Robustness and sensitivity analysis of risk measurement procedures

Rama Cont - Columbia University

Romain Deguest - Ecole Polytechnique Paris

Giacomo Scandolo - University of Florence

- How to detect and quantify **model risk** in the measurement of risk?
- When is a risk measure ”***robust***” with respect to small changes of the underlying model (i.e. the distribution of the risk factors) ?
- A **risk measure** is just one ingredient of a **risk measurement procedure**. The **estimation method** must be taken into account as well.
- Different results under different estimation methods. In general, coherency and robustness seem to be conflicting properties.

$L \subseteq L^0(\Omega, \mathcal{F}, \mathbb{P})$: subspace of r.v. containing all the constants

Definitions

- **Risk measure**: $\rho : L \to \mathbb{R}$
- **Coherent risk measure** (Artzner et al, '99): ρ satisfying:
  1. *Translation invariance*: $\rho(X+m) = \rho(X) - m \quad \forall m \in \mathbb{R}$;
  2. *Monotonicity*: $\rho(X) \leqslant \rho(Y)$ whenever $X \geqslant Y$
  3. *Positive homogeneity*: $\rho(\lambda X) = \lambda \rho(X) \quad \forall \lambda \geqslant 0$
  4. *Subadditivity*: $\rho(X+Y) \leqslant \rho(X) + \rho(Y)$.
- **Law-invariant risk measure**: ρ satisfying

$$F_X = F_Y \quad \Rightarrow \quad \rho(X) = \rho(Y).$$

(so that we will write $\rho(X) = \rho(F)$ when $F = F_X$)

We focus on the following class of risk measures:

$$\boxed{\rho_m(X) \triangleq -\int_0^1 q_u(X)\, m(du)}$$

where

- m is a normalized (non-negative) measure on $(0,1)$ and
- $q_\alpha(X) = \inf\{x \in \mathbb{R} : F_X(x) > \alpha\}$ is the (upper) quantile of order $\alpha \in (0,1)$.

For any m, ρ_m satisfies 1,2,3. It is subadditive (hence coherent) if and only if:

1. m is absolutely continuous w.r.t. Lebesgue measure, i.e.

   $$m(du) = \phi(u)du$$

   for some density ϕ, and
2. the map $u \mapsto \phi(u)$ is "decreasing".

1. $m = \delta_\alpha$ (for a fixed $\alpha \in (0,1)$) corresponds to **Value at Risk**:

$$\boxed{\mathbb{V}\mathrm{aR}_\alpha(X) = -q_\alpha(X)}$$

2. $m(du) = \frac{1}{\alpha} I_{(u \leqslant \alpha)}\, du$ corresponds to **Average Value at Risk**:

$$\boxed{\mathbb{AV}\mathrm{aR}_\alpha(X) = \int_0^\alpha \mathbb{V}\mathrm{aR}_u(X)\, du}$$

3. more generally, $m(du) = \phi(u)\, du$ with ϕ decreasing corresponds to a **spectral risk measure** (of **spectrum** ϕ)

$$\boxed{\rho_\phi(X) = \int_0^\alpha \mathbb{V}\mathrm{aR}_u(X)\phi(u)\, du}$$

- Spectral risk measures are subadditive (hence coherent), VaR is not.

- Spectral risk measures are a fairly broad class
  - ”All” comonotonic additive and coherent r.m. are spectral r.m. [Foellmer/Schied, ’03]
  - All coherent r.m. take the (more general) form:

$$\rho(X) = \sup_{\phi \in \mathcal{A}} \rho_\phi(X)$$

    for some subset $\mathcal{A}$ of spectra. [Kusuoka, ’02]

- VaR is defined over all L^0, spectral risk measures are not (for instance, AVaR is only defined on L^1). No coherent risk measure can be defined over all L^0 [Delbaen, ’02].

Coming to estimation, if $X = h(\mathbf{Z})$, where $\mathbf{Z} \in \mathbb{R}^K$ are risk factors:

	step 1		step 2		step 3	
Data	$\longrightarrow$	Factors Df	$\longrightarrow$	Portfolio Df	$\longrightarrow$	Risk estimate
$(\mathbf{z}^1, \ldots, \mathbf{z}^N)$		$\widehat{F}_{\mathbf{Z}} \in \mathcal{D}(\mathbb{R}^K)$		$\widehat{F}_{h(\mathbf{Z})} \in \mathcal{D}(\mathbb{R})$		$\rho(\widehat{F}_{h(\mathbf{Z})}) \in \mathbb{R}$

Step 3 depends only on the chosen risk measure ρ. Three different approaches for steps 1 and 2:

Historical. No assumptions are made about the underlying model (except that the future reflects the past!)

Parametric The underlying model belongs to a parametric family and parameters are fitted to data.

Monte Carlo. Typically used in the second step when portfolio is non-linear.

For simplicity we consider $K = 1$ and $h = id$.
Moreover, $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ are realizations of N IID r.v. distributed as X.

	step 1		step 2	
Data	$\longrightarrow$	Portfolio Df	$\longrightarrow$	Risk estimate
$\mathbf{x} \in \mathbb{R}^N$		$\widehat{F}_X \in \mathcal{D}(\mathbb{R})$		$\rho(\widehat{F}_X) = \widehat{\rho}(\mathbf{x}) \in \mathbb{R}$

1. $\mathbf{x} \mapsto \widehat{F}_X \in \mathcal{D}(\mathbb{R})$ is the estimator of the model

2. $\widehat{\rho}$ is the **risk estimator** associated to the couple $(\widehat{F}, \rho)$ (a risk measurement procedure)

3. we shall study properties of $\widehat{\rho}$ as an estimator of ρ

Fix a risk measure ρ. Good properties of a risk estimator $\widehat{\rho}$ are:

Consistency : As N increases, the probability of large estimation errors vanishes.

Qualitative robustness the distribution of $\widehat{\rho}$ is "stable" w.r.t. a small change in the true underlying distribution.

Quantitative robustness (or good sensitivity properties) $\widehat{\rho}$ is stable w.r.t. the addition of an observation, when n is already large.

Warning: Neither VaR nor any coherent risk measure is everywhere continuous w.r.t. convergence in law (Bauerle-Muller, '05)

Let

$$F^{\mathbf{x}}(x) = \sum_{n=1}^{N} I\{x \leqslant x_n\}$$

be the empirical df associated with datum $\mathbf{x} \in \mathbb{R}^N$.

For a fixed risk measure ρ the **historical risk estimator** is

$$\boxed{\widehat{\rho}_h(\mathbf{x}) = \rho(F^{\mathbf{x}}) \qquad \mathbf{x} \in \mathbb{R}^N}$$

1. The historical estimator of $\mathbb{V}\mathrm{aR}_\alpha$, $\alpha \in (0,1)$ is:

$$\widehat{\rho}_h(\mathbf{x}) = \mathbb{V}\mathrm{aR}_\alpha(F^{\mathbf{x}}) = -x_{([n\alpha]+1)}, \qquad \mathbf{x} = (x_1, \ldots, x_n),$$

2. The historical estimator of $\mathbb{A}\mathbb{V}\mathrm{aR}_\alpha$ is:

$$\widehat{\rho}_h(\mathbf{x}) = \mathbb{A}\mathbb{V}\mathrm{aR}_\alpha(F^{\mathbf{x}}) = -\frac{1}{n\alpha}\left(\sum_{i=1}^{[n\alpha]} x_{(i)} + x_{([n\alpha]+1)}(n\alpha - [n\alpha])\right).$$

3. The historical estimator of the spectral risk measure ρ_ϕ is:

$$\widehat{\rho}_h(\mathbf{x}) = \rho_\phi(F^{\mathbf{x}}) = -\sum_{i=1}^{n} w_i\, x_{(i)}, \qquad \mathbf{x} = (x_1, \ldots, x_n),$$

where, for any i

$$w_i \triangleq \int_{(i-1)/n}^{i/n} \phi(u)\, du.$$

Fixed: a risk measure ρ, a risk estimator $\widehat{\rho}$, a df F.

The estimator $\widehat{\rho}$ is **consistent** with ρ at F if

$$\operatorname*{p-lim}_{N\to\infty} \widehat{\rho}(X_1, \ldots, X_N) = \rho(F),$$

where (X_n) is a sequence of IID r.v. distributed as F.

- The historical estimator of $\mathbb{V}\mathrm{aR}_\alpha$ is consistent at F iff F has a unique quantile of order α. The ”if” part is a consequence of the Glivenko-Cantelli Theorem.

- The historical estimator of $\mathbb{A}\mathbb{V}\mathrm{aR}_\alpha$ or of any spectral r.m. ρ_ϕ is consistent at any F where ρ is defined. Consequence of a general result about L-estimators.

The weak topology on $\mathcal{D}$ is metrized by the Levy distance $d = d_L$, or even by the Prohorov distance $d = d_P$.

If $\widehat{\rho}$ is a risk estimator and $F \in \mathcal{D}$, then the sampling df is

$$\mathcal{L}_N(\widehat{\rho}, F) = \text{Law}\{\widehat{\rho}(X_1, \ldots, X_N)\},$$

where (X_n) are IID with df F.

Definition [see Hampel, '71]

Given a subset $\mathcal{C} \subseteq \mathcal{D}$, a risk estimator $\widehat{\rho}$ is $\mathcal{C}$-**robust** at $F \in \mathcal{C}$ if

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta > 0 \;\; \text{s.t.}$$

$$d(G, F) \leqslant \delta, \quad G \in \mathcal{C} \quad \Rightarrow \quad \sup_{N \geqslant 1} d(\, \mathcal{L}_N(\widehat{\rho}, G),\ \mathcal{L}_N(\widehat{\rho}, F)\,) \leqslant \varepsilon.$$

Equivalently, if the sequence of maps $\mathcal{L}_N(\widehat{\rho}, \cdot) : \mathcal{D} \to \mathcal{D}$ is equi-continuous at F.

Theorem

Given a risk measure ρ, a set $\mathcal{C} \subset \mathcal{D}$ and a df $F \in \mathcal{C}$; assume that $\widehat{\rho}_h$ is consistent with ρ in a ($\mathcal{C}$-)neighborhood of F. Then:

$$\widehat{\rho}_h \text{ is } \mathcal{C}\text{-robust at } F \iff \rho|_{\mathcal{C}} \text{ is continuous at } F$$

(when $\mathcal{C} = \mathcal{D}$ see Hampel, '71)

Let:

$$\rho_m(F) = \int_0^1 \mathbb{V}\mathrm{aR}_u(F)\, dm(u),$$

and

$$\mathcal{C} = \{F \in \mathcal{D} \,:\, F \text{ has a unique quantile of order } \alpha \text{ whenever } m(\{\alpha\}) > 0\}$$

Theorem [Huber, '81]

Let $\alpha_- \leqslant \alpha_+$ be the upper and lower extremes of the support of m

1. *if $\alpha_-, \alpha_+ \in (0,1)$ then $\rho_m|_{\mathcal{C}}$ is continuous at any $F \in \mathcal{C}$;*
2. *if $\alpha_- = 0$ or $\alpha_+ = 1$, then $\rho_m|_{\mathcal{C}}$ is not continuous at any $F \in \mathcal{C}$.*

As a consequence:

- the historical estimator of $\mathbb{V}\mathrm{aR}_\alpha$ is $\mathcal{C}$-robust at any $F \in \mathcal{C}$, where

$$\mathcal{C} = \{F \in \mathcal{D} \,:\, F \text{ has a unique quantile of order } \alpha\}$$

- if $\mathcal{C}$ is any open set in $\mathcal{D}$ then:
  - the historical estimator of $\mathbb{A}\mathbb{V}\mathrm{aR}_\alpha$ (and of any other spectral r.m.) is not $\mathcal{C}$-robust at any $F \in \mathcal{C}$
  - if $\varepsilon \in (0, \alpha)$, the historical estimator of:

$$\mathbb{A}\mathbb{V}\mathrm{aR}_\alpha^\varepsilon(F) \triangleq \frac{1}{\alpha - \varepsilon} \int_\varepsilon^\alpha \mathbb{V}\mathrm{aR}_u(F)\, du.$$

    is $\mathcal{C}$-robust at any $F \in \mathcal{C}$

Given a risk measure ρ.
Fix a parametric family $(F_\theta)_{\theta \in \Theta}$ and an estimator of θ, $\widehat{\theta} : \mathbb{R}^N \to \Theta$.

The **parametric estimator** of ρ is:

$$\boxed{\widehat{\rho}(\mathbf{x}) = \rho(F_\theta), \quad \text{for } \theta = \widehat{\theta}(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^N,}$$

We focus on M-estimators for θ, coming in the form

$$\widehat{\theta}(\mathbf{x}) = \text{arg.max}_{\theta \in \Theta} \sum_{n=1}^{N} \Psi(x_n, \theta),$$

for some $\Psi = \Psi(x, \theta)$.

The **MLE** corresponds to $\Psi(x, \theta) = \log f_\theta(x)$, where f_θ is the density of F_θ.

Many families of practical interest are of **scale-location** type:

$$\mathcal{D}_F = \{F_{\mu,\sigma} \,:\, \mu \in \mathbb{R},\, \sigma > 0\},$$

where:

- F is a given initial df (usually of mean 0 and variance 1), for instance:

  1. standard gaussian
  2. double exponential: $f(x) = e^{-|x|}/2$
  3. Cauchy: $f(x) = (\pi(1+x^2))^{-1}$

- $F_{\mu,\sigma}$ is defined as

$$F_{\mu,\sigma}(x) \triangleq F\left(\frac{x-\mu}{\sigma}\right)$$

  or, equivalently,

$$f_{\mu,\sigma}(x) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$$

Let $\mathcal{D}_F$ be a scale location family. As any ρ_m is PH and TI:

$$\rho_m(F_{\mu,\sigma}) = -\mu + \sigma\rho(F).$$

If $\widehat{\mu}$ and $\widehat{\sigma}$ are the MLE of μ and σ, then the MLE of ρ_m is

$$\widehat{\rho}(\mathbf{x}) = -\widehat{\mu} + \widehat{\sigma}\rho(F)$$

$c = \rho(F)$ is independent of the data; examples ($\alpha < .5$):

$$\begin{aligned} \mathbb{V}\mathrm{aR}_\alpha(F) &= |z_\alpha| && \text{(gaussian case)} \\ &= -\log(2\alpha) && \text{(d. exp. case)} \\ &= \tan(\pi(\alpha - 1/2)) && \text{(Cauchy case)} \end{aligned}$$

Let $\widehat{\rho}(\mathbf{x}) = \rho(F_{\widehat{\theta}(\mathbf{x})})$ be a parametric risk estimator.
We define the **effective risk measure**

$$\rho^{\text{eff}}(F) = \operatorname*{p-lim}_{N\to\infty} \widehat{\rho}(X_1, \ldots, X_N), \quad (X_n) \text{ IID } \sim F,$$

at any F where the limit exists and is a constant.

Example: the effective r.m. associated to the gaussian MLE of $\mathbb{V}\mathrm{aR}_\alpha$ is:

$$\rho^{\text{eff}}(F) = -\mu(F) + |z_\alpha|\,\sigma(F), \quad F \in \mathcal{D}^2.$$

Immediate, but crucial remarks:

- $\rho^{\text{eff}}(F^{\mathbf{x}}) = \widehat{\rho}(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^N$
- $\widehat{\rho}$ is the historical estimator of ρ^{eff} and is consistent with it at any F (where ρ^{eff} is defined)
- $\rho^{\text{eff}} \equiv \rho$ on the parametric family, but they may be otherwise very different

Let $\mathcal{C} \subseteq \mathcal{D}$ be fixed. In order to assess $\mathcal{C}$-robustness of a risk estimator $\widehat{\rho}$ at some $F \in \mathcal{C}$ we have to:

- compute the effective risk measure ρ^{eff}
- assess whether $\rho^{\text{eff}}|_{\mathcal{C}}$ is continuous at F
- $\widehat{\rho}$ is the historical estimator of ρ^{eff} and is consistent with it: we may apply the previous Th.
- $\widehat{\rho}$ is $\mathcal{C}$-robust at F iff $\rho^{\text{eff}}|_{\mathcal{C}}$ is continuous at F

Choosing $\mathcal{C} = \mathcal{D}_F$ for some F, we obtain:

- The ML parametric estimators of VaR, AVaR and spectral r.m. are not $\mathcal{C}$-robust at any $F \in \mathcal{C}$ when the parametric family is:
  - **gaussian**,
  - **double exponential**.
- It is robust at any $F \in \mathcal{C}$ when the parametric family is **Cauchy** (note however that in this case spectral r.m. are not defined on $\mathcal{C}$!)

risk estimator	Robust ?
Historical VaR	**yes**
Gaussian VaR	**no**
Double exponential VaR	**no**
Historical AVaR (or spectral)	**no**
Gaussian or d. exp. AVaR	**no**
Gaussian or d. exp. spectral	**no**
Cauchy VaR or spectral	**yes**

Two messages:

- We propose an approach to assess robustness of a risk measure: it involves the whole estimation procedure. Different estimation procedures lead to different results.

- In general it seems that coherence and robustness are conflicting requirements.

Further problems:

- Convex risk measures
- Parametric families with more than 2 parameters (e.g. asymmetric exponential)
- Weaker forms of subadditivity that may allow for robustness.

www.dmd.unifi.it/∼scandolo