuitive to get an anti-diversification 'benefit'!

  $$\text{Risk}(A+B) <= \text{Risk}(A) + \text{Risk}(B)$$

- MONOTONICITY

  If a portfolio is always worth more than another, it cannot be riskier. Consider company ABC – if one strategy gave modeled financial results better than another for any given return period, the risk, as measured by a coherent risk measure, must be better.

  $$\text{Risk}(A) <= \text{Risk}(B) \text{ if } A >= B$$

- POSITIVE HOMOGENEITY

  Scaling a portfolio by a constant will change the risk by the same proportion. For example changing the currency being used or buying a quota share should change the risk by the exchange rate and quota share retention respectively. This also equates to a specific case of the sub-additivity condition: when combining two perfectly correlated portfolios, the total risk should be the sum of the risk of the parts.

  $$\text{Risk}(kA) = k\text{Risk}(A), \text{ for any constant } k$$

- TRANSLATION INVARIANCE

  Adding a risk free portfolio to an existing portfolio creates no change in risk. For example, when focusing on capital adequacy, adding additional capital to a company will not change the portfolio's volatility (although it will reduce the risk to policyholders). Similarly, with the company ABC example, where the underwriting result is being measured, adding additional non-variable income (e.g. premium from a rating increase), the financial result will be improved by this amount in all outcomes.

  $$\text{Risk}(A+k) = \text{Risk}(A) + k \text{ ,for any constant } k$$

3.1 Risk/Capital Allocation

It is usual to refer to capital allocation methods. However, the same principles apply if a risk tolerance is used that does not come close to equating to a capital value – the measurement may be more abstract but this does not necessarily limit the potential benefit to decision support problems.

A capital (or risk) allocation methodology shall be deemed to be Coherent if it satisfies the following properties:

- NO UNDERCUT

  The allocation for a sub-portfolio (or coalition of sub-portfolios) should be no greater than if it was considered separately. A sub-portfolio's allocation should be no more than its standalone capital requirement and, similarly, a sub-portfolio's allocation should be at least as great as its marginal (last in) allocation, i.e. the other portfolios shouldn't be better off without the sub-portfolio. This is a key property which encourages individual behaviour to be in a group's interests.

- SYMMETRY

  If the risk of two sub-portfolios is the same (as measured by the risk measure), the allocation should be the same for each

- RISKLESS ALLOCATION

  Cash in a sub-portfolio reduces allocation accordingly. This is akin to the translation invariance property for a risk measure.

A crucial point to note, also, is that an allocation method will not be coherent unless the risk measure chosen is coherent

4. RISK MEASUREMENT ALTERNATIVES

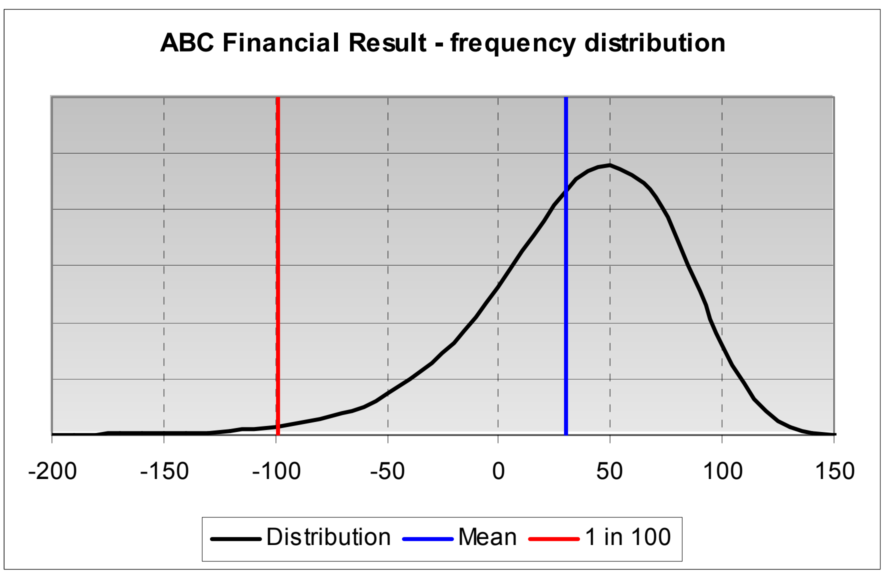
Rather than give a long list of every conceivable risk measure, an attempt has been made to group those which have similar characteristics. Also, for comparative purposes, the way in which each measure might produce a value approximately equal to the 1 in 100 downside result is shown for company ABC.

It may be that the risk measurement value is used as a benchmark for capital adequacy.

Where this is the case it should be noted that, as in the case of the company ABC, any financial loss is after allowing for any income (i.e. 75 net of expenses). For more near-term risk tolerances, it is quite possible for the risk measurement not to represent a net loss scenario (or set of such scenarios), e.g. ABC still makes a profit at the 1 in 4 downside result. In such cases the 'risk' from the losses is not greater than the premium. This is not necessarily an unreasonable result but can create some confusion and anomalies when using risk measurement values in certain practical situations.

4.1 Point Measures

The value of a distribution of outcomes at a single point. This point might be defined using a given percentile (this is akin to Value at Risk (VaR), which is commonly used in the banking sector) or the probability of an outcome being worse than a given monetary threshold (e.g. probability of ruin).



The result for company ABC is forecast to be a loss of -98.4 or worse every 100 years

4.1.1 Technical Evaluation

This risk measure only focuses on a single point of the distribution. It captures no information for the decision-maker regarding how the tail of the distribution behaves (nor any other part of the distribution for that matter). Unless there is an overriding purpose that may justify such a narrow perspective of the 'risk', there is a serious danger that such a metric can lead to a decision-maker making inappropriate inferences regarding the behaviour of other parts of the distribution.

Such measures are not coherent as they fail the sub-additivity requirement. This can be illustrated using a simple example of two similar uncorrelated portfolios. There is a 99.1% chance that each portfolio will produce a profit of 10 and a 0.9% chance of a loss of 100. The 1 in 100 result for each is a profit of 10 – added together gives a total profit of 20. However, when considered as a single combined portfolio the 1 in 100 result is a loss of 90, one of the original portfolios making a profit, the other a loss.

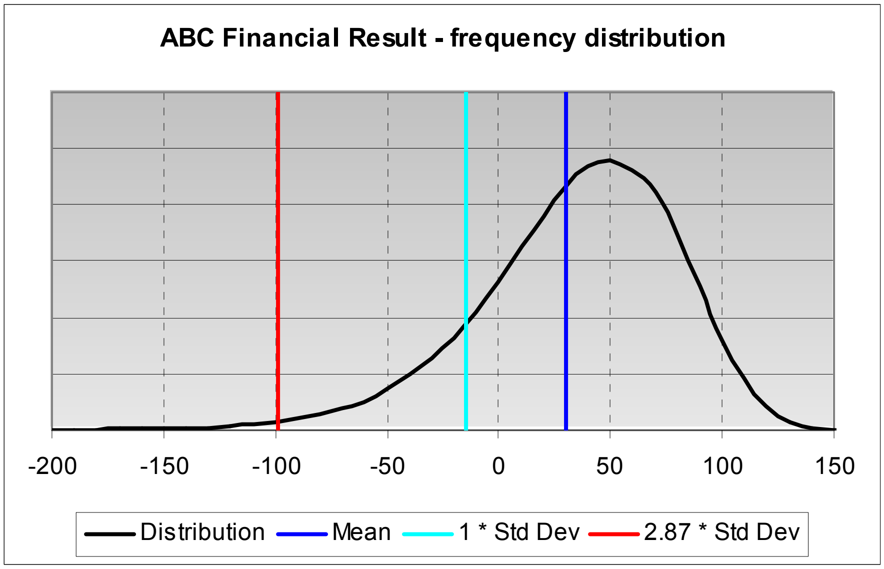
It might be thought that this sub-additivity issue is only material for fairly contrived distributions and it is certainly true that anomalies do come to the fore for relatively 'chunky' distributions (e.g. recoveries to excess of loss reinsurance contracts). However, this issue of the sum of the risk of the parts being greater than the whole has been observed with cat modelling results with the cat modelling company concerned citing this phenomenon as the cause – the effect was material.

4.1.2 Practical Evaluation

For all its technical limitations, the point value is the most widely used risk measure (or set of risk measures) in the insurance industry and is likely to remain so:

- Everyone understands it (or at least thinks they do)
- It is intuitively easy – 'what is the chance of that happening?'
- You don't need to know (or be able to estimate) more than a single point of a probability distribution. For decision-makers, such as regulators, who require their measures to be adopted industry-wide, this makes this a very persuasive argument in favour of this form of risk metric.

But what about the technical limitations? These can in part be overcome by considering more than one value using this measure – looking at two or more points on the distribution. However, the sub-additivity issue remains at best a trap for the unwary and at worst a fundamental flaw for certain applications. In particular the use of point values is not a recommended for detailed pricing techniques nor as the basis for capital allocation exercises (as the measure itself is not coherent, neither will the allocation).



Standard Deviation (and Higher Moments)

Standard deviation is a statistical measure of the spread of a distribution and is the square root of the variance, which is also known as the 2nd central moment. The variance is perhaps simply described in words as a probability-weighted sum of the squares of the deviation from the mean for all potential outcomes.

As a descriptor of the shape of a distribution, standard deviation is limited. It is possible to construct starkly different distributions which have the same standard deviation but would elicit strong variations in views regarding perceived 'risk'. The description of a distribution can be enhanced by looking at higher moments – the third and fourth moments are used to define skewness and kurtosis (peakiness). Insurance distributions are typically highly skewed making such additional metrics very relevant.

If standard deviation was to be used as a risk measure to set capital for ABC then 2.87 times the standard deviation (44.7) from the mean would give the equivalent of our previously used 1 in 100 figure.

It should be noted that the factor applied to the standard deviation will give different numbers for different distributions. For example, if ABC's financial outcome was normally distributed, the 1 in 100 result would equate to 2.33 standard deviations. For more extreme, highly skewed distributions (e.g. individual risks or ABC without reinsurance), the equivalent standard deviation multiple would be much higher.

4.2.1 Technical Evaluation

As outlined above, standard deviation is limited to giving a measure of spread. Fuller descriptions of the risk behaviour, such as the degree of skewness of a distribution, require reference to higher moments and the immediate elegance of a single metric is lost.

On the plus side, standard deviation does take into account the entire distribution, usefully giving a greater influence to more extreme outcomes. However, this is a take it or leave it situation – the decision-maker who is only interested in part of the curve has no choice.

There is also an issue that the processes of squaring, cubing or whatever makes such measures harder to manage algebraically and introduces corrupting features not present in more direct measures. One of the fallouts from this is that standard deviation fails the monotonicity requirement and is therefore not coherent. This can be illustrated in the following simple example:

Consider 2 businesses with the following sets of financial outcomes:

A: Equal likelihood of a loss or profit of 100 (mean 0, standard deviation = 141)

B: Always a loss of 100 (mean -100, standard deviation 0)

This suggests that not only is A more volatile, if capital were to be allocated based on the mean plus a factor of more than 0.7 of the standard deviation, it would attract more capital than B.

Another related unhelpful feature that there is nothing to stop a 'capital' value derived in such a way from exceeding a physical constraint such as the sum insured.

4.2.2 Practical Evaluation

As the school textbook measure of volatility, standard deviation is a widely used risk measure – non-technical decision makers do not feel overly intimidated despite its relatively technical persona. One aspect of false confidence, which has particularly helped its popularity in the banking/investment sector, is the association with the normal distribution, i.e. if you know the standard deviation then you know the value at any given percentile. The assumption of normality has received some bad press in the banking/investment sector and is demonstrably inappropriate for the highly skewed distributions typically found in insurance.

For many users a standard deviation is somewhat abstract – it doesn't tend to garner the same strength of feeling of, say, a 1 in 100 result estimate. Its abstract nature can also lead to inappropriate estimates being accepted without proper challenge. One particularly common but poorly recognised phenomenon is that the Coefficient of Variation (COV – standard deviation divided by mean) will be higher for small sample sizes, i.e. a small premium portfolio will typically have a much higher relative standard deviation than will a large premium portfolio.

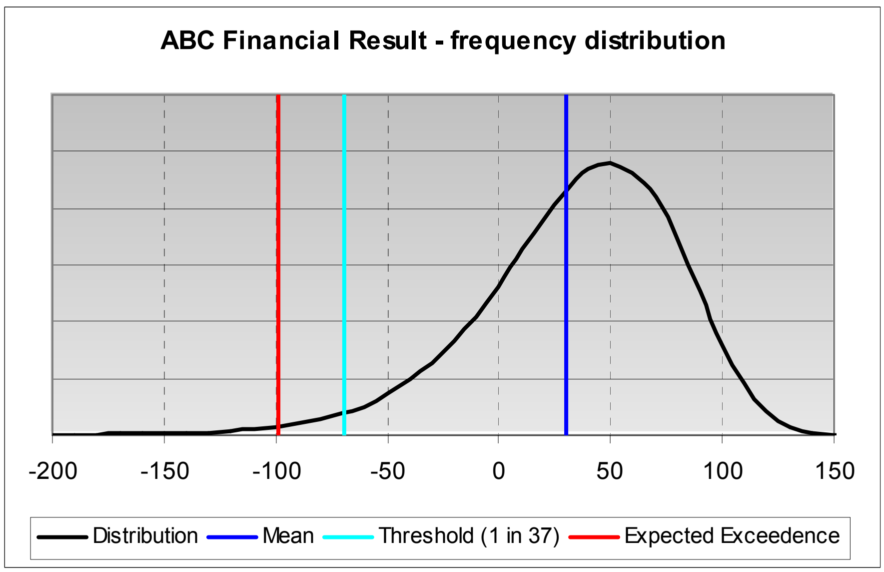
However, it does give a measure that takes in the whole of a distribution and it is fairly easy to calculate using spreadsheets and other computer software. These 'quick and dirty' benefits make it hard to ignore but it is dangerous if it ends up in ill-informed hands and/or is the only measure of risk used. Its limitations also make it a poor choice for use as a basis for more advanced decision support applications, such as capital/risk allocation.

4.3 Expected Exceedence Measures

These are a family of measures based on the expected result given the result is beyond a given threshold, i.e. the average of all the values beyond a given point of the distribution. Names include Tail Conditional Expectation (TCE), Tail Value at Risk (TVaR) and the related measure, Expected Shortfall (ES).

A key issue is how the threshold is defined and whether or not the distribution concerned is continuous. There is no problem with a percentile (or return period etc) but there can be ambiguities if a monetary value is used – if the cumulative distribution is flat for the selected monetary threshold, should all, none or part of the percentile range be included or excluded from the expected value calculation? Note that this problem also arises if a percentile is used to obtain a monetary value which in turn is used for the expected exceedence calculation.

In order to clarify the difference between measures of this form and using a single point on a distribution (4.1), it is worth comparing the interpretation of a 1 in 100 percentile/threshold. With the percentile approach the measure implies the value will be exceeded every 100 years whereas the expected exceedence measures give the expected result every 100 years, the latter being greater as it is influenced by the outcomes of more remote return periods than 1 in 100. This can be illustrated further by considering the exceedence threshold necessary for company ABC to give a result equivalent to the original 1 in 100 exceedence:



The expected result every 37 years (the average of scenarios with a result of -70 or worse) is equal to the 1 in 100 result exceedence. Obviously the return period threshold that gives an expected exceedence to match, say, a 1 in 100 exceedence, will vary according to the shape of the distribution.

A further specific example of this form of measure is Excess Tail Value at Risk (XTVaR) where the threshold is defined by the mean.

Expected Shortfall (ES) is similar but with a key difference – it represents the expected amount of the shortfall beyond the threshold, i.e. in the chart above, the Expected Exceedence less the Threshold. Expected Policyholder Deficit (EPD) is a specific case of the Expected Shortfall measure where the threshold is the company's surplus capital.

4.3.1 Technical Evaluation

This family of risk measures has strong technical properties and satisfies all the coherence criteria. The only note of caution is that of ensuring the measure is well defined if the threshold is expressed at a monetary value.

By definition the measure focuses on a single tail of the distribution. By focusing on more than a single point it gives a much richer description of 'risk' than the percentile measure. However, unlike some other measures, such as standard deviation, only part of the distribution influences the risk value – the upside characteristics, however good or bad, have no impact.

4.3.2 Practical Evaluation

Aside from their technical strengths, these measures are intuitively appealing – the expected result every x years (or beyond a given threshold) is little more difficult to understand than the omni-present return period exceedence measure. Indeed it might be suggested that this additional perspective helps challenge people's understanding of what they mean when referring to a '1 in x' result.

Its focus only on the tail of a distribution may limit applications where a more sophisticated capture of upside risk behaviour is considered desirable. However, for certain decision support requirements this is less of an issue as they are very much downside focused, e.g. capital adequacy, cat pricing etc.

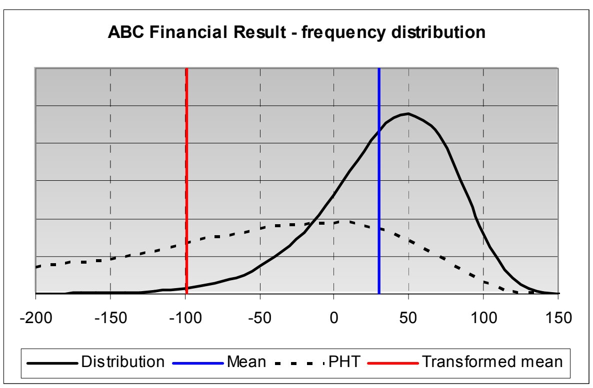
One practical issue is the calculation of these values. A more complete view of the distribution of outcomes, or at the very least the tail, is required. This is an issue for anyone who cannot express the tail of the distribution of results either algebraically or through using computer simulation. At least one regulator has indicated an interest in using the Expected Shortfall measure and it will be interesting to see how companies with less sophisticated modelling capabilities cope with calculating such a measure. For those working with probability distributions and good computer software, deriving values based on these measures is relatively straightforward.

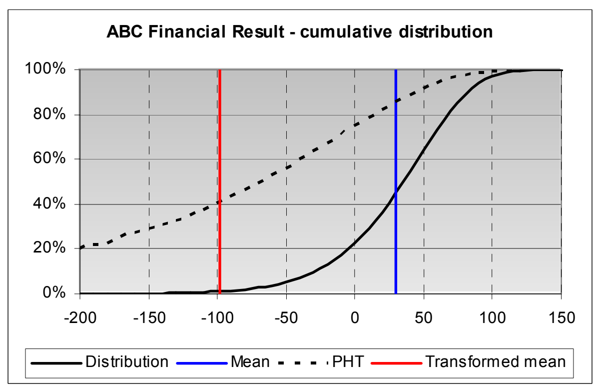
4.4 Transform Measures

In its most generic form this type of measure involves applying adjustments to different parts of the original distribution to form a new, transformed, distribution. Upside or downside outcomes that are considered to be particularly critical to a particular decision or application are given much greater weight. The mean of the resultant transformed distribution is calculated and its difference from the original mean gives a measure of risk.

From a practical calculation perspective the weightings can be affected by transforming the original percentiles (leaving the outcome values unchanged) or by being applied to the outcome values (leaving the percentiles unaltered). Where weightings are applied to the outcome values they may be standardised in order to give a mean weighting of 1.

Consider approaches involving transforming the original percentiles. **Proportional Hazards Transform** (Wang/Christofides) is one well-known example. This involves raising the cumulative distribution by a selected power. The following shows the transform undergone for company ABC using a power of approximately one fifth (the new mean matching the original 1 in 100 return period level):

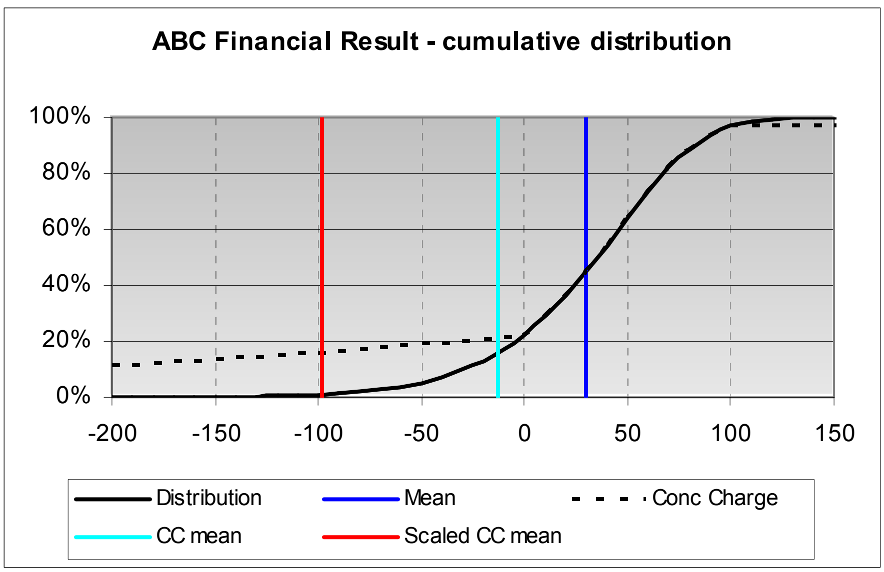




The **Wang transform** is slightly more complex, being based on a shift of the inverse Normal (0,1) transform of the distribution. A more recent version incorporates an adjustment for parameter uncertainty and has been shown to match closely the risk margins observed in catastrophe bond pricing [4].

Mango's **Concentration Charge** [1] showcases the related approach of applying weightings to different outcome values depending on how much influence it is felt each should have on the risk measurement. It is possible to be relatively crude or extremely sophisticated when deriving a Concentration Charge measure. At the same time the weightings rule can be easily understandable and accessible to the ultimate decision-maker or left as a somewhat abstract measure of 'risk'.

To illustrate the approach, consider an example using company ABC. We do not wish to dismiss any outcomes entirely so we start with a default weighting of 1 for all. Let us say that outcomes which involve losing money are considered highly undesirable – a decision is made to give these a weighting of 8. At the same time there is an interest in encouraging high upside outcomes (the appeal of the lottery effect), albeit not to the same extent as discouraging downside results – outcomes of greater than 100 are given a weighting of 4.



The resultant mean of the Concentration Charge can be scaled in order to equate to any particular less abstract risk value, in this case our 1 in 100 'capital' charge. Note that the scaling can be carried out on the absolute Concentration Charge value or relative to, say, the mean, the latter being more appropriate as a basis for risk/capital allocation – see later.

Note that the Concentration Charge is totally flexible in the way it is defined. For example it might be completely downside (or upside) focused, or attempt to introduce some sort of targeted balance between good and bad outcomes.

The expected exceedence measures are in fact related examples of this type of measure. For example with Tail Conditional Expectation / Tail Value at Risk, all outcomes beyond the prescribed threshold are given an equal positive weighting, all others a zero weighting –

albeit the zero outcomes are not included in the subsequent mean calculation (i.e. measures represent the expected result **given** the threshold is exceeded).

4.4.1 Technical Evaluation

This form of measure satisfies all the coherence properties with the potential exception of Translation Invariance. This limitation is not an issue for full distribution transforms (e.g. PHT) but may need to be considered where values have been scaled beyond the transformation of the distribution (e.g. when using non-standardised weights for the Concentration Charge). However, it is likely to be relatively easy to manage this limitation by taking care over how a measure is used.

The form of the transform/weights is effectively unlimited but the generic properties still hold in all cases.

4.4.2 Practical Evaluation

The flexibility of measures of this form is a huge benefit but also comes with the challenge of how to come up with appropriate weights for a given application. However, there is an argument that the process of thinking through what 'risk' really means in a given circumstance is important and is easily overlooked with more 'off the shelf' metrics.

There are the usual benefits of not making matters unnecessarily complicated and it is undoubtedly possible for such measures to become too abstract and intimidating for many non-technical decision makers.

Mango's as yet unpublished 'shared asset' concept is one way of introducing a rational basis to a tiered structure of 'concentration charges'. Beyond a basic capital exposure charge, a higher weight may be given to scenarios which result in capital actually being destroyed. The weight applied is increased further where capital beyond that originally notionally allocated is required.

These measures are very well suited to calculation using scenario-based simulation, particularly the clear weightings of outcomes used in a Concentration Charge. A key benefit is the notion that each outcome makes an explicit contribution towards the final measurement. This provides an elegant transparency that helps overcome user apprehensiveness and is similarly advantageous when such measures are used as the basis for risk/capital allocation.

4.5 Performance Measures

Performance ratios are perhaps the odd ones out in the company of the other measures covered in this paper in that they relate more to performance than to pure risk. As well as embodying something relating to capital (or at least some measure of downside risk) they incorporate some measure of upside performance. The typical form is to give the upside performance per unit of risk accepted, i.e. a risk adjusted measure of return.

The traditional such measure used in investment theory is known as the Sharpe Ratio (introduced by Professor William Sharpe in 1966). This is calculated by dividing the average return in excess of the risk free return by the standard deviation of the return. This gives a simple, albeit imperfect, metric by which the performance of different assets can be compared.

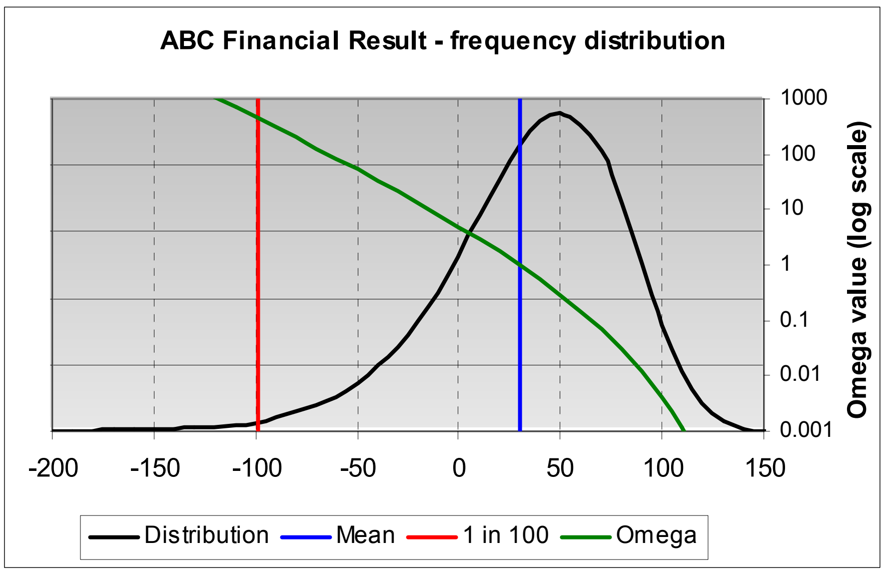
Clearly similar measures can be constructed for insurance applications but it is important to be mindful of the strengths and limitations of the component parts – as highlighted in this paper. The Sharpe Ratio for insurance equates to the gradient on a mean / standard deviation risk reward chart – for any given point on the chart how much additional profit can be made for accepting a small amount of extra risk?

Given the earlier comments regarding the limitations in using standard deviation to measure risk what alternatives exist?

One worth noting is Ruhm's **Risk Coverage Ratio (RCR)** [5] (also known as **R2R – Reward To Risk**). The return (on a present value basis) is as for Sharpe but the denominator is replaced by the expected downside result multiplied by the probability of a downside result (note downside in this context refers to any negative outcome - the denominator is in effect a measure of the contingent use of capital). The ratio's intuitive meaning is how many times the risk is "covered" by the expected return, hence the name.

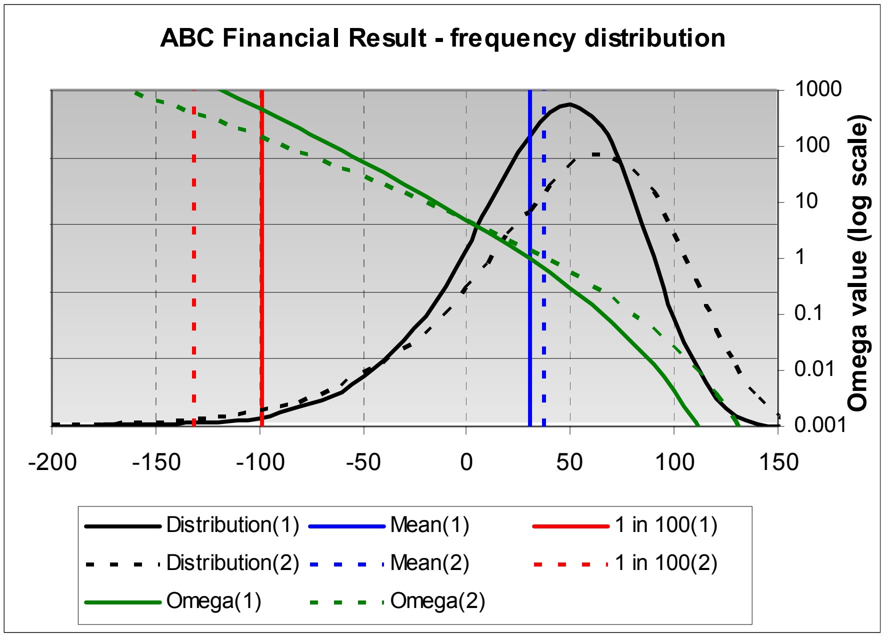
Another performance ratio worth looking at is Keating and Shadwicks' **Omega** function [6]. This is simply, for a given threshold, the expected upside (i.e. the expected result given it is better than the threshold) divided by the expected downside. Put another way the numerator can be thought of as the chance of winning multiplied by the expected amount of a win once it happens. Likewise the denominator, as for the Risk Coverage Ratio, is the chance of losing multiplied by the expected loss in such a scenario.

So what does Omega look like in practice and how should it be interpreted? A key issue is what threshold to select. This could be achieving plan, making a profit or be related to some other specific objective, and is likely to vary according to the application / decision being made. The Omega value will always be positive (numerator and denominator are based on deviations away from the threshold) and will be greater or less than 1 depending on whether the threshold is worse or better than the mean. This can be illustrated by looking at the full range of Omega values (plotted on a separate log scale) for company ABC based on all possible threshold values:



Note that the x axis represents the threshold for the plot of the Omega value 'curve'.

These values are somewhat abstract when looking for an immediate interpretation for a given threshold. However, they become more meaningful when it comes to comparing alternative risks / strategies. An alternative strategy might be considered for ABC which has a higher profit expectation (mean losses for each of A, B and C reduced by 2.5) but is more volatile (each of the three loss standard deviations increased by 5). The additional statistics (dashed) are included in the following chart:



The new distribution has a higher mean but a significantly poorer 1 in 100 result. We can see how the Omega measure interprets these differences. The lines cross just above the break even point (1.33) and if the selected threshold is below this level, the original distribution gives higher Omega values and is therefore valued as being preferable. Alternatively, for higher thresholds the new strategy is viewed as more appealing.

It should be noted that by plotting the full range of Omega values it is possible to achieve a more balanced view of the perspectives of different stakeholders. It is too simplistic to suggest, for example, that a regulator's sole interest is in the probability of ruin or that management is only concerned with nearer term bonus or career risk type return periods. The Omega measure gives a perspective of the entire distribution of outcomes based on any number of selected threshold view points.

As a way of bringing together the implications from the full range of values of the Omega function, the Financial Development Centre has developed 'Omega metrics'. These incorporate the slope of the log of the Omega function and balance the right or left bias of the distribution, thereby penalising fat downside tails and rewarding fat upside tails. This captures an important natural risk aversion dynamic within a single statistic.

4.5.1 Technical Evaluation

The technical properties of any given ratio will depend on its construct and the purposes for which it might be used.

Both the Risk Coverage Ratio and Omega function are constructed from measures which themselves have robust characteristics. They reflect both upside and downside behaviours in a single value, with the Omega measure having the additional flexibility as to the threshold by which 'upside' and 'downside' are defined.

Considering ratios in the context of the coherence properties outlined in section 3 is probably inappropriate as these are inherently performance rather than risk measures. However, they can still be used as a basis for capital allocation – see 5.6.

4.5.2 Practical Evaluation

Despite their relatively simple foundations, ratios are typically more complex to explain than other measures. The value of the measure does not directly represent a loss or profit amount, but rather a semi-abstract hybrid. However, they bring together the upside and downside characteristics of distributions in an elegant and neat way, and have a valuable role to play in the ongoing challenge of adequately evaluating risk reward tradeoffs.

Again these measures are very well suited to calculation using scenario-based simulation.

4.6 Other Measures

Other measures probably fall into two categories.

The first would be specific and/or hybrid versions from the five main categories highlighted above. For example there might a very specific transform or ratio which is viewed as meeting the requirements of a particular decision support application. Alternative different risk measurement concepts might be combined – perhaps the expected exceedence of a transformed distribution?

It is inappropriate to attempt to identify every conceivable form that an actual risk measure might take. However, it is suggested that it will be based on one or more of the main risk measurement concepts outlined, and thereby share the corresponding strengths and limitations.

The second category might not be considered risk measures at all but for the fact they are, on occasion, used to measure risk. They might better be classified as 'measures' without the 'risk' prefix. The quantification of anything can subsequently end up being associated with risk – weight, size, length, speed, power etc.

In insurance it is common to see real world items such as premium, sum insured, wageroll etc all being used a proxies for exposure and hence risk. Such measures do have an important role but their limitations in terms of risk measurement can be considerable.

5. RISK / CAPITAL ALLOCATION ALTERNATIVES

This section considers ways of allocating the risk measurement values derived from section 4 across contributing sub-portfolios. If a risk measurement value happens to equate to some form of capital value then these can be considered to be capital allocation methods.

The methods are again illustrated using company ABC. A number of the risk measures are used in each case (all equating to the 1 in 100 return period benchmark used throughout section 4). For illustrative purposes we will attempt to allocate the 128.4 shortfall below the mean rather than the absolute loss of 98.4. Note that any attempt to allocate the shortfall below an absolute value encounters a further difficulty, as to what threshold to use for each sub-portfolio (less of a problem if 0 used). Allocating the overall threshold between portfolios can be as awkward as the main allocation challenge.

A recap of the measures and tolerances used follows:

- 1^{st} percentile
- Standard Deviation * 2.8706
- $TCE_{2.728\%}$ (expected result beyond 2.728%ile)
- $PHT_{5.1375}$ (Proportional Hazards Transform with power 5.1375)
- Scaled Concentration Charge (relative weights: less than 0 = 8, over 100 = 4, otherwise 1)

As was highlighted at the end of section 3, an allocation method can only be coherent if the risk measure used for the allocation is coherent. This that means certain of these risk measures will be a poor choice in practice.

It should be noted that there is no guarantee of avoiding one or more sub-portfolios getting a negative risk allocation, i.e. an allocation with a different sign to the overall risk measurement. Consider our company ABC example where we are modelling the overall financial result. It may be that one particular sub-segment is relatively small, highly profitable, not overly volatile and non-correlating with the overall company risk. In such circumstances it is possible that the premium may be more than likely an adequate buffer against any adverse claims behaviour. Would a negative allocation be unreasonable in such circumstances? The key issue is not that it can occur but how this phenomenon should be managed in the context of a particular decision support application.

The recommended approach is the Aumann-Shapley method which is covered in 5.4. However, the preceding methods highlight some of the common weakness and illustrated how the thinking behind Aumann-Shapley values evolved.

The Covariance Share method (5.5) is also worth noting as is the potential to use performance ratios, mentioned in 5.6.

5.1 Independent "First In"

The independent method estimates the risk for each sub-portfolio as if it were a stand-alone business unit. In practice the sub-portfolios are unlikely to be perfectly correlated so there should be some diversification benefit at the overall portfolio level (i.e. the sum of the parts is likely to be greater than the whole). This diversification benefit is shared pro-rata according to the independent risk values.

The following table shows the calculations based on the 1st percentile risk measure for ABC:

1st percentile	Actual	%	Scaled
A	60.1	33.3%	42.8
B	60.1	33.3%	42.8
C	60.1	33.3%	42.8
Total	180.3	100.0%	128.4

Note that the sum of the parts is significantly higher than the 128.4 value for ABC as a whole but the standalone values are scaled accordingly to give our first allocation candidate. This approach can be repeated using other risk measures:

2.87 * Std Dev	Actual	%	Scaled
A	57.4	33.3%	42.8
B	57.4	33.3%	42.8
C	57.4	33.3%	42.8
Total	172.3	100.0%	128.4

$TCE_{2.728\%ile}$	Actual	%	Scaled
A	60.0	33.3%	42.8
B	60.0	33.3%	42.8
C	60.0	33.3%	42.8
Total	179.9	100.0%	128.4

This is equivalent to allocating risk as a fixed premium percentage – it is the simplest approach, is easy to understand, and always generates positive capital requirements for every business line. However it does not penalise highly correlated portfolios nor reward those which give rise to an overall diversification effect. This is not a coherent allocation method (failing the 'no undercut' criterion) and is not recommended.

5.2 Marginal "Last In"

The marginal method considers the risk of a portfolio, with and without each sub-portfolio, for which the allocation is to be undertaken – the difference being the 'marginal' risk. The sum of the marginal risk values will be less than the overall value unless all risks are perfectly correlated – the marginal values can be scaled accordingly.

The calculations for ABC are as follows:

1st percentile	Excluding portfolio	Marginal impact	%	Scaled
A	79.6 (BC)	48.8 (ABC-BC)	46.1%	59.2
B	79.7 (AC)	48.7 (ABC-AC)	46.1%	59.2
C	120.2 (AB)	8.2 (ABC-AB)	7.8%	10.0
Total		105.8	100.0%	128.4

2.87 * Std Dev	Excluding portfolio	Marginal impact	%	Scaled
A	81.2 (BC)	47.2 (ABC-BC)	43.2%	56.1
B	81.2 (BC)	47.2 (ABC-AC)	43.2%	56.1
C	114.8 (AB)	13.6 (ABC-AB)	12.6%	16.2
Total		107.9	100.0%	128.4

$TCE_{2.728\%ile}$	Excluding portfolio	Marginal impact	%	Scaled
A	79.2 (BC)	49.2 (ABC-BC)	46.0%	59.1
B	79.2 (BC)	49.2 (ABC-AC)	46.0%	59.1
C	119.9 (AB)	8.5 (ABC-AB)	7.9%	10.2
Total		106.8	100.0%	128.4

This approach goes to the other extreme. Uncorrelated sub-portfolios are over rewarded, leaving a relatively harsh allocation for correlating sub-portfolios. Again this is not a coherent allocation method (again failing the 'no undercut' criterion) and is not recommended.

The following methods give rise to more equitable allocations which lie between the extremes of the independent and marginal approaches.

5.3 Shapley

Game Theory is a relatively young branch of mathematical analysis (foundations in the 1940s) developed to study decision-making in conflict situations. Practical applications can be found in many fields including economics and finance. The relevance to risk allocation is obvious – a sub-portfolio benefits from being part of a larger diversified portfolio but, at the same time gives a diversifying benefit to other portfolios. The inherent tension in this set up gives rise to the challenge of finding an equitable and stable way of sharing the overall diversification benefit across each sub-portfolio.

Denault [3] considered how Game Theory concepts could be applied to this problem, in particular Shapley and Aumann-Shapley values (see 5.4).

Lloyd Shapley introduced the concept of the "Shapley Value" in 1953 [7] as a stable solution to coalitional games involving any number of players. The key limitation in applying this to risk allocation problems is the issue of having a whole number of players, i.e. a sub-portfolio is treated as a whole – more on this later.

The calculation of Shapley values is a natural extension of the independent and marginal methods already discussed, and is based on the average of the "1st in", last in" and all the intermediate "ins" This is best illustrated by looking at the calculations for ABC where the "2nd in" is the only intermediate value to consider:

1st percentile	"1st in"	Average "2nd in"	"Last in"	Average
A	60.1	39.8	48.8	49.6
B	60.1	39.8	48.7	49.6
C	60.1	19.5	8.2	29.3
Total	180.3	99.2	105.8	128.4

"2nd in" calculations	
60.1 (AB-B)	19.6 (AC-C)
60.1 (AB-A)	19.5 (BC-C)
19.6 (AC-A)	19.5 (BC-B)

2.87 * Std Dev	"1st in"	Average "2nd in"	"Last in"	Average
A	57.4	40.6	47.2	57.4
B	57.4	40.6	47.2	57.4
C	57.4	23.8	13.6	31.6
Total	172.3	105.0	107.9	128.4

"2nd in" calculations	
57.4 (AB-B)	23.8 (AC-C)
57.4 (AB-A)	23.8 (BC-C)
23.8 (AC-A)	23.8 (BC-B)

$TCE_{2.728\%ile}$	"1st in"	Average "2nd in"	"Last in"	Average
A	60.0	39.6	49.2	49.6
B	60.0	39.6	49.2	49.6
C	60.0	19.2	8.5	29.2
Total	179.9	98.5	106.8	128.4

"2nd in" calculations	
59.9 (AB-B)	19.3 (AC-C)
60.0 (AB-A)	19.3 (BC-C)
19.2 (AC-A)	19.2 (BC-B)

No scaling is required – the parts naturally add up to match the whole. So long as the risk measure used is strongly sub-additive (see Denault) in addition to satisfying the general coherence properties (as outlined in 3.1), then this method of allocation is coherent (3.2). Note that only the latter measure, $TCE_{2.728\%ile}$, is coherent amongst the examples shown.

Shapley values are undoubtedly elegant but have two important practical failings.

First, they are computationally challenging – an allocation based on 4 sub-portfolios requires calculations for 15 combinations and 5 requires 31 combinations etc. This quickly becomes impractical.

Second, we have the issue regarding whole number of players. If, say, A was split into two and the above calculations were performed on the 4 sub-portfolios, the Shapley values for B and C would change! This is a poor property and the next method moves on to looking at games involving fractional players.

Note that Covariance Share (see 5.5) is related to Shapley Values. This uses variance as the underlying risk measure (note not coherent) but overcomes the computational problems highlighted above.

5.4 Aumann-Shapley

Aumann and Shapley extended the concept of Shapley values to non-atomic (fractional) games in their original book [8] published in 1974. The result was called the Aumann-Shapley value. In words, the Aumann-Shapley value represents the rate of increase in the risk /capital allocation, i.e. how much additional overall risk comes from a sub-portfolio, for a tiny increase in size.

It is interesting to note that a number of academics and practitioners have come separately to the same conclusion – that this is the most appropriate and, as we will see, practical method for risk allocation. The Myers-Read methodology is in effect based on calculating Aumann-Shapley values using a probability of ruin risk measure.

Aumann-Shapley values are particularly easy to calculate using simulation techniques. This has been formalised by the Ruhm-Mango-Kreps algorithm (RMK) [9]. This considers the selected risk measure in the form of a weight applied to the outcome for each trial (as explicitly happens for Mango's concentration charge). It is then possible to calculate the risk-weighted average, see how much it differs from the original average and, most importantly, see what contribution was made to this difference by each sub-portfolio.

Allocations have been calculated for ABC based on four of the risk measures covered in section 4:

	1st percentile (0.57%-1.5%)	**$TCE_{2.728\%ile}$**	**$PHT_{5.1375}$**	**Concentration Charge**
A	54.7	54.7	58.9	63.3
B	54.8	54.7	58.3	63.3
C	19.0	19.0	11.2	1.8
Total	128.4	128.4	128.4	128.4

The 1st percentile calculation follows the approach taken by Ruhm and Mango in their paper on implementing Myers-Read using simulation [10]. A collar is taken round the point estimate so as to capture a reasonable number of trials from which to base the calculation of the contribution made by A, B and C to the overall measure. This collar might simply be +/- 0.5% (the odd range here is somewhat pedantic and used in order to ensure an exact match up with the 1 in 100 figure).

The Tail Conditional Expectation and Proportional Hazards Transform measures are both coherent and using these with the Aumann-Shapley values ensures a coherent allocation. The TCE measure is particularly intuitively appealing when used with Aumann-Shapley: what is the expected downside result every x years and, when things go wrong, whom do we expect to have caused it?

The Concentration Charge results are interesting and illustrate the importance of thinking through the weights carefully, and understanding the subsequent results. The average weighted outcome for C is little different from the original mean. As C is uncorrelated with A and B, the large downside weights applied at ABC level presumably come from scenarios dominated by A and B. In such circumstances C is having as many upside outcomes as downside – the weighting effect is therefore largely averaged out.

5.5 Covariance Share

Mango's 1998 paper 'An application of Game Theory: Property Catastrophe Risk Load' [11] examines the use of standard deviation and variance in calculating risk loads for property catastrophe risks. In considering the additivity problems associated with these measures Mango turns to Game Theory and, in particular, Shapley values. He notes that the computational complexity of calculating Shapley values is overcome when using variance.

In the case of company ABC, the Shapley value for A is simply the variance of the outcomes for A plus the covariance of the outcomes for A and BC (i.e. ABC excluding A).

Note that the increase in the variance cause by adding A to BC (the marginal variance) is the variance for A plus twice the covariance of A and ABC excluding A. However when you add the marginal variance for A to the marginal variance for BC this overshoots the overall variance of ABC by twice the covariance of A and BC:

Overall variance of ABC	Var(A) + Var(BC) + 2*Cov(A,BC)
Marginal variance of A	Var(A) + 2*Cov(A,BC)

Marginal variance of BC Var(BC) + 2*Cov(A,BC)

The Shapley value effectively shares the '2*Cov(A,BC)' term equally between A and BC. Applying this approach to company ABC gives the following results:

	Variance	Covariance with rest	Variance + Covariance	Scaled Total
A	400.2	400.0	800.2	51.4
B	400.3	400.1	800.3	51.4
C	399.9	0.5	400.4	25.7
Sum A, B, C	1,200.3	800.5	2,000.9	128.4
ABC	2,000.9			

However, the relative contribution to the covariance term may be far from equal, suggesting a more equitable sharing may be appropriate. The Covariance Share method gives a more generic approach to sharing the covariance term. Mango's paper focuses on catastrophe events and suggests the overall share is based on each sub-portfolio's share of the overall loss for each event.

Note that the limitations of using variance as a risk measure will be an issue in certain practical situations (the measure is not coherent, being super-additive rather the sub-additive). Perhaps at least as important is the fact that variance (and covariance) will be a relatively abstract concept for many end users which may prove a barrier to acceptance. However the computational elegance makes Covariance Share an interesting method to consider where full Aumann-Shapley calculations are not feasible.

5.5 Others

Performance ratios can be used as a basis for capital allocation, as illustrated by the '**Constant R2R**' or 'Constant Risk Coverage Ratio' method (see 4.5). This involves adjusting the profit margin in each sub-portfolio until the R2R ratios are equivalent. This results in an allocation of the required profit margin. A similar principle can be used with Omega values (for a given threshold).

Also worth reviewing is Lemaire's paper on applying Game Theory to cost allocation problems [12]. This may not be an immediately obvious place to look for risk allocation alternatives. However, cost and risk allocation problems have a lot in common and Lemaire gives an eloquent explanation of why many of the more obvious allocation possibilities are inadequate. Examples include (taking insurance equivalents):

- Premium
- Policy count
- Sub-portfolio count

Beyond methods covered already in this paper, more exotic concepts include:

- The Nucleolus (David Schmeidler, 1969)
- The Proportional Nucleolus (Young et al, 1980)
- The Disruptive Nucleolus (Littlechild and Vmdya, 1976, and Michener, Yuen and Sakuraz, 1981)

These are all optimisation methods with variations in the way the constraints are articulated. The difficulties in computing solutions for these methods limit their practical value. For the enthusiast, the Proportional Nucleolus is the pick of the bunch!

6. RISK MEASUREMENT IN PRACTICE

Some of the measures and methods described have been derived with specific applications in mind, ranging from enterprise-wide financial modelling to individual insurance/reinsurance pricing. It is useful not to pigeonhole any particular concept into a single application and form, as this can discourage thinking as to how similar ideas might be applied in other areas.

It is important to consider all the methods illustrated as being part of a risk measurement tool kit. Sometimes a single simple metric will do; other times a more exotic combination of ideas may be appropriate. The issue of what risk tolerance levels, thresholds, weightings etc to use with each of the concepts shown, and in what circumstances, has deliberately not been commented upon as this requires a full appreciation of the circumstances in which each risk measure is being used.

Selecting the right tools and ensuring they are appropriately calibrated for the purpose for which they are being used is a critical part of the implementation process. Likewise it is essential that, for all the great theory, intelligent judgements are made regarding how such metrics influence decisions:

- How well does the measure reflect the risk/performance in the context of the decision being made?
- Do the decision-makers / interpreters of the risk measurements understand the information presented?
- What about limitations in the underlying risk model, either risk behaviours poorly captured or not included in the model in the first place?!

Acknowledgment

Thanks are given to all the people whose work has been referred to in compiling this paper, particularly Don Mango for reviewing an earlier version. The contributions of colleagues at Benfield are also noted and appreciated, particularly Colin Kerley, David Simmons, Kirsty Howard-Price, Paul Maitland, Nathan Schwartz and John Pelican.

REFERENCES

[1] Mango, Donald, 1998 CAS Ratemaking Call Paper "The Concentration Charge: Reflecting Property Catastrophe Exposure Accumulation in Rates"

[2] Artzner, Ph., F. Delbaen, J.-M. Eber, D. Heath, "Coherent measures of risk", Mathematical Finance, vol. 9, no. 3 (1999), 203–228.

[3] Denault, Michel, "Coherent allocation of risk capital", 2001 (update of 1999 original)

[4] Wang, Shaun, "Cat Bond Pricing Using Probability Transforms", published in Geneva Papers, 2004

[5] Ruhm, David, 2001 ASTIN Colloquium "Risk Coverage Ratio: A Leverage-Independent Method of Pricing based on Distribution of Return"

[6] Keating, Con, William F. Shadwick, "A Universal Performance Measure", The Finance Development Centre, London, 2002

[7] Shapley, Lloyd, "A value for n-person games", Contributions to the theory of games, volume II (Annals of mathematics studies, 28, Tucker and Luce, eds.), 1953, Princeton University Press, Princeton, New Jersey

[8] Aumann, Robert, Lloyd Shapley, "Values of Non-Atomic Games", 1974, Princeton University Press, Princeton, New Jersey.

[9] Ruhm, David, Donald Mango, "A Risk Charge Calculation Based on Conditional Probability", The 2003 Bowles Symposium

[10] Ruhm, David, Donald Mango, "A Method of Implementing Myers-Read Capital Allocation in Simulation", 2003

[11] Mango, Donald, "An application of Game Theory: Property Catastrophe Risk Load", 1998

[12] Lemaire, Jean, "An application of Game Theory: Cost Allocation", ASTIN, 14, 1984, 61-81

Abbreviations and notations

EPD, Expected Policyholder Deficit
PHT, Proportional Hazards Transform
RCR, Risk Coverage Ratio
TVaR, Tail Value at Risk (= TCE)
XTVaR, Excess Tail Value at Risk
ES, Expected Shortfall
R2R, Reward to Risk
TCE, Tail Conditional Expectation
VaR, Value at Risk

Biography

Paul Kaye is a London based member of the ReMetrics team within Benfield, a leading reinsurance intermediary. His role focuses on non-transactional consultancy services including reinsurance strategy reviews, Risk Based Capital exercises, business plan evaluation and other general financial modelling assignments.

Prior to joining Benfield in 2000, he worked for Commercial Union/CGU for 13 years in the UK, holding a number of front-line and head office positions. He has worked on a number of short and longer term consultancy projects with a number of insurance and reinsurance customers around the world.

He holds a Mathematics degree from Nottingham University and is also an Associate of the Chartered Insurance Institute. He was a member of the Insurance Institute of London's Advanced Study Group on the Future Developments of Excess of Loss Reinsurance, which published its report in 2000.