How Superadditive Can a Risk Measure Be?

Ruodu Wang
(wang@uwaterloo.ca)

Department of Statistics and Actuarial Science
University of Waterloo, Canada

Seminar on Probability and Statistics
University of Bern    Bern, Switzerland    May 9th, 2014

Joint work with Valeria Bignozzi and Andreas Tsanakas
Outline

1. VaR and ES
2. The Holy Triangle of Risk Measures
3. How Superadditive Can a Risk Measure Be?
4. Discussion
5. References
From Basel Committee on Banking Supervision:

R1: Consultative Document, May 2012, 
Fundamental review of the trading book

R2: Consultative Document, October 2013, 
R1, Page 41, Question 8:

"What are the likely constraints with moving from VaR to ES, including any challenges in delivering robust backtesting, and how might these be best overcome?"

- Cont, Deguest and Scandolo (2010): ES is not robust, while VaR is.
- Gneiting (2011): ES is not elicitable, while VaR is.
**R1**, Page 41, Question 8:

"What are the likely constraints with moving from VaR to ES, including any challenges in delivering robust backtesting, and how might these be best overcome?"

- Cont, Deguest and Scandolo (2010): ES is not robust, while VaR is.
- Gneiting (2011): ES is not elicitable, while VaR is.

In this talk, $X > 0$ is interpreted as a loss.

**Definition**

$\text{VaR}_p(X)$, for $p \in (0, 1)$,

$$\text{VaR}_p(X) = F_X^{-1}(p) = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\}.$$  

**Definition**

$\text{ES}_p(X)$, for $p \in (0, 1)$,

$$\text{ES}_p(X) = \frac{1}{1 - p} \int_p^1 \text{VaR}_\delta(X) d\delta = \mathbb{E} \left[ X | X > \text{VaR}_p(X) \right].$$
### VaR versus ES: Summary

**Value-at-Risk**

1. **Always** exists
2. **Only** frequency
3. Non-coherent risk measure *(diversification problem)*
4. Backtesting **straightforward**
5. Estimation: far in the tail
6. Model uncertainty: sensitive to dependence
7. (Almost) robust with respect to **weak topology**

**Expected Shortfall**

1. **Needs** first moment
2. Frequency and **severity**
3. Coherent risk measure *(diversification benefit)*
4. Backtesting an issue *(non-elicitability)*
5. Estimation: data limitation
6. Model uncertainty: sensitive to tail modeling
7. Robust with respect to **Wasserstein distance**
A risk measure $\rho : \mathcal{X} \rightarrow [-\infty, \infty]$. $\mathcal{X} \supset L^\infty$ is a set which is closed under addition and scaler multiplication.

There are many desired properties of a good risk measure. Some properties are without debate:

- cash-invariance: $\rho(X + c) = \rho(X) + c, c \in \mathbb{R}$;
- monotonicity: $\rho(X) \leq \rho(Y)$ if $X \leq Y$;
- zero-normalization: $\rho(0) = 0$;
- law-invariance: $\rho(X) = \rho(Y)$ if $X =_d Y$.

(A standard risk measure; those properties are not restrictive)

Another one is listed here as debatable:

- positive homogeneity: $\rho(\lambda X) = \lambda \rho(X), \lambda \geq 0$. 

In my opinion, in addition to being standard, the three key elements of being a good risk measure are

(C) Coherence (subadditivity): \( \rho(X + Y) \leq \rho(X) + \rho(Y) \).

[aggregate regulation/capturing the tail/capital allocation/convex optimization]
In my opinion, in addition to being standard, the three key elements of being a good risk measure are

(C) Coherence (subadditivity): $\rho(X + Y) \leq \rho(X) + \rho(Y)$. [aggregate regulation/capturing the tail/capital allocation/convex optimization]

(A) Comonotonic additivity: $\rho(X + Y) = \rho(X) + \rho(Y)$ if $X$ and $Y$ are comonotonic. [economical interpretation/distortion representation/non-diversification identity]
The Holy Triangle of Risk Measures

In my opinion, in addition to being standard, the three key elements of being a good risk measure are

(C) Coherence (subadditivity): $\rho(X + Y) \leq \rho(X) + \rho(Y)$. [aggregate regulation/capturing the tail/capital allocation/convex optimization]

(A) Comonotonic additivity: $\rho(X + Y) = \rho(X) + \rho(Y)$ if $X$ and $Y$ are comonotonic. [economical interpretation/distortion representation/non-diversification identity]

(E) Elicitability [statistical advantage/backtesting straightforward].
The War of the Two Kingdoms

- Financial mathematicians
  - appreciate coherence (subadditivity);
  - favor ES in general.

- Financial statisticians
  - appreciate backtesting and statistical advantages;
  - favor VaR in general.

A natural question is to find a standard risk measure which is both coherent (subadditive) and elicitable.
Expectiles

For $0 < p < 1$ and $X \in L^2$, the $p$-expectile is

$$e_p(X) = \arg\min_{x \in \mathbb{R}} \mathbb{E}[p(X - x)^2_+ + (1 - p)(x - X)^2_+]$$.

- $e_p(X)$ is the unique solution $x$ of the equation for $X \in L^1$:
  $$p\mathbb{E}[(X - x)_+] = (1 - p)\mathbb{E}[(x - X)_+]$$.
- $e_{1/2}(X) = \mathbb{E}[X]$.
- If we allow $p = 1$: $e_1(X) = \text{ess-sup}(X)$. 
Expectiles

The risk measure $e_p$ has the following properties:

1. positive homogeneous and standard;
2. subadditive for $1/2 \leq p < 1$, superadditive for $0 < p \leq 1/2$;
3. elicitable;
4. coherent for $1/2 \leq p < 1$;
5. not comonotonic additive in general.

In summary:

- VaR has (A) and (E): often criticized for not being subadditive: *diversification/aggregation problems and inability to capture the tail!*
In summary:

- VaR has (A) and (E): often criticized for not being subadditive: *diversification/aggregation problems and inability to capture the tail!*

- ES has (C) and (A): criticized for *estimation, backtesting and robustness problems!*
The War of the Three Kingdoms

In summary:

- VaR has (A) and (E): often criticized for not being subadditive: *diversification/aggregation problems and inability to capture the tail!*

- ES has (C) and (A): criticized for *estimation, backtesting and robustness problems!*

- Expectile has (C) and (E): criticized for *lack of economical meaning, difficulty to conceptualize, distributional computation and over-diversification benefits!***
The War of the Three Kingdoms

The following hold:

- if $\rho$ is coherent, comonotonic additive and elicitable, then $\rho$ is the mean (Ziegel, 2014);

- if $\rho$ is coherent, and elicitable with a convex scoring function, then $\rho$ is an expectile (Bellini and Bignozzi, 2014);

- if $\rho$ is comonotonic additive, and elicitable, then $\rho$ is a VaR or the mean (Kou and Peng, 2014, W. and Ziegel, 2014).
In summary:

The only standard risk measure that has (C), (A) and (E) is the **mean**, which is not a tail risk measure, and does not have a risk loading.
The War of the Three Kingdoms

In summary:

The only standard risk measure that has (C), (A) and (E) is the mean, which is not a tail risk measure, and does not have a risk loading.

- Remark: the very old-school risk measure/pricing principle $\rho(X) = (1 + \theta)\mathbb{E}[X], \theta > 0$ has (C-subadditivity), (A) and (E), although it is not standard.
Subadditivity

Subadditivity has to do with

- diversification benefit - "a merger does not create extra risk".
- aggregation - manipulation of risk: $X \rightarrow Y + Z$;
- capturing the tail;
- convex optimization and capital allocation.
Subadditivity

Subadditivity has to do with

- diversification benefit - "a merger does not create extra risk".
- aggregation - manipulation of risk: $X \to Y + Z$
- capturing the tail;
- convex optimization and capital allocation.

It is questioned from different aspects:

- aggregation penalty - convex risk measures;
- robustness and backtesting;
- financial practice - "a merger creates extra risk".
How Superadditive Can a Risk Measure be?

Question: given an non-subadditive risk measure,

How superadditive can it be?
How Superadditive Can a Risk Measure be?

Question: given an non-subadditive risk measure,

How superadditive can it be?

Motivation:

- Measure model uncertainty.
- Quantify worst-scenarios.
- Trade subadditivity for statistical advantages such as robustness or elicitability.
- Understand better about subadditivity.
For a law-invariant (always assumed) risk measure $\rho$, and risks $X = (X_1, \cdots, X_n)$, the **diversification ratio** is defined as

$$\Delta^X(\rho) = \frac{\rho(X_1 + \cdots + X_n)}{\rho(X_1) + \cdots + \rho(X_n)}.$$ 

For the moment, the denominator is assumed positive.
Diversification ratio

For a law-invariant (always assumed) risk measure $\rho$, and risks $X = (X_1, \cdots, X_n)$, the diversification ratio is defined as

$$\Delta^X(\rho) = \frac{\rho(X_1 + \cdots + X_n)}{\rho(X_1) + \cdots + \rho(X_n)}.$$ 

For the moment, the denominator is assumed positive.

- $\Delta^X(\rho)$ is important in modeling portfolios.
- We want to know how large $\Delta^X(\rho)$ can be.
- $\Delta^X(\rho) \leq 1$ for subadditive risk measures.
- We cannot take a supremum over all possible $X$, which often explodes for any non-superadditive risk measure.
Diversification ratio

We define a law-invariant version of the diversification ratio:

\[ \Delta^F_n(\rho) = \sup \left\{ \frac{\rho(X_1 + \cdots + X_n)}{\rho(X_1) + \cdots + \rho(X_n)} : X_1, \ldots, X_n \sim F \right\}. \]

Here we assumed homogeneity in \( F_i \) for:

- mathematical tractability;
- that it makes sense to let \( n \) vary;
- that it also catches superadditivity of \( \rho \) for inhomogeneous portfolio.
Define

\[ \mathcal{S}_n(F) = \{ X_1 + \cdots + X_n : X_1, \ldots, X_n \sim F \} . \]

Let \( X_F \sim F \). Then

\[ \Delta_n^F(\rho) = \frac{1}{n \rho(X_F)} \sup \{ \rho(S) : S \in \mathcal{S}_n(F) \} . \]

- Known to be a difficult problem; explicit solution for \( \Delta_n^F(\text{VaR}_p) \) (under some strong conditions) given in W., Peng and Yang (2013).
- Numerical calculation for \( \Delta_n^F(\text{VaR}_p) \) given in Embrechts, Puccetti and Rüschendorf (2013).
We are interested in the global superadditivity ratio

$$\Delta^F(\rho) = \sup_{n \in \mathbb{N}} \Delta^F_n(\rho) = \sup_{n \in \mathbb{N}} \frac{1}{n \rho(X_F)} \sup \{\rho(S) : S \in \mathcal{S}_n(F)\}.$$ 

$$\Delta^F(\rho)$$ characterizes how superadditive $$\rho$$ can be for a fixed $$F$$. 
Extreme-aggregation measure

We are interested in the global superadditivity ratio

$$\Delta^F(\rho) = \sup_{n \in \mathbb{N}} \Delta^F_n(\rho) = \sup_{n \in \mathbb{N}} \frac{1}{n \rho(X_F)} \sup \{ \rho(S) : S \in \mathcal{G}_n(F) \}.$$ 

$\Delta^F(\rho)$ characterizes how superadditive $\rho$ can be for a fixed $F$.

The real mathematical target:

$$\sup_{n \in \mathbb{N}} \frac{1}{n} \sup \{ \rho(S) : S \in \mathcal{G}_n(F) \}.$$ 

A closely related quantity:

$$\limsup_{n \to \infty} \frac{1}{n} \sup \{ \rho(S) : S \in \mathcal{G}_n(F) \}.$$
Extreme-aggregation measure

**Definition**

An extreme-aggregation measure induced by a law-invariant risk measure $\rho$ is defined as

$$\Gamma_\rho : \mathcal{X} \to [-\infty, \infty], \quad \Gamma_\rho(X_F) = \limsup_{n \to \infty} \frac{1}{n} \sup \{\rho(S) : S \in \mathcal{S}_n(F)\}.$$
Extreme-aggregation measure

Definition

An extreme-aggregation measure induced by a law-invariant risk measure $\rho$ is defined as

$$\Gamma_\rho : \mathcal{X} \to [-\infty, \infty], \quad \Gamma_\rho(X_F) = \limsup_{n \to \infty} \frac{1}{n} \sup \{ \rho(S) : S \in \mathcal{G}_n(F) \}.$$ 

- $\Gamma_\rho$ quantifies the limit of $\rho$ for worst-case aggregation under dependence uncertainty.
- $\Gamma_\rho$ is a law-invariant risk measure.
Proposition

If $\rho$ is (i) comonotonic additive, or (ii) convex and zero-normalized, then

$$\Gamma_{\rho}(X_F) = \sup_{n \in \mathbb{N}} \frac{1}{n} \sup \{ \rho(S) : S \in \mathcal{G}_n(F) \} \geq \rho(X_F)$$

If $\rho$ is subadditive then $\Gamma_{\rho} \leq \rho$. If it also satisfies (i), or (ii), then $\Gamma_{\rho} = \rho$. 
Proposition

If \( \rho \) is (i) comonotonic additive, or (ii) convex and zero-normalized, then

\[
\Gamma_\rho(X_F) = \sup_{n \in \mathbb{N}} \frac{1}{n} \sup \{ \rho(S) : S \in \mathcal{G}_n(F) \} \geq \rho(X_F)
\]

If \( \rho \) is subadditive then \( \Gamma_\rho \leq \rho \). If it also satisfies (i), or (ii), then \( \Gamma_\rho = \rho \).

Remark

\( \Gamma_\rho \) inherits monotonicity, cash-invariance, positive homogeneity, subadditivity, convexity, or zero-normalization from \( \rho \) if \( \rho \) has the corresponding properties.
Question: given a non-subadditive risk measure \( \rho \),

What is \( \Gamma_\rho \)?
Question: given a non-subadditive risk measure $\rho$, 

What is $\Gamma_\rho$?

- Known motivating result (Wang and W., 2014): as $n \to \infty$, 

$$ \frac{\sup\{\text{VaR}_\rho(S) : S \in \mathcal{G}_n(F)\}}{\sup\{\text{ES}_\rho(S) : S \in \mathcal{G}_n(F)\}} \to 1. $$

Note that

$$ \sup\{\text{ES}_\rho(S) : S \in \mathcal{G}_n(F)\} = n\text{ES}_\rho(X_F), $$

leading to $\Gamma_{\text{VaR}_\rho} = \Gamma_{\text{ES}_\rho} = \text{ES}_\rho$. 
Distortion risk measures:

\[ \rho(X_F) = \int_0^1 F^{-1}(t)dh(t). \]

\( h \): probability measure on \((0, 1)\). A distortion function.
Distortion risk measures

Distortion risk measures:

\[ \rho(X_F) = \int_0^1 F^{-1}(t)dh(t). \]

\( h \): probability measure on \((0,1)\). A distortion function.

- We assume random variables are bounded from below: \( F^{-1}(0) > -\infty \).
- Standard risk measures are comonotonic additive if and only if it is a distortion risk measure (a property of Choquet integral; see Yaari, 1987).
- ES, VaR are special cases of distortion risk measures.
Distortion risk measures

Theorem (Extreme-aggregation for distortion risk measures)

Suppose $\rho$ is a distortion risk measure with distortion function $h$, then $\Gamma_\rho$ is

(a) the smallest coherent risk measure dominating $\rho$;
(b) a coherent distortion risk measure with a distortion function as the largest convex distortion function dominated by $h$. 

Example: $\Gamma_{\text{VaR}} p = \text{ES} p$. A proof quite complicated.
Distortion risk measures

Theorem (Extreme-aggregation for distortion risk measures)

Suppose $\rho$ is a distortion risk measure with distortion function $h$, then $\Gamma_\rho$ is

(a) the smallest coherent risk measure dominating $\rho$;
(b) a coherent distortion risk measure with a distortion function as the largest convex distortion function dominated by $h$.

- Example: $\Gamma_{\text{VaR}_p} = \text{ES}_p$.
- A proof quite complicated.
Main result for distortion risk measures

For distortion risk measures,

\[ \Delta^F(\rho) = \frac{\Gamma_{\rho}(X_F)}{\rho(X_F)}. \]
Main result for distortion risk measures

For distortion risk measures,

\[ \Delta^F(\rho) = \frac{\Gamma_\rho(X_F)}{\rho(X_F)} . \]

**Theorem (Coherence and extreme-aggregation)**

Suppose \( \rho \) is distortion risk measure. The following are equivalent:

(a) \( \rho \) is coherent.

(b) \( \Gamma_\rho(X_F) = \rho(X_F) \) for all distributions \( F \).

(c) \( \Gamma_\rho(X_F) = \rho(X_F) \) for some continuous distribution \( F \), \( \rho(X_F) < \infty \).

(d) \( \Delta^F(\rho) = 1 \) for all distributions \( F \), \( \rho(X_F) \in (0, \infty) \).

(e) \( \Delta^F(\rho) = 1 \) for some continuous distribution \( F \), \( \rho(X_F) \in (0, \infty) \).
Shortfall risk measures:

\[ \rho(X) = \inf\{y \in \mathbb{R} : \mathbb{E}[\ell(X - y)] \leq l(0)\} \text{.} \]

\( \ell \): convex and increasing function. A loss function.

- Motivation from indifference pricing theory.
Theorem (Extreme-aggregation for shortfall risk measures)

Suppose \( \rho \) is a shortfall risk measure with loss function \( \ell \), then \( \Gamma_\rho \) is

(a) the smallest coherent risk measure dominating \( \rho \);

(b) a coherent \( p \)-expectile, where

\[
p = \lim_{x \to \infty} \ell'(x) / \left( \lim_{x \to \infty} \ell'(x) + \lim_{x \to -\infty} \ell'(x) \right)
\]
Theorem (Extreme-aggregation for shortfall risk measures)

Suppose \( \rho \) is a shortfall risk measure with loss function \( \ell \), then \( \Gamma_\rho \) is

(a) the smallest coherent risk measure dominating \( \rho \);
(b) a coherent \( p \)-expectile, where

\[
p = \lim_{x \to \infty} \ell'(x)/(\lim_{x \to \infty} \ell'(x) + \lim_{x \to -\infty} \ell'(x))
\]

- Example: \( \Gamma_{\text{ER}_\beta} = e_1 = \text{ess-sup} \), where \( \text{ER}_\beta \) is the entropy risk measure: with loss function \( \ell(x) = \exp(\beta x) - 1 \).
- A proof of one page.
A convex risk measure is standard and convex:

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y), \quad \lambda \in (0, 1).$$

**Theorem (Extreme-aggregation for convex risk measures)**

Suppose $\rho$ is a convex risk measure, then $\Gamma_{\rho}$ is a coherent risk measure.
Convex risk measures

A law-invariant convex risk measure has the following representation

\[ \rho = \sup_{h \in \mathcal{P}[0,1]} \left\{ \int \text{ES}_p \, dh(p) - v(h) \right\}, \]

where

- \( \mathcal{P}[0,1] \) is the set of all probability measures on \([0, 1]\);
- \( v : \mathcal{P}[0,1] \rightarrow \mathbb{R} \cup \{ +\infty \} \) is a convex function;
- \( \rho(0) = 0 \) is equivalent to \( \inf_{h \in \mathcal{P}[0,1]} v(h) = 0 \).
Extreme-aggregation for convex risk measures

\( \Gamma_\rho \) is a coherent risk measure with representation

\[
\Gamma_\rho = \sup_{h \in Q} \left\{ \int ES_p dh(p) \right\},
\]

where \( Q = \{ h \in \mathcal{P}[0, 1] : v(h) > -\infty \} \).

- \( \Gamma_\rho \) is the smallest coherent risk measure dominating \( \rho \).
We define the extreme-division measure \( W., 2014 \) for a risk measure \( \rho \):

\[
\Psi_\rho(X) = \inf \left\{ \sum_{i=1}^{n} \rho(X_i) : n \in \mathbb{N}, X_i \in \mathcal{X}, i = 1, \ldots, n, \sum_{i=1}^{n} X_i = X \right\}.
\]

- \( \Psi_\rho(X) \) is the least amount of capital requirement according to \( \rho \) if the risk \( X \) can be divided arbitrarily.
- \( \Psi_\rho \leq \rho \).
- \( \Psi_\rho = \rho \) for subadditive risk measures.
- Not relevant to this topic, just wanted to show an interesting duality.
Extreme-division for convex risk measures

Ψ_ρ is a coherent risk measure with representation

\[ Ψ_ρ = \sup_{h \in Q} \left\{ \int \text{ES}_p dh(p) \right\} , \]

where \( Q = \{h \in \mathcal{P}[0, 1] : v(h) = 0\} \).

- \( Ψ_ρ \) is the largest coherent risk measure dominated by \( ρ \).
- When \( ρ \) is a distortion risk measure, \( Ψ_ρ \) is a coherent risk measure, but not necessarily a distortion.
- \( Ψ_{\text{VaR}_p} = -\infty \) for all \( p \in (0, 1) \).
Discussion

- $\Gamma_\rho$ is very often a coherent risk measure for all commonly used standard risk measures. However, counter-example can be built.

- $\Gamma_\rho$ often gains positive homogeneity, convexity, and subadditivity even if $\rho$ does not have these properties.

- A universal axiomatic proof of this phenomenon is not available yet.

- Characterize the class of risk measures which induce coherent extreme-aggregation measures?

- What happens to shortfall risk measures with non-convex loss functions?
Some take-home message:

Coherence is indeed a natural property desired by a good risk measure. Even when a non-coherent risk measure is applied to a portfolio, its extreme behavior under dependence uncertainty leads to coherence.

When we allow arbitrary division of a risk, the extreme behavior also leads to coherence.

This contributes to the Basel question and partly supports the use of coherent risk measures.
References I


Thank you for your kind attendance!
**R1, Page 20, *Choice of risk metric*:

“... However, a number of weaknesses have been identified with VaR, including its inability to capture “tail risk”. The Committee therefore believes it is necessary to consider alternative risk metrics that may overcome these weaknesses.”
We focus on the mathematical and statistical aspects, avoiding discussion on practicalities and operational issues.

R1, Page 3:

“"The Committee recognises that moving to ES could entail certain operational challenges; nonetheless it believes that these are outweighed by the benefits of replacing VaR with a measure that better captures tail risk.""
R2, Page 3, Approach to risk management:

"the Committee has its intention to pursue two key confirmed reforms outlined in the first consultative paper [May 2012]: Stressed calibration . . . Move from Value-at-Risk (VaR) to Expected Shortfall (ES)."