Coherent Risk Measures

C VIJAY ADITHYA 06/12/2024

On the notion of risk and its measurement:

- Risk as the possibility of an undesirable event(- early hints towards the usage of probability theory)
- Risk measure defined as the amount of transactable material to be held with you as agreed by all regulators in a given location
- Risk measures in history built initially on the base of statistical measures of dispersion, some famous methods of measurement as follows:
  1. Markowitz M-V Framework (1952)
  2. Sharpe’s CAPM (1964)
  3. Black-Scholes option pricing model “Greek” (1973)
  4. stress testing methods (late 80s)
- Above are some pre-1990s models which were pertinent

Post 1990s measures and pertinence:

1. Value at Risk

Let X be a profit and loss distribution (loss negative and profit positive). The VaR at level $\alpha \in (0,1)$ is the smallest number y such that the probability that $Y := -X$ does not exceed y is at least $1-\alpha$. Mathematically, $\mathrm{VaR}_\alpha(X)$ is the $(1-\alpha)$-quantile of Y, i.e.

$$\mathrm{VaR}_\alpha(X) = -\inf\{x \in \mathbb{R} : F_X(x) > \alpha\} = F_Y^{-1}(1-\alpha).$$ [14][15]

This formula can only be used when ∃ parametric distribution for F (applies to both 1 and 2)

2. Tail Value At Risk (or Tail conditional expectation)

Given a random variable X which is the payoff of a portfolio at some future time and given a parameter $0 < \alpha < 1$ then the tail value at risk is defined by[5][6][7][8]

$$\mathrm{TVaR}_\alpha(X) = \mathrm{E}[-X | X \leq -\mathrm{VaR}_\alpha(X)] = \mathrm{E}[-X | X \leq x^\alpha],$$

Post 1990s measures and pertinence:

4. Worst conditional expectation (screenshot from artzner et al 1999)

DEFINITION 5.2. *Worst conditional expectation.* Given a base probability measure $\mathbb{P}$ on Ω, a total return r on a reference instrument, and a level α, the worst conditional expectation is the measure of risk defined by

$$WCE_{\alpha}(X) = -\inf\{\mathbf{E}_{\mathbb{P}}[X/r \mid A] \mid \mathbb{P}[A] > \alpha\}.$$

This formula can only be used when ∃ parametric distribution for F (applies to both 4 and 5)

2. Conditional Value at Risk (or expected shortfall)

$$\mathrm{ES}_{\alpha}(X) = E[-X \mid X \leq -\mathrm{VaR}_{\alpha}(X)]$$

Typical values of α in this case are 5% and 1%

Risk as a natural and intuitive notion

ON THE INTUITIVE EXPLANATION OF RISK AND ITS AXIOMS AS PROPOSED BY ARTZNER ET AL 1999

Set up of risk measures: (Artzner et al 1999)

- G as the set of all risks, i.e. The set of all real valued functions on the real numbers
- Acceptance set A (for a given country) as the set of all net worths accepted by all the regulators in a given country

We shall call Ω the set of states of nature, and assume it is finite. Considering Ω as the set of outcomes of an experiment, we compute the final net worth of a position for each element of Ω. It is a random variable denoted by X.

- If ρ(X) >0, ρ(X) is the minimum extra cash the agent has to add to the risky position X to get into an acceptance set
- If ρ(X)<0, we do the opposite, and the case of ρ(X)=0 is trivial (the last two points are motivational definitions)

Set up of risk measures contd.

DEFINITION 2.2. *Risk measure associated with an acceptance set.* Given the total rate of return r on a reference instrument, the risk measure associated with the acceptance set $\mathcal{A}$ is the mapping from $\mathcal{G}$ to $\mathbb{R}$ denoted by $\rho_{\mathcal{A},r}$ and defined by

$$\rho_{\mathcal{A},r}(X) = \inf\{m \mid m \cdot r + X \in \mathcal{A}\}.$$

DEFINITION 2.3. *Acceptance set associated with a risk measure.* The acceptance set associated with a risk measure ρ is the set denoted by $\mathcal{A}_\rho$ and defined by

$$\mathcal{A}_\rho = \{X \in \mathcal{G} \mid \rho(X) \leq 0\}.$$

Does a good risk measure need to follow axioms?

ANSWER TO THE QUESTION: WHAT ARE COHERENT MEASURES OF RISK?

Axiom T: Translation invariance

AXIOM T. Translation invariance. *For all $X \in \mathcal{G}$ and all real numbers α, we have* $\rho(X + \alpha \cdot r) = \rho(X) - \alpha$.

REMARK 2.6. Axiom T ensures that, for each X, $\rho(X + \rho(X) \cdot r) = 0$. This equality has a natural interpretation in terms of the acceptance set associated with ρ (see Definition 2.3 above).

Axiom S: Subadditivity

AXIOM S. Subadditivity. *For all* X_1 *and* $X_2 \in \mathcal{G}$, $\rho(X_1 + X_2) \le \rho(X_1) + \rho(X_2)$.

We contend that this property, which could be stated in the brisk form "a merger does not create extra risk," is a natural requirement:

Try to remember this axiom

Axiom PH: Positive Homogeneity

AXIOM PH. Positive homogeneity. *For all* $\lambda \geq 0$ *and all* $X \in \mathcal{G}$, $\rho(\lambda X) = \lambda \rho(X)$.

REMARK 2.8. If position size directly influences risk (e.g., if positions are large enough that the time required to liquidate them depends on their sizes) then we should consider the consequences of lack of liquidity when computing the future net worth of a *position*. With this in mind, Axioms S and PH, about mappings from *random variables* into the reals, remain reasonable.

Axioms M and R: Monotonicity and relevance

REMARK 2.10. Axioms T and PH imply that, for each α, $\rho(\alpha \cdot (-r)) = \alpha$.

AXIOM M. Monotonicity. *For all X and $Y \in \mathcal{G}$ with $X \leq Y$, we have $\rho(Y) \leq \rho(X)$.*

AXIOM R. Relevance. *For all $X \in \mathcal{G}$ with $X \leq 0$ and $X \neq 0$, we have $\rho(X) > 0$.*

Definition of coherent risk measure:

DEFINITION 2.4. *Coherence.* A risk measure satisfying the four axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity is called coherent.

Why should I care about coherence?

ON THE IMPLICATIONS, CLAIMS, AND PROOFS THAT FOLLOW

Implications on Value at Risk (VaR)

Ref: Acerbi, Carlo. 2007. “Coherent Measures of Risk in Everyday Market Practice†.” Quantitative Finance 7 (4): 359–64. doi:10.1080/14697680701461590.

- “…by imposing, via an axiomatic framework, specific mathematical conditions which enforce some basic principles that a sensible risk measure should always satisfy…as the first serious attempt to give a precise definition of financial risk itself, via a deductive approach.” (Acerbi, Carlo, 2007)

In conclusion, the basic reasons to reject the value at risk measure of risks are the following:

(a) value at risk does not behave nicely with respect to the addition of risks, even independent ones, thereby creating severe aggregation problems.

(b) the use of value at risk does not encourage and, indeed, sometimes prohibits diversification because value at risk does not take into account the *economic consequences* of the events, the probabilities of which it controls.

Implications on Value at Risk (VaR) contd.

- Value at risk, not a coherent measure of risk, due to its lack of subadditivity (Artzner)
  Artzner's proof by counterexample is as follows:

- Probability density function (PDF) for X_1 and X_2:

$$f(x) = \begin{cases} 0.90 & \text{for } x \in [0, 1], \\ 0.05 & \text{for } x \in [-2, 0], \\ 0 & \text{otherwise.} \end{cases}$$

This is a simplification of the example into probability notation, however conveys the message.
The 10% VaR for both RVs will be 0, however the VaR of the sum of the RVs will be >0 and hence it is not sub additive.

- Finite mean, finite variance, however, not a distribution which might look like normal or normal type-coincidence?

Implications on Value at Risk (extended proof)

Definition 4. Let $T : \mathbb{R}^n \to \mathbb{R}^n, \mathbf{x} \mapsto A\mathbf{x} + \mu$, $A \in \mathbb{R}^{n\times n}$, $\mu \in \mathbb{R}^n$ be an affine map. $\mathbf{X}$ has an elliptical distribution if $\mathbf{X} = T(\mathbf{Y})$ and $\mathbf{Y} \sim S_n(\phi)$.

Since the characteristic function can be written as

$$\begin{aligned}\psi(\mathbf{t}) &= \mathbb{E}[\exp(i\mathbf{t}^t\mathbf{X})] = \mathbb{E}[\exp(i\mathbf{t}^t(A\mathbf{Y} + \mu))] \\ &= \exp(i\mathbf{t}^t\mu)\exp(i(A^t\mathbf{t})^t\mathbf{Y}) = \exp(i\mathbf{t}^t\mu)\phi(\mathbf{t}^t\Sigma\mathbf{t}),\end{aligned}$$

where $\Sigma := AA^t$, we denote the elliptical distributions

$$\mathbf{X} \sim \mathrm{E}_n(\mu, \Sigma, \phi).$$

Implications on Value at Risk (extended proof)

- Any linear combination of an elliptically distributed random vector is also elliptical with the same characteristic generator ϕ. If $\mathbf{X} \sim \mathrm{E}_n(\mu, \Sigma, \phi)$ and $B \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, then

$$B\mathbf{X} + b \sim \mathrm{E}_m(B\mu + b, B\Sigma B^t, \phi).$$

Implications on Value at Risk (extended proof)

Theorem 1. *Suppose* $\mathbf{X} \sim E_n(\mu, \Sigma, \phi)$ *with* $\sigma^2[X_i] < \infty$ *for all* i. *Let*

$$\mathcal{P} = \{Z = \sum_{i=1}^{n} \lambda_i X_i \mid \lambda_i \in \mathbb{R}\}$$

be the set of all linear portfolios. Then the following are true.

1. *(Subadditivity of VaR.) For any two portfolios* $Z_1, Z_2 \in \mathcal{P}$ *and* $0.5 \leq \alpha < 1$,

$$VaR_\alpha(Z_1 + Z_2) \leq VaR_\alpha(Z_1) + VaR_\alpha(Z_2).$$

Implications on Value at Risk (extended proof)

Proof. The main observation is that $(Z_1, Z_2)^t$ has an elliptical distribution so Z_1, Z_2 and $Z_1 + Z_2$ all have distributions of the same type.

1. Let q_α be the α-quantile of the standardised distribution of this type. Then

$$\begin{aligned}
\mathrm{VaR}_\alpha(Z_1) &= \mathbb{E}[Z_1] + \sigma[Z_1]q_\alpha, \\
\mathrm{VaR}_\alpha(Z_2) &= \mathbb{E}[Z_2] + \sigma[Z_2]q_\alpha, \\
\mathrm{VaR}_\alpha(Z_1 + Z_2) &= \mathbb{E}[Z_1 + Z_2] + \sigma[Z_1 + Z_2]q_\alpha.
\end{aligned}$$

   Since $\sigma[Z_1 + Z_2] \leq \sigma[Z_1] + \sigma[Z_2]$ and $q_\alpha \geq 0$ the result follows.

Proof taken from: Embrechts, Paul & McNeil, Alexander. (2000). Correlation And Dependence In Risk Management: Properties And Pitfalls. Rev. Econ. Stat.. 86.

Implications on Conditional Value at Risk (CVaR):

S.P. Uryasev (ed.), Probabilistic Constrained Optimization, 272-281

The *conditional value-at-risk* $CVaR_\alpha$ is defined as the solution of an optimization problem

$$\mathrm{CVaR}_\alpha(Y) := \inf\{a + \frac{1}{1-\alpha}\mathbb{E}[Y-a]^+ : a \in \mathbb{R}\}. \tag{2}$$

Here $[z]^+ = \max(z, 0)$. Uryasev and Rockafellar (1999) have shown that CVaR equals the conditional expectation of Y, given that $Y \geq \mathrm{VaR}_\alpha$, i.e.

$$\mathrm{CVaR}_\alpha(Y) = \mathbb{E}(Y | Y \geq \mathrm{VaR}_\alpha(Y)). \tag{3}$$

Implications on Conditional Value at Risk (CVaR):

Artzner, Delbaen, Eber and Heath call a risk measure *coherent*, if it is translation-invariant, convex, positively homogeneous and monotonic w.r.t. $\prec_{SD(1)}$. One sees that CVaR_α is coherent in this sense.

Now we prove (iv). Let a_i be such that $\mathrm{CVaR}_\alpha(Y_i) = a_i + \frac{1}{1-\alpha}\mathbb{E}[Y_i - a_i]^+$. Since $y \mapsto [y-a]^+$ is convex, we have

$$\begin{aligned}
& \mathrm{CVaR}_\alpha(\lambda Y_1 + (1-\lambda)Y_2) \\
\leq \quad & \lambda a_1 + (1-\lambda)a_2 + \frac{1}{1-\alpha}\mathbb{E}[\lambda Y_1 + (1-\lambda)Y_2 - \lambda a_1 + (1-\lambda)a_2)]^+ \\
\leq \quad & \lambda a_1 + (1-\lambda)a_2 + \frac{\lambda}{1-\alpha}\mathbb{E}[Y_1 - a_1]^+ + \frac{1-\lambda}{1-\alpha}\mathbb{E}[Y_2 - a_2]^+ \\
\leq \quad & \lambda\ \mathrm{CVaR}_\alpha(Y_1) + (1-\lambda)\ \mathrm{CVaR}_\alpha(Y_2).
\end{aligned}$$

Representation theorems and applications

TOWARDS THE EASE OF CHECKING FOR COHERENCE

Representation of coherent risk measures by scenarios: (with proof)

PROPOSITION 4.1. *Given the total return r on a reference investment, a risk measure ρ is coherent if and only if there exists a family $\mathcal{P}$ of probability measures on the set of states of nature, such that*

$$\rho(X) = \sup\{\mathbf{E}_{\mathbb{P}}[-X/r] \mid \mathbb{P} \in \mathcal{P}\}.$$

(2) The sets Ω and $\mathcal{M}$ of Huber are our set Ω and the set of probabilities on Ω. Given a risk measure ρ we associate with it the functional E^* by $E^*(X) = \rho(-r \cdot X)$. Axiom M for ρ is equivalent to Property (2.7) of Huber for E^*, Axioms PH and T together are equivalent to Property (2.8) for E^*, and Axiom S is Property (2.9).

Ref: Huber(1991)

Representation of coherent risk measures by scenarios: (with proof)

(3) The "if" part of our Proposition 4.1 is obvious. The "only if" part results from the "representability" of E^*, since Proposition 2.1 of Huber states that

$$\rho(X) = E^*(-X/r) = \sup\{\mathbf{E}_{\mathbb{P}}[-X/r] \mid \mathbb{P} \in \mathcal{P}_\rho\},$$

where $\mathcal{P}_\rho$ is defined as the set

$$\begin{aligned}\{\mathbb{P} \in \mathcal{M} \mid \text{ for all } X \in \mathcal{G}\colon\ \mathbf{E}_{\mathbb{P}}[X] \leq E^*(X) = \rho(-r \cdot X)\} \\ = \{\mathbb{P} \in \mathcal{M} \mid \text{ for all } Y \in \mathcal{G}\colon\ \mathbf{E}_{\mathbb{P}}[-Y/r] \leq \rho(Y)\}.\end{aligned}$$

Ref: Huber(1991)

Application of coherent risk measures:

"...Casualty actuaries have long been computing pure premium for policies with deductible, using the conditional average of claim size, given that the claim exceeds the deductible…. We prove below that one of the suggested methods gets us close to *coherent* risk measures."

DEFINITION 5.1. *Tail conditional expectation (or "TailVaR").* Given a base probability measure $\mathbb{P}$ on Ω, a total return r on a reference instrument, and a level α, the tail conditional expectation is the measure of risk defined by

$$TCE_\alpha(X) = -\mathbf{E}_\mathbb{P}\left[X/r \mid X/r \leq -VaR_\alpha(X)\right].$$

DEFINITION 5.2. *Worst conditional expectation.* Given a base probability measure $\mathbb{P}$ on Ω, a total return r on a reference instrument, and a level α, the worst conditional expectation is the coherent measure of risk defined by

$$WCE_\alpha(X) = -\inf\{\mathbf{E}_\mathbb{P}\left[X/r \mid A\right] \mid \mathbb{P}[A] > \alpha\}.$$

More representations without explicit proofs:

PROPOSITION 5.1. *We have the inequality* $TCE_\alpha \leq WCE_\alpha$.

PROPOSITION 5.2. *For each risk* X *one has the equality*

$$VaR_\alpha(X) = \inf\{\rho(X) \mid \rho \text{ coherent and } \rho \geq VaR_\alpha\}$$

We do not go any more deeper into coherent risk measures as the context for the final part has been built

Back to elliptical distributions?

THE ROLE OF ELLIPSES IN RISK MEASUREMENT

Statistical implications of such distributions:

Correlation as a measure of dependence fails in non elliptical distributions (Joe 1997)

Markowitz' mean variance model only suitable for elliptically distributed returns if that is the case

Severe underestimation of risk occurs otherwise –notice the definition of correlation

VaR not coherent due to lack of convexity (causes a problem in optimization) when placed in elliptical distributions, and the lack of subadditivity when non- elliptical

Need for a measure of dependence under non-elliptic conditions:

Introduction to scalar risk measures and copulas

What about fat tailed distributions?

Majority of the models discussed will not work on fat tailed distributions due to the limitations in the definition of the law of large numbers- mainly the models dependent on finite moments

Copulas used extensively as dependence measures for fat tailed distributions

Do there exist alternative risk measures (that are coherent?)