Loss reserving using loss aversion functions

Weihao Choo and Piet de Jong

Working Paper 09-07
The Centre for Financial Risk brings together researchers in the Faculty of Business & Economics on uncertainty in capital markets. It has two strands. One strand investigates the nature and management of financial risks faced by institutions, including banks and insurance companies, using techniques from statistics and actuarial science. It is directed by Associate Professor Ken Siu. The other strand investigates the nature and management of financial risks faced by households and by the economy as a whole, using techniques from economics and econometrics. It is directed by Professor Geoffrey Kingston. The co-directors promote research into financial risk, and the exchange of ideas and techniques between academics and practitioners.
Loss reserving using loss aversion functions

Weihao Choo, Piet de Jong

Abstract
This article discusses the determination of risk capital based on “aversion” functions. Aversion functions weigh different outcomes according to perceived severity. Many practical and popular risk measures are usefully viewed in terms of aversion functions including those arising from distortion operators and risk margin loadings. The approach of this paper builds on, unifies, and extends existing disparate approaches discussed in the literature. Analytical and computer-generated illustrations are given as well as suggestions for the practical determination of aversion functions.

Key words: Distortion operators; Loss aversion; Risk measure; Percentile rank aversion; Standard deviation principle; Premium loading, Expected Maximum Loss.

1. Introduction
Loss reserving assigns risk capital where the loss outcome is as yet unknown. In many practical contexts the number of risks is small and/or risks are dependent. Hence there is limited scope for diversification and capital must be set aside to cover the expected value plus a risk margin. In this article the amount of risk capital is determined using a “risk aversion” function.

A formula translating a random variable to a number is often called a “risk measure.” Risk measures are used in insurance banking and finance, for premium calculation and option valuation. The terminology “reserving methods” is used in this article to emphasize the specific interest in insurance reserving.

This article presents a unified treatment of risk measures based on weighted premiums (Venter, 1991; Furman and Zitikis, 2008):

\[ R(x) \equiv \mathbb{E}\{x\psi(x)\} , \]

where \( x \geq 0 \) is a random loss and \( \psi \) is an appropriately defined “aversion” function. The risk measure (1) appears in various guises – see for example the recent articles Tsanakas (2008) and Furman and Zitikis (2008) which point to the
relationship of (1) to distortion operators in the insurance premium literature introduced by Denneberg (1990) and Wang (1996).

Our treatment unifies and extends the existing literature. The existing literature suffers from limited cross referencing and attribution even though approaches are closely related. For example expressions in the works of Wang (1995), Wang (1996), Wirch and Hardy (1999), Acerbi (2001) Acerbi (2002) and the recent works Tsanakas (2008)) and Furman and Zitikis (2008)) are weighted premiums in the sense of (1). Other close connections of (1) to suggestions appearing elsewhere in the literature are displayed in the sections below.

The further sections of this article are structured as follows. The next section imposes appropriate properties of \( \psi \). Section 3 displays the connection to expectations under adjusted densities and risk margin calculations. Section 4 discusses the specialization \( \psi = \phi \circ F \), called percentile rank aversion, where \( F \) is the distribution of the loss \( x \) and \( \phi \) is an appropriately defined function. Section 5 discusses the connection to distortion operators while §6 considers the issue of coherence. Risk margins are further explored in §7 while §8 discusses Expected Maximum Loss and its variants. Section 9 deals with further generalizations based on mixing distributions while §10 discusses further variants. Illustrative examples are given in §11 while a practical method for the determination of \( \psi \) is given in §12. A final section gives conclusions.

2. Aversion function properties

The import of \( \psi \) in (1) is to adjust losses \( x \) by an “aversion factor” \( \psi(x) \). Generally, larger losses \( x \) are assigned larger aversions \( \psi(x) \). Special cases of the above setup are \( \psi \equiv 1 \), in which case \( R(x) = E(x) \), and the Value–at–Risk (VaR) where \( \psi \) is the Dirac delta function picking out the appropriate percentile of \( x \).

Three conditions are imposed on \( \psi \). First, \( \psi \geq 0 \) implying \( x \psi(x) \geq 0 \) and \( R(x) \geq 0 \). Second, \( E(\psi) = 1 \) and hence on average, there is no aversion adjustment. Third, \( \psi' > 0 \) where \( \psi' \) is the derivative of \( \psi \) with respect to \( x \). Hence aversion increases with \( x \), a reasonable practical requirement. The condition \( \phi' > 0 \) is not satisfied for example in the case of VaR where \( \psi \) is the Dirac delta function.

If \( \psi \) is the aversion function depending on a random parameter \( \theta \) and \( \psi = E_\theta(\psi_\theta) \), then

\[
R(x) \equiv E(x\psi) = E\{xE_\theta(\psi_\theta)\} = E_\theta\{E_x(x\psi_\theta)\} = E_\theta\{R_\theta(x)\},
\]

with the last equation serving to define \( R_\theta(x) \). Hence averages or mixtures of aversion functions yield new aversion functions and associated risk measures. This is exploited in §9. Throughout the further sections of this paper the dependence of \( \psi \) on \( x \) is left implicit.
3. Alternative representations

If \( x \) has density \( f \) then the first two conditions on \( \psi \) stated in the previous section imply \( f(x) \equiv \psi(x)f(x) \) is a valid density since \( f(x) \geq 0 \) and

\[
1 = E(\psi) = \int \psi(x)f(x)dx = \int \tilde{f}(x)dx .
\]

Hence under these conditions the risk measure in (1) is an expectation under the modified density \( \tilde{f} \):

\[
R(x) = \int_0^\infty x\psi(x)f(x)dx = \int_0^\infty x\tilde{f}(x)dx . \tag{2}
\]

Risk measures which are expectations under a modified density occur in the work of Venter (1991), in the context of premium calculation, and Ruhm et al. (2003) and Kreps (2005) in the context of capital allocation.

A second way of rewriting \( R(x) \) in (1) is to use \( E(\psi) = 1 \) to derive

\[
R(x) = E(x)E(\psi) + \text{cov}(x, \psi) = E(x) + \sigma_x \sigma_\psi \text{cor}(x, \psi) . \tag{3}
\]

where \( \text{cov} \) and \( \text{cor} \) denote covariance and correlation, respectively, and \( \sigma_x \) and \( \sigma_\psi \) are the standard deviation of \( x \) and \( \psi \equiv \psi(x) \), respectively. Thus \( R(x) \) is the expected loss plus the covariance between \( x \) and the aversion adjustment \( \psi(x) \), which closely relates to the “covariance” principle (Venter, 1991). The final form is a modification of the “standard deviation” principle (Embrechts and Center, 1996). The final term on the right in (3) is called the risk margin or risk loading as it is the excess of \( R(x) \) over \( E(x) \). From (3) it follows

\[
R^*(x) \equiv \frac{R(x) - E(x)}{\sigma_x} = \sigma_\psi \text{cor}(x, \psi) , \tag{4}
\]

called the standardized risk margin, that is the risk margins in units of standard deviation of the loss \( x \). The standardized risk margin is independent of units of measurement and equals \( \sigma_\psi \) times the correlation between the loss and aversion \( \psi \). Risk margins are further discussed in §7.

Writing \( R(x) \) as in (1), (2) or (3) make for flexible and practical interpretations. Formulating risk through \( \psi \) emphasizes that risk measurement is subjective. Subjectivity is incorporated by taking expectations under aversion adjusted losses or aversion adjusted likelihood. Since \( E(\psi) = 1 \), there is no aversion to the loss on an ex-ante basis, i.e. before the actual loss is known. Alternatively there is no aversion to the loss distribution as a whole. The function \( \psi(x) \geq 0 \) reduces or amplifies different losses \( x \) or their likelihood, in such a way that, on average, the amplification or reduction is 1: \( E(\psi) = 1 \). Relative aversion to an outcome \( x \) occurs if \( \psi(x) > 1 \) resulting in amplification, and vice versa. Neutrality to a particular outcome occurs if \( \psi(x) = 1 \).

Adjustment of losses or their likelihood using aversion functions is justified as follows. Aversion to particular outcomes indicates a belief that the severity or
likelihood of those outcomes are higher than actually the case. And vice versa: outcomes to which there is little aversion tend are “weighed down.” Measuring risk according to (1) implicitly assumes that the loss severity is $x\psi(x)$ rather than $x$, or that the loss is modelled by the modified density $\tilde{f}$ rather than $f$. The latter is closely related to the approach of distorted risk operators discussed in §5.

4. Loss reserving based on percentile rank aversion

If $\psi = 1$ then $R(x) = E(x)$ and the modified density is the original density: $\tilde{f}(x) = f(x)$, while $\text{cor}(x, \psi) = 0$ implying the risk margin is zero. In this case there is an equal aversion to all losses $x$ and in particular all percentile ranks $u \equiv F(x)$ since $\psi = \phi \circ F$ where $\phi = 1$. This is a trite example of aversion depending only on the percentile rank of the outcome: $\psi = \phi \circ F$.

Another example of percentile rank aversion is the VaR $F^{-}(u)$ where $\psi$ is the Dirac delta function, having a single spike at $F(x) = \pi$, written as $\psi(x) = \{F(x) = \pi\}$. In this case $\psi = \phi \circ F$ with $\phi(u) = (u = \pi)$. A richer example is Conditional Tail Expectation (CTE) also called expected shortfall (Acerbi, 2002; McNeil et al., 2005), where $\psi$ is a step function, equal to 0 if $F(x) \leq \pi$ and $1/(1 - \pi)$ otherwise, written as

$$\psi(x) = \frac{\{F(x) > \pi\}}{1 - \pi}.$$ 

This form implies $\psi = \phi \circ F$ with $\phi(u) = (u > \pi)/(1 - \pi)$. Note $E(\psi) = 1$ and

$$R(x) = E\{x|F(x) > \pi\} = \frac{1}{1 - \pi} \int_{F^{-}(\pi)}^{\infty} x dF(x).$$

$$= \frac{1}{1 - \pi} \int_{-\pi}^{1} F^{-}(u) du = F^{-}(\pi) + \frac{1}{1 - \pi} \int_{F^{-}(\pi)}^{\infty} \{1 - F(x)\} dx.$$

The equalities assume sufficient regularity on $F$: for details see Acerbi and Tasche (2002) and references therein. Similar to VaR, the aversion function depends on $x$ only through $F$.

Any risk measure (1) with $\psi = \phi \circ F$ can be written as weighted average of VaRs (Acerbi, 2002) or weighted average of CTE’s. Both results are established by change of variable in integration. For the first result

$$R(x) = \int_{0}^{\infty} x \phi\{F(x)\} dF(x) = \int_{0}^{1} \phi(u) F^{-}(u) du,$$

and hence $R(x)$ is a weighted average of VaRs $F^{-}(u)$, using the weighting function $\phi(u)$ with $\phi \geq 0$ and $E\{\phi(u)\} = 1$. Hence $\phi(u)$ indicates the aversion to different VaR’s. Acerbi (2002) calls $R(x)$ as defined in (5) the “spectral risk measure” generated by $\phi$. 

4
Further, if $\phi$ has derivative $\phi' \geq 0$, then

$$R(x) = \int_0^1 (1 - u)\phi'(u)E\{x|F(x) > u\}du \equiv E[\{x|F(x) > u\}]$$ \hspace{1cm} (6)$$

where it is assumed $\phi(0) = 0$. Hence $R(x)$ is a weighted average of CTEs using weights $(1 - u)\phi'(u)$ where,

$$E\{(1 - u)\phi'(u)\} = \int_0^1 (1 - u)\phi'(u)du = E\{\phi(u)\} = 1.$$  

Hence $E$ in (6) is an expectation with respect to $(1 - u)\phi'(u)$. If $\phi$ is not differentiable then a mixture of continuous and discrete averaging is employed.

Aversion based on percentile rank implies

$$\text{cor}(x, \psi) = \text{cor}\{F^{-}(u), (\phi \circ F)(x)\} = \text{cor}\{F^{-}(u), \phi(u)\}.$$  

Hence the correction factor in (3) measures the linearity of $F^{-}$ when plotted against $\phi$. If $F^{-}$ is a scaled version of $\phi$ then the correction factor is one and the loss $F^{-}(u)$ and aversion $\phi(u)$ increase in an identical manner as the percentile rank $u$ increases. In this case $\sigma_x\sigma_\psi = c\sigma^2_x$ and hence the risk margin is proportional to the variance of $x$. Generally $F^{-}(u)$ and $\phi(u)$ increase differently as $u$ increases. The correction factor is thus less than one and, in addition, varies with $F$ for a given $\phi$.

The condition $\psi = \phi \circ F$ implies aversion depends on the loss $x$ only through its percentile rank $u = F(x)$. An immediate consequence is $R(\lambda x) = \lambda R(x)$ where $\lambda$ is any nonnegative constant. Multiplying by $\lambda$ corresponds to say a change in the monetary unit and further properties of $R(x)$ when $\psi = \phi \circ F$ are discussed in §6.

For given $\phi$, the aversion to a one million and a one billion dollar loss arising from two separate loss distributions are equal if they have the same percentile rank in their respective distributions. This stresses that aversion is measured relatively, based on the relative position of a loss within its probability distribution. The aversion to two identical losses arising from two distinct loss distributions are generally not equal. The condition $E(\psi) = 1$ implies no adjustment overall: some losses are magnified and others diminished.

An argument for basing aversion on percentile ranks is as follows. Risk measurement is carried out after a risk is identified and assumes the risk is unavoidable, implying there is no aversion to the risk on average, i.e. $E(\psi) = 1$. In addition the aversion to a particular outcome only depends on its relative magnitude compared to all other outcomes suggesting the positions of outcomes in the probability distribution drives aversion rather than absolute magnitudes. An example of a measure of position is the percentile rank $F(x)$. Percentile ranks are distribution free quantities and therefore the specification of $\phi$ is independent of $F$ and, in addition, applies to all risks accepted by the entity and implying $\psi$ depends on $F$. Under this formulation, the only input required is $\phi(u)$ describing aversion to percentile ranks. To ensure a valid risk measure
requires \( \phi(u) \geq 0 \) and, as before, \( E\{\phi(u)\} = 1 \) where \( u = F(x) \). An additional condition to ensure coherence is that \( \phi \) is non-decreasing Acerbi (2002) – see §6 below.

5. Connection to distortion operators

Another way of writing the risk measure \( R(x) \) in (1) with \( \psi = \phi \circ F \) is in terms of distortion operators (Denneberg, 1990; Wang, 2001). In particular suppose \( \Phi \) is the indefinite integral of \( \phi \geq 0 \) with \( 0 = \Phi(0) \leq \Phi(u) \leq \Phi(1) = 1 \).

Then \( \phi = \Phi' \) and

\[
\int_0^\infty [1 - \Phi\{F(x)\}]dx = E[x\phi\{F(x)\}],
\]

The left hand side is the expectation of \( x \) under the distorted distribution \( \Phi \circ F \). The right hand side is \( R(x) \) and follows from the left hand side using integration by parts. Thus given \( \psi = \phi \circ F \) the equivalent distortion operator on \( F \) is given by the indefinite integral of \( \phi \). Generally it is assumed \( \Phi'' \equiv \phi' \geq 0 \) implying \( R(x) \geq E(x) \) (Wang et al., 1997). Further properties and details of distortion operators are explored in Wang (1996).

Often the survival function \( 1 - F \) is distorted, achieved with the distortion \( 1 - \{\Phi \circ (1 - I)\} \) where \( I(u) = u \). A well known special case of distorting the survival function is where survival probabilities are transformed, via the standard normal distribution, to \( z \)-scores which are then uniformly shifted and mapped back to the original scale via the inverse standard normal. In the actuarial literature this is called the “Wang transform” further discussed in §8.

Wirch and Hardy (1999) discusses the “dual power” distortion where

\[
R(x) = \int_0^\infty \{1 - F^n(x)\}dx = E\{\max(x_1, \ldots, x_n)\} = E\{xnF^{n-1}(x)\},
\]

where the \( x_i \) are \( n \) independent copies of \( x \). The right hand side follows by noting \( \Phi(u) = u^n \) implying \( \phi(u) = nu^{n-1} \). This risk measure is further discussed and extended below. Other distortion operators resulting in other risk measures include the PH transform (Wang, 1995), \( \Phi(u) = 1 - (1 - u)^{1/\gamma} \) where \( \gamma \geq 1 \) and the beta transform (Wirch and Hardy, 1999) where \( \Phi \) is the beta distribution. The PH transform yield the percentile aversion function \( (1 - u)^{(1/\gamma)-1}/\gamma \).

6. Coherence

A risk measure \( R(x) \) such as defined in (1) is coherent (Artzner et al., 1999) if it is translation invariant: \( R(c + x) = c + R(x) \), positive homogeneous: \( R(\lambda x) = \lambda R(x) \), subadditive \( R(x + y) \leq R(x) + R(y) \) for any two \( x > 0 \) and \( y > 0 \), and comonotonic: \( R(x) \geq R(y) \) if \( x \geq y \). Here \( c \) and \( \lambda \) are constants while \( x \geq y \) means \( x \) stochastically dominates \( y \).

The risk measure in (7), derived from distortion operators, is coherent (Wang et al., 1997) if \( \Phi \) is continuous, non-decreasing, satisfies \( \Phi(0) = 0 \) and \( \Phi(1) = 1 \),
and is convex. This is the case if $\Phi$ is the indefinite integral of a non-decreasing $\phi > 0$ with $E[\phi \{ F(x) \}] = 1$. Hence $\psi = \phi \circ F$ yields a coherent risk measure. An alternative proof is to note CTE is coherent, any weighted average of coherent risk measures is coherent, and any risk measure with $\psi = \phi \circ F$ is a weighted average of CTE’s as in (6) and is hence coherent.

The following additional property is sometimes imposed as a requirement for coherence: $\min(x) \leq R(x) \leq \max(x)$. The risk measure $R(x)$ in (1) satisfies this property for any $\psi > 0$ since the modified density $\tilde{f}$ has a smaller support than the original density $f$.

7. Risk margins

From (3), the risk margin or risk loading associated with $R(x)$ is

$$R(x) - E(x) = \sigma_x \sigma_\psi \text{cor}(x, \psi) \ .$$

The relation (9) indicates the close relationship to the standard deviation principle used in actuarial pricing, where the risk margin is a constant multiple of standard deviation. The difference here is the inclusion of the correction factor. The standard deviation $\sigma_\psi$ measures “conservatism”. Large values of $\sigma_\psi$ indicate values of $\psi$ are far from $E(\psi) = 1$ and hence a high degree of adjustment. In addition the term $\sigma_x \text{cor}(x, \psi)$ is a “corrected” standard deviation of the loss $x$ where the correlation $\text{cor}(x, \psi)$ is the correction factor. The standard deviation $\sigma_x$ captures the entire variability of $x$ while $\text{cor}(x, \psi) \leq 1$ represents the proportion of $\sigma_x$ relevant to the risk margin calculation. While loss variability, $\sigma_x^2$, increases the risk margin, it is moderated by $\text{cor}(x, \psi)$, the correlation between loss and aversion. If $\text{cor}(x, \psi) = 1$ then all of $\sigma_x$ is relevant to the risk margin calculation.

Note that $\psi \equiv \psi(x)$ depends functionally on $x$ and hence $\text{cor}(x, \psi)$ is a measure of the linearity of $\psi$ as a function of $x$. It is interesting to note that even though the present risk margin development has no basis in either linear modeling or normality, nevertheless the correlation $\text{cor}(x, \psi)$ is central to the calculation of the risk margin.

If $\psi = \phi \circ F$ where $\phi$ does not depend on $F$ then the distribution of $\psi$ depends only on $\phi$, noting $F(x)$ is uniformly distributed over $[0, 1]$. Hence in the case of percentile rank aversion $\sigma_\psi = \sigma_\phi$ is constant for all loss distributions $F$ and

$$R^*(x) \equiv \frac{R(x) - E(x)}{\sigma_x} = \sigma_\phi \text{cor}\{x, \phi(u)\} \ ,$$

where $u = F(x)$ and $\sigma_\phi$ does not depend on the distribution of $F$ of $x$.

The risk margin $R(x) - E(x)$ is zero if either $\sigma_x = 0$ or $\sigma_\psi = 0$, arising when $x$ and $\psi$ are respectively constant. A constant $\psi$ holds if and only if $\psi = 1$ indicating a completely cavalier attitude to the risk. Apart from the trivial situations where $x$ or $\psi$ are constant, $\text{cor}(x, \phi) > 0$ since $\phi' > 0$. 

7
An alternative representation of the risk margin is arrived at using iterated expectations:

\[
R(x) = E\{\psi E(x|\psi)\} = E(x) + \text{cov}\{\psi, E(x|\psi)\}.
\]

This is a rewriting of (3).

A special case of (10) is where \(\psi\) has only two values, as with CTE. Then \(\text{cov}\{\psi, E(x|\psi)\} = 1\) and the risk margin reduces to \(\sigma_{x}\sigma_{E(x|\psi)}\) and direct calculations show

\[
\sigma_{\psi} = \sqrt{\frac{\pi}{1 - \pi}}, \quad \sigma_{E(x|\psi)} = \sqrt{\pi(1 - \pi)} \{E(x|u > \pi) - E(x|u \leq \pi)\},
\]

where \(u \equiv F(x)\). Hence for CTE the risk margin is

\[
\pi \{E(x|u > \pi) - E(x|u \leq \pi)\}.
\]

Furthermore since \(\sigma_{x}^{2} = E(\sigma_{x|\psi}^{2}) + \sigma_{E(x|\psi)}^{2}\) it follows

\[
\sigma_{x}\text{cor}(x, \psi) = \sigma_{E(x|\psi)} = \sqrt{\sigma_{x}^{2} - \pi \sigma_{x|u \leq \pi}^{2} - (1 - \pi)\sigma_{x|u > \pi}^{2}}.
\]

Hence the adjusted measure of loss variability, \(\sigma_{x}\text{cor}(x, \psi)\), ignores the variability of losses above and below the \(\pi\)–quantile and focusses only on the variability of expected loss above and below the \(\pi\)–quantile.

Since any risk measure with \(\psi = \phi \circ F\) can be expressed as a weighted average of CTE’s it follows

\[
R(x) - E(x) = \mathcal{E}\{\pi \{E(x|u > \pi) - E(x|u \leq \pi)\}\},
\]

where \(\mathcal{E}\) is expectation with respect to the density \((1 - \pi)\phi'(\pi)\).

8. Expected maximum loss and variants

Following from (8) the choice \(\phi(u) = nu^{n-1}\) leads to

\[
R^{*}(x) = \frac{E\{\max(x_{1}, \ldots, x_{n})\} - E(x)}{\sigma_{x}} = \frac{n - 1}{\sqrt{2n - 1}} \text{cor}(x, u^{n-1}),
\]

where \(x_{1}, \ldots, x_{n}\) are \(n\) independent copies of \(x\) and the final equality follows from a straightforward calculation. This risk measure is called Expected Maximal Loss (EML) and has a convenient and practical interpretation as the expected worst outcome in \(n\) independent and identical trials. As \(n\) increases the correlation between \(x\) and \(u^{n-1}\) increases to 1 and the risk margin increases. For example if \(n = 1000\) then the risk margin is about 7 standard deviations.
The CTE and EML risk measures can be combined yielding the so-called Conditional Expected Maximal Loss (CEML) – the expected maximal loss in $n$ copies given all the copies exceed a given threshold. Define

$$\phi(u) = \frac{n(u - \pi)^{n-1}}{(1 - \pi)^n} (u > \pi)$$

where $n$ is a positive integer and $0 \leq \pi \leq 1$. Hence losses below $F^-(\pi)$ are ignored while for losses greater than $F^-(\pi)$ only the worst outcome in $n$ copies is considered:

$$R(x) = E\{\max(x_1, \ldots, x_n) | u_1 > \pi, \ldots, u_n > \pi \},$$

where $u_i = F(x_i)$. The modified distribution associated with the given $\phi$ is

$$\left\{ \frac{F(x) - \pi}{1 - \pi} \right\}^n \{F(x) > \pi\},$$

and is the distribution of the maximum of $n$ independent copies of the loss given all losses exceed $F^-(\pi)$. The CEML risk measure reduces to EML if $\pi = 0$ and to CTE if $n = 1$. Setting $\pi = 0$ and $n = 1$ yields $E(x)$.

The Expected Maximum Conditional Loss (EMCL) is the expected maximum loss conditional on the maximum exceeding the $\pi$ percentile. In other words it is conditional expectation in the maximum distribution where the maximum is taken over $n$ independent copies and conditional on the maximum exceeding the $\pi$ percentile. In this case

$$\phi(u) = \frac{(u^n > \pi)nu^{n-1}}{1 - \pi}, \quad R(x) = E(x^*|u^* > \pi),$$

where $x^* \equiv \max(x_1, \ldots, x_n)$ and $u^* \equiv \max\{F(x_1), \ldots, F(x_n)\}$, the percentile rank of $x^*$.

The CEML and EMCL are examples of the following setup. Suppose $\Phi_1$ and $\Phi_2$ are two distortion operators in the sense of (7). Then $\Phi = \Phi_1 \circ \Phi_2$ is a distortion operator with derivative $\phi = (\phi_1 \circ \phi_2) \times \phi_2$ which defines a composite percentile aversion function. The case $\phi_1(u) = nu^{n-1}$ and $\phi_2(u) = (u > \pi)/(1 - \pi)$ yields CEML discussed above while interchanging the definitions of $\phi_1$ and $\phi_2$ yields the Expected Maximum Conditional Loss (EMCL).

9. Generalized aversion functions based on mixing

An interesting risk measure is derived from the well known Esscher transform of $u \equiv F(x)$:

$$\phi(u) = \frac{e^{\lambda u}}{E(e^{\lambda u})} = \frac{\lambda e^{\lambda u}}{e^\lambda - 1}.$$

Using the expansion for $e$ yields

$$R(x) = \frac{E(x e^{\lambda u})}{E(e^{\lambda u})} = E\{\max(x_1, \ldots, x_n)\} = E(x) + \frac{\lambda}{e^\lambda - 1} \text{cov}(x, e^{\lambda u}),$$

9
where \( n \) has the zero truncated Poisson distribution with parameter \( \lambda \) and the expectation in the third expression is taken over both the \( x_i \) and \( n \). Since \( u \) is uniform the variance of \( e^{\lambda u} \) is

\[
\mathbb{E}(e^{\lambda u} - 1)^2 = \left( \frac{e^{\lambda} - 1}{\lambda} \right)^2 \left( \frac{\lambda(e^{\lambda} + 1)}{2(e^{\lambda} - 1)} - 1 \right).
\]

Hence the standardized risk margin is

\[
\left\{ \sqrt{\frac{\lambda(e^{\lambda} + 1)}{2(e^{\lambda} - 1)} - 1} \right\} \operatorname{cor}(x, e^{\lambda u}).
\]

The associated modified density and distribution are respectively

\[
f_\lambda(x) \equiv f(x) \phi\{ F(x) \} = \frac{\lambda f(x) e^{\lambda F(x)}}{e^{\phi - 1}} \quad \Rightarrow \quad F_\lambda(x) = \frac{e^{\lambda F(x)} - 1}{e^{\lambda u} - 1}.
\]

Alternatively \( F_\lambda = \Phi \circ F \) where

\[
\Phi(u) = \int_0^u \phi(v) dv = \frac{e^{\lambda u} - 1}{e^{\phi} - 1}.
\]

For example if \( x \) is exponentially distributed: \( F(x) = 1 - e^{-x} \) then

\[
\tilde{F}(x) = \frac{e^{\lambda(1-e^{-x})} - 1}{e^{\phi} - 1} = \frac{e^{-\lambda e^{-x}} - e^{-\phi}}{1 - e^{-\lambda}},
\]

and the modified distribution is the complementary log–log distribution, truncating negative values.

The last example can be generalized to other distributions on \( n \). Suppose

\[
\phi(u) = \mathbb{E}(nu^{n-1}|u) = \sum_{k=1}^{\infty} p_k ku^{k-1},
\]

where \( n \) is a positive random integer with the indicated probabilities and the expectation is with respect to \( n \). Hence \( \phi(u) \) is a probabilistic mixture of \( ku^{k-1} \) and

\[
R^*(x) = \frac{E\{\max(x_1, \ldots, x_n)\} - E(x)}{\sigma_x} = E \left\{ \frac{n - 1}{\sqrt{2n-1}} \operatorname{cor} (x, u^{n-1}|n) \right\}.
\]

In the middle expression the expectation is with respect to \( n \) and the \( x_i \) copies of \( x \), while the extreme right expectation is with respect to \( n \). Reversing this result, for a given \( \phi \) the polynomial expansion of \( \phi(u) \) can be used to give insight into the reasonableness of the given choice.

The Wang distortion (Wang, 2000) is, using the notation of §5, \( \Phi(u) = \mathcal{N}\{\mathcal{N}^-(u) - \lambda\} \) where \( \lambda > 0 \) is given and \( \mathcal{N} \) is the standard normal distribution.
function with inverse $\mathcal{N}^-$. If $x$ is normal then $u = \mathcal{N}^-\{x - \mu/\sigma\}$ where $\mu$ and $\sigma$ are the mean and standard deviations of $x$, respectively. Hence

$$\mathcal{N}^{-}(u) - \lambda = \frac{x - \mu}{\sigma} - \lambda = \frac{x - \mu - \lambda \sigma}{\sigma}$$

implying $R(x) = \mu + \lambda \sigma$, the standard deviation principle of setting risk margins. The percentile aversion function is

$$\phi = \frac{\mathcal{N}' \circ (\mathcal{N}^- - \lambda)}{\mathcal{N}' \circ \mathcal{N}^-}$$

where $\mathcal{N}'$ is the derivative of $\mathcal{N}$. Other risk measures are arrived at by replacing the standard normal distribution with other distributions such as the chi-squared or “t”.

Percentile aversion implies the value of $(\phi \circ F)(x)$ at each outcome $x$ is unchanged if $x$ is strictly monotonically increased. If the monotonic transformation is $\psi = \phi \circ F$ then

$$R\{\psi(x)\} = E\{\psi^2(x)\} = 1 + \sigma_{\psi}^2.$$ 

Hence $\sigma_{\psi}^2$ is the risk margin associated with the transformed loss $\psi(x)$.

10. Further risk measures

This section consider further risk measures based on (1) some of which occur in the recent work of Furman and Zitikis (2008):

- If $\psi(x) = x^\lambda/E(x^\lambda)$ where $\lambda > 0$ then

$$R(x) = \frac{E(x^{1+\lambda})}{E(x^{\lambda})} = E(x) + \sigma_x \frac{\sigma_{x^\lambda}}{E(x^{\lambda})} \text{cor}(x, x^\lambda).$$

Hence the standardized risk margin it the coefficient of variation of $x^\lambda$ multiplied by the correlation between $x$ and $x^\lambda$. If $\lambda = 1$ the standardized risk margin reduces to the coefficient of variation of $x$. This is the equivalent to the standard deviation principle or premium setting.

- Suppose $\psi(x)$ is proportional to $\{(1 + x)^\lambda - 1\}/\lambda$ with $\lambda > 0$. Then $\psi(x)$ converges to $\ln(1 + x)$ as $\lambda \to 0$ and $\psi(0) = 0$ for all $\lambda \geq 0$. In this case

$$R(x) = \frac{E\{x \ln(1 + x)\}}{E\{\ln(1 + x)\}} = E(x) + \sigma_x \frac{\sigma_{\ln(1+x)}}{E\{\ln(1 + x)\}} \text{cor}\{x, \ln(1 + x)\}.$$ 

- Tail value at risk (TVaR) is attained if $\psi(x)$ proportional to $x\{F(x) > \pi\}$. This illustrates the procedure of deriving a new aversion function from the product of two aversion functions: in this case $x$ and $\{F(x) > \pi\}$, ignoring normalising constants. Hence in this case aversion is proportional to $x$. 

11
provided the loss is in the upper tail defined by \( \pi \). Using arguments similar to the above

\[
R(x) = E\{x|x > F^-(\pi)\} + \frac{\sigma^2_{x|x > F^-(\pi)}}{E\{x|x > F^-(\pi)\}}.
\]

This is analogous to the \( \psi(x) = x/E(x) \) case with the additional conditioning on \( x > F^-(\pi) \).

- **Esscher premiums** are attained with \( \psi(x) = e^{\lambda x}/E(e^{\lambda x}) \). In this case

\[
R(x) = \frac{E(xe^{\lambda x})}{E(e^{\lambda x})} = E(x) + \sigma_x \frac{e^{\lambda x}}{E(e^{\lambda x})} \text{cor}(x, e^{\lambda x}).
\]

Expanding \( e^{\lambda x} \) shows the Esscher premium can be viewed as a mixture of risks with aversion function \( \psi(x) = x^n E(x^n) \) for non-negative integer values of \( n \), with weights proportional to \( \lambda^n/n! \).

- A variant of the Esscher premium is the Kamps premium where \( \psi(x) = 1 - e^{-\lambda x} \) yielding the standardized risk margin

\[
R^*(x) \equiv \frac{R(x) - E(x)}{\sigma_x} = \frac{e^{-\lambda x}}{E(e^{-\lambda x}) - 1} \text{cor}(x, e^{-\lambda x}).
\]

Note \( 1 - e^{-\lambda x} \) is the distribution of the Exponential distribution, and the Kamps premium is generalized by setting \( \psi(x) \) as a loss distribution with mean normalised to one. In this case, \( \psi(x) = G(x) \), where \( G \) is a distribution with mean one.

- A final example is \( \psi = \phi \circ F \) where \( \phi(u) = G^-(u)/\mu_G \). Here \( G^- \) is an inverse distribution with mean \( \mu_G \). It is then useful to view aversion as the pseudo loss \( G^- (u) \) varying with percentile rank \( u \), but normalised so that the expectation is one. Then

\[
R^*(x) \equiv \frac{R(x) - E(x)}{\sigma_x} = \frac{\sigma_G}{\mu_G} \text{cor}\{G^-(u), F^-(u)\},
\]

where \( \sigma_G \) is the standard deviation of \( G^- (u) \). Hence the standardized risk margin is a product of the coefficient of variation of a pseudo loss \( G^- (u) \), and the extent to which \( x \) and the pseudo loss are identical after allowing for scale differences. Here \( x \) and \( G^- (u) \) are comonotonic but not identical due to differences between \( F \) and \( G \). Note \( G \) is fixed but \( F \) varies depending on the loss in question. When \( F = G \), ignoring scale differences, then \( \psi(x) = x/E(x) \), discussed above. The example allows aversion to approach infinity as the loss does, that is, when the percentile rank approaches one. Wang (2000) discusses this property in the context of premium calculation.
11. Application to gamma and Burr loss distributions

This section illustrates the application of aversion to distributions displayed in the two top panels of Figure 1: a gamma and Burr density (Klugman et al., 2004) proportional to $ xe^{-x} $ and $ x/(1 + x^2)^{3} $, respectively and where $ x > 0 $. The riskiness of the two densities as indicated by the risk measure $ R(x) $ are not directly comparable. However they are made comparable by considering the standardized risk margins $ R^{*}(x) $, as defined in (4), which are independent of location and scale. Hence, in effect, the two densities are made comparable by standardizing on the same mean and variance: 0 and 1. This emphasizes that for risk comparison purposes, it is the shapes of the two distributions which are critical.

The bottom two panels of Figure 1 display aversion adjusted densities

$$ \tilde{f}(x) \equiv \psi(x)f(x) = \phi\{F(x)\} f(x) $$

where $ f, F $ and $ \phi $ are the density, distribution and aversion function, respectively. The risk measure $ R(x) $ is the expected loss under the aversion adjusted density $ \tilde{f}(x) $, which is then standardized to yield $ R^{*}(x) $. VaR, CTE, EML and CEML, as illustrated in Figure 1, correspond to different choices of $ \phi $ (ignoring normalizing constants):

$$ (u = 0.95), \quad (u > 0.9), \quad u^{0}, \quad (u - 0.8)^{4}(u > 0.8) $$

Thus VaR is at 0.95, CTE is at 0.90 while EML is the expected worst outcome in $ n = 10 $ copies. Finally CEML is at $ \pi = 0.80 $ and $ n = 5 $ and hence corresponds to the maximum expected loss in 5 copies given all exceed the 0.80 percentile.

Table 1 displays the means and variances of the two distributions as well as the standardized risk margins $ R^{*}(x) $, under the VaR, CTE, EML and CEML aversion adjustments.

<table>
<thead>
<tr>
<th>distribution</th>
<th>$ E(x) $</th>
<th>$ \sigma_{x} $</th>
<th>VaR</th>
<th>CTE</th>
<th>EML</th>
<th>CEML</th>
</tr>
</thead>
<tbody>
<tr>
<td>gamma</td>
<td>2.00</td>
<td>1.42</td>
<td>1.94</td>
<td>2.18</td>
<td>1.85</td>
<td>2.68</td>
</tr>
<tr>
<td>Burr</td>
<td>0.79</td>
<td>0.61</td>
<td>1.74</td>
<td>2.18</td>
<td>1.84</td>
<td>2.91</td>
</tr>
</tbody>
</table>

The following features of the standardized risk margins $ R^{*}(x) $ are worth noting. First, the gamma has a significantly higher standardized risk margin than the Burr using VaR at $ \pi = 0.95 $. This suggests the gamma is the more “risky” distribution given the same mean and variance. However, there is little difference in “riskiness” between the gamma and Burr based on the CTE and EML. Finally the standardized risk margins computed from CEML at the indicated parameters suggest the Burr distribution is more “risky” than the gamma. These somewhat conflicting messages suggest that in practice, the choice of the aversion function, and hence the risk measure, is important when determining
risk margins, and may dominate the choice of the risk distribution. Finally the knowledge of the sensitivity of the risk margin to the risk measure gives a more complete, and hence perhaps more comforting, view of the risk profile.

12. Revealed aversion

This section discusses an approach to determining an appropriate aversion function \( \psi \) in practical settings. Suppose, \( x_0 \) is an outcome of \( x \), used as a reference loss. For other possible outcomes of \( x \), a risk manager is asked the relative aversion to the loss, ie the value of \( x \psi(x)/x_0 \psi(x_0) \). For example in millions of dollars, \( x_0 = 1 \) with losses of 2, 4 and 8 evaluated as 3, 10 and 50 times worse than 1, respectively. Hence it is possible to evaluate \( \psi(x)/\psi(x_0) \) for each outcome \( x \), and, given the probability distribution of losses,

\[
\begin{align*}
\mathbb{E}\left\{ \frac{\psi(x)}{\psi(x_0)} \right\} &= \mathbb{E}\{\psi(x)\} = \frac{1}{\psi(x_0)}, \\
\mathbb{E}\left\{ \frac{x \psi(x)}{x_0 \psi(x_0)} \right\} &= \mathbb{E}\{x \psi(x)\} = \frac{\mathbb{R}(x)}{x_0 \psi(x_0)}.
\end{align*}
\]
The first equation is used to determine $\psi(x_0)$. Given $x_0$ and $\psi(x_0)$, the second equation is used to determine $R(x)$. Further $\psi(x)$ is determined from the relativities $x\psi(x)/\{x_0\psi(x_0)\}$.

Table 2 illustrates the steps in the calculations. The first two columns indicate possible losses (in millions of dollars say) and the associated probabilities, respectively. The third displays the cumulative distribution $F(x)$ and the fourth column serves to compute the mean $E(x) = 1.64$. The fifth column indicates the assumed relative severities relative to the reference case $x_0 = 1$. Thus a 2 million dollar loss is judged 3 times more severe than a 1 million dollar loss, while 4 million is 10 times as severe, and 8 million is 50 times as severe. These severities together with the values of $x$ and $x_0$ are then used to infer the ratios $\psi(x)/\psi(x_0)$. Imposing the condition $E\{\psi(x)\} = 1$ determines $\psi(x_0)$ and hence the seventh column. The next two columns display the calculations for $\tilde{f}(x) = \psi(x)f(x)$ and the $R(x) = E\{x\psi(x)\} = 2.17$.

For this example $R(x) = 2.17 \approx F^{-}(0.9)$ and the standardized risk margin is $(2.17-1.64)/1.40=0.38$. Further computations show $\sigma_\psi = 0.50$ and hence $\text{cor}(x, \psi) = 0.76$.

Table 2: Losses, probabilities, aversions and associated calculations

<table>
<thead>
<tr>
<th>x</th>
<th>f(x)</th>
<th>F(x)</th>
<th>xf(x)</th>
<th>xψf(x)</th>
<th>ψ(x)</th>
<th>f(x)</th>
<th>xf(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>1.00</td>
<td>0.75</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>2.00</td>
<td>0.30</td>
<td>0.90</td>
<td>0.60</td>
<td>3.00</td>
<td>1.12</td>
<td>0.34</td>
<td>0.67</td>
</tr>
<tr>
<td>4.00</td>
<td>0.09</td>
<td>0.99</td>
<td>0.36</td>
<td>10.00</td>
<td>2.50</td>
<td>0.17</td>
<td>0.67</td>
</tr>
<tr>
<td>8.00</td>
<td>0.01</td>
<td>1.00</td>
<td>0.08</td>
<td>50.00</td>
<td>6.25</td>
<td>0.05</td>
<td>0.37</td>
</tr>
<tr>
<td>Total</td>
<td>1.00</td>
<td>1.64</td>
<td>1.00</td>
<td>2.17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The plot of $\psi(x)$ versus $u = F(x)$ is $\phi(u)$, the implied percentile rank aversion function. From Table 2, $\phi(0.6) = 0.75$, $\phi(0.9) = 1.12$, $\phi(0.99) = 1.87$ and $\phi(1) = 4.67$. If $\phi(u) \approx nu^{n-1}$, that is $\ln \phi(u) \approx \ln n + (n-1) \ln u$, then $E\{x\psi(x)\} \approx E\{\max(x_1, \ldots, x_n)\}$ where the $x_i$ are $n$ independent copies of $x$. With the Table 2 illustrative data, a least squares regression of $\ln \psi(x)$ on $\ln F(x)$, or equivalently $\ln \phi(u)$ on $u$, yields

$$\ln \phi(u) \approx 0.90 + 2.54 \ln u \approx \ln 3 + (3 - 1) \ln u,$$

suggesting $R(x) = 2.19$, attained with the revealed aversion function, is approximately equivalent to setting reserves equal to the expected worst outcome in 3 independent trials.

To throw further light on $\psi(x)$ of Table 2, suppose required is the reserving equivalent to the expected worst outcome in 10 independent trials. Then $\psi(x)$ is proportional to $10\{F(x)\}^5$. Using the above $F(x)$ values this implies $\psi(x)$ equals 0.004, 0.168, 0.395 and 0.433 at the Table 2 given $x$ values. Hence the relative $\psi(x)$ values in this case are 1, 38.4, 90.6 and 99.2.
13. Conclusion

This article has considered risk measurement or loss reserving based on aversion functions. Aversion functions adjust different possible outcomes according to perceived severity. The loss reserve is then calculated as the expected value of the actual loss multiplied by the aversion adjusted loss. A specialization is where aversion depends only on the percentile rank of the outcome. Many practical risk measures are of this form, including Value at Risk, Conditional Tail Expectation and Expected Maximum Loss in independent copies. The risk measures can be rewritten as either expectation under a distorted density or as the expected value loaded with a quantity proportional to the correlation between the loss and the aversion adjusted loss. Conditions on the aversion function to ensure coherence are discussed. The different representations of the risk measures permit the construction of generalized risk measures based on multiplying or mixing the relatively well known measures. A method is outlined to elicit appropriate aversion functions in practical settings.

References


