Comparative and qualitative robustness for law-invariant risk measures

Alexander Schied
Universität Mannheim

LUH-Kolloquium
Versicherungs- und Finanzmathematik
Leibniz Universität Hannover
23. Januar 2014

Based on joint work with
Volker Krätschmer (Duisburg) and Henryk Zähle (Saarbrücken)

1 Introduction

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be atomless and let $\mathscr{X} \subset L^0 := L^0(\Omega, \mathscr{F}, \mathbb{P})$ be a vector space containing the constants. A map $\rho : \mathscr{X} \to \mathbb{R}$ is a convex risk measure when the following conditions are satisfied:

(i) monotonicity: $\rho(X) \geq \rho(Y)$ for $X, Y \in \mathscr{X}$ with $X \leq Y$;

(ii) convexity: $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ for all $X, Y \in \mathscr{X}$ and $\lambda \in [0,1]$;

(iii) cash additivity: $\rho(X+m) = \rho(X) - m$ for $X \in \mathscr{X}$ and $m \in \mathbb{R}$.

Note: it is possible here to replace axiom (iii) by the following weaker notion:

(iii') cash coercivity: $\rho(-m) \longrightarrow +\infty$ when $m \in \mathbb{R}$ tends to $+\infty$.

When ρ is law-invariant,

$$\rho(X) = \rho(\widetilde{X}) \text{ whenever } X \text{ and } \widetilde{X} \text{ have the same law under } \mathbb{P},$$

it makes sense to estimate $\rho(X)$ by means of a Monte Carlo procedure or from a sequence of historical data.

Let

$$\mathscr{M}(\mathscr{X}) := \{\mathbb{P} \circ X^{-1} \,|\, X \in \mathscr{X}\}$$

Law invariance of a risk measure $\rho : \mathscr{X} \to \mathbb{R}$ is equivalent to the existence of a map

$$\mathscr{R}_\rho : \mathscr{M}(\mathscr{X}) \to \mathbb{R}$$

such that

$$\rho(X) = \mathscr{R}_\rho(\mathbb{P} \circ X^{-1}), \qquad X \in \mathscr{X}. \tag{1}$$

This map $\mathscr{R}_\rho$ will be called the risk functional associated with ρ.

When $\widehat{\mu}_n$ estimates the law $\mu = \mathbb{P} \circ X^{-1}$ of X then

$$\widehat{\rho}_n := \mathscr{R}_\rho(\widehat{\mu}_n) \tag{2}$$

is an estimator for $\rho(X)$. A typical choice is the empirical distribution of a sequence $X_1, \ldots, X_n$ of historical observations or Monte Carlo simulations

$$\widehat{\mu}_n = \widehat{m}_n := \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k}$$

Questions:

- **Consistency:** do we have $\widehat{\rho}_n \to \rho(X)$ as $n \uparrow \infty$?
- **Continuity:** is $\mu \mapsto \mathscr{R}_\rho(\mu)$ continuous?
- **Asymptotic analysis:** what can be said about the asymptotic distribution of the estimation error $\widehat{\rho}_n - \rho(X)$?
- **Robustness:** is the law of $\widehat{\rho}_n$ stable with respect to small perturbations of the law generating the $X_1, \ldots, X_n$?

When $\widehat{\mu}_n$ estimates the law $\mu = \mathbb{P} \circ X^{-1}$ of X then

$$\widehat{\rho}_n := \mathscr{R}_\rho(\widehat{\mu}_n) \tag{2}$$

is an estimator for $\rho(X)$. Typical choice is the empirical distribution of a sequence $X_1, \dots, X_n$ of historical observations or Monte Carlo simulations

$$\widehat{\mu}_n = \widehat{m}_n := \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k}$$

Questions:

- **Robustness:** is the law of $\widehat{\rho}_n$ stable with respect to small perturbations of the law generating the $X_1, \dots, X_n$?

  Cont, Deguest, and Scandolo (2010):

  When ρ is **coherent**, the risk functional $\mathscr{R}_\rho$ cannot be qualitatively robust in the sense of Hampel (1971). However, $\mathscr{R}_{V@R}$ essentially is qualitatively robust.

Cont, Deguest, and Scandolo (2010):
When ρ is **coherent**, the risk functional $\mathscr{R}_\rho$ cannot be qualitatively robust in the sense of Hampel (1971). However, $\mathscr{R}_{V@R}$ essentially is qualitatively robust.

Sounds like a strong argument in favor of $V@R$ and against coherent risk measures.

Cont, Deguest, and Scandolo (2010):
When ρ is **coherent**, the risk functional $\mathscr{R}_\rho$ cannot be qualitatively robust in the sense of Hampel (1971). However, $\mathscr{R}_{V@R}$ essentially is qualitatively robust.

Sounds like a strong argument in favor of $V@R$ and against coherent risk measures.

But let's review the argument.

- The argument is weak when compared to the main argument againt $V@R$: $V@R$ creates toxic incentives.

Cont, Deguest, and Scandolo (2010):
When ρ is **coherent**, the risk functional $\mathscr{R}_\rho$ cannot be qualitatively robust in the sense of Hampel (1971). However, $\mathscr{R}_{V@R}$ essentially is qualitatively robust.

Sounds like a strong argument in favor of $V@R$ and against coherent risk measures.

But let's review the argument.

- The argument is weak when compared to the main argument againt $V@R$: $V@R$ creates toxic incentives.

- By Hampel's theorem, qualitative robustness of $\mathscr{R}_\rho$ implies that $\mathscr{R}_\rho$ is continuous with respect to weak convergence in $\mathscr{M}_1(\mathbb{R})$.

  Since the compactly supported probability measures are dense in $\mathscr{M}_1(\mathbb{R})$ with respect to weak convergence, Hampel's theorem thus implies that $\mathscr{R}_\rho(\mu)$ must be insensitive to the tail behavior of μ

Cont, Deguest, and Scandolo (2010):
When ρ is **coherent**, the risk functional $\mathscr{R}_\rho$ cannot be qualitatively robust in the sense of Hampel (1971). However, $\mathscr{R}_{V@R}$ essentially is qualitatively robust.

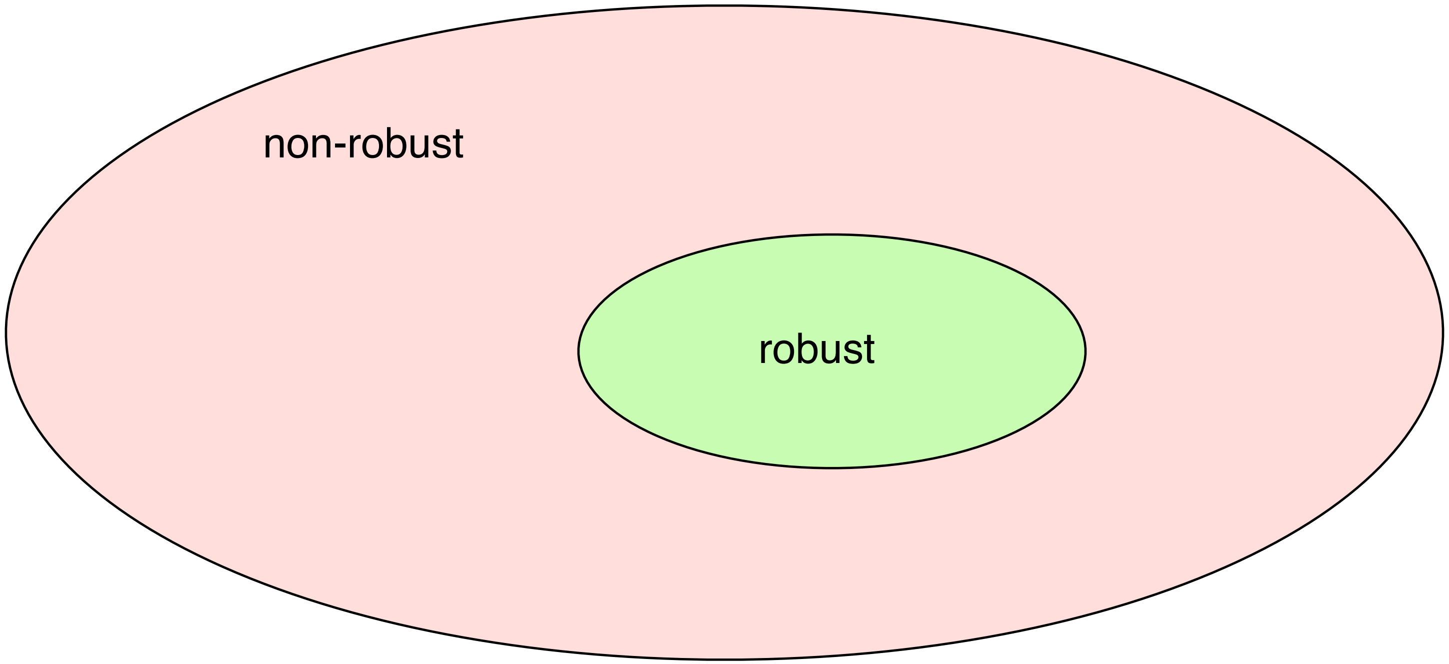
Sounds like a strong argument in favor of $V@R$ and against coherent risk measures.

But let's review the argument.

- The argument is weak when compared to the main argument againt $V@R$: $V@R$ creates toxic incentives.
- By Hampel's theorem, qualitative robustness of $\mathscr{R}_\rho$ implies that $\mathscr{R}_\rho$ is continuous with respect to weak convergence in $\mathscr{M}_1(\mathbb{R})$.

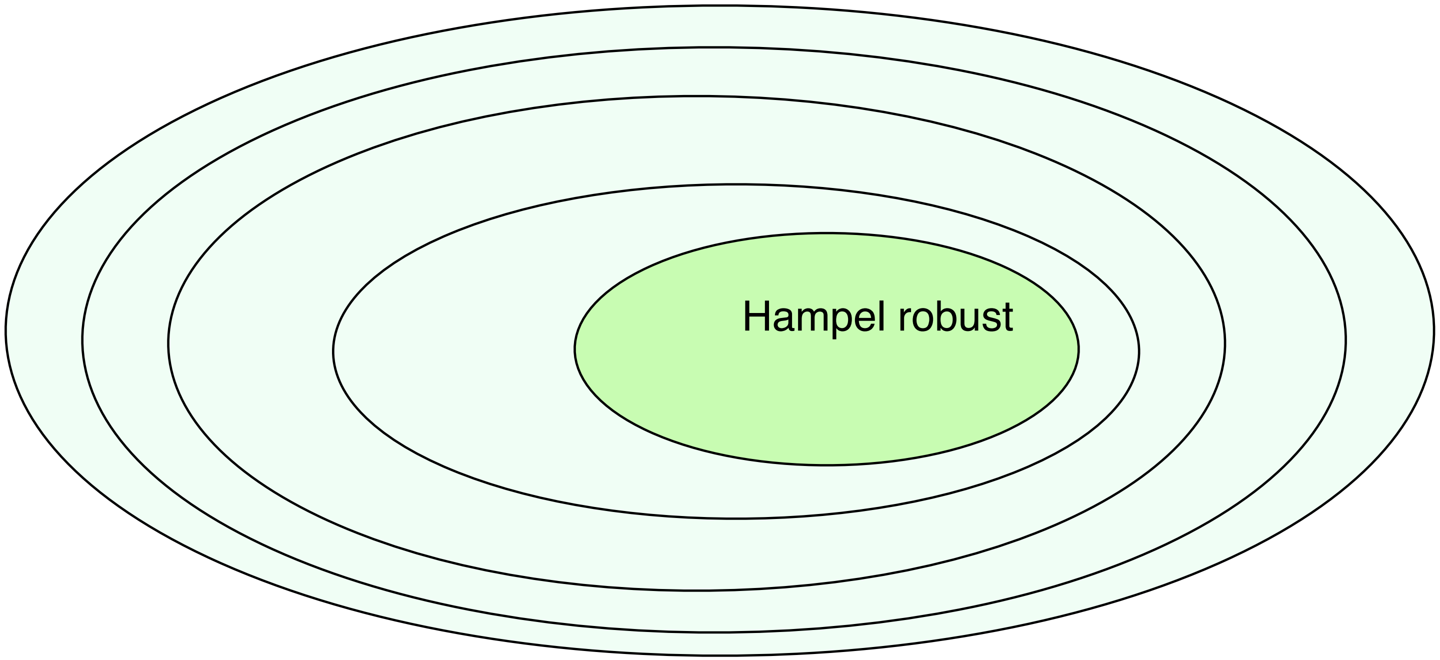
  Since the compactly supported probability measures are dense in $\mathscr{M}_1(\mathbb{R})$ with respect to weak convergence, Hampel's theorem thus implies that $\mathscr{R}_\rho(\mu)$ must be insensitive to the tail behavior of μ.
- $\rho(X) = -\mathbb{E}[\,X\,]$ is a coherent risk measure and thus corresponds to a non-robust risk functional. But $\rho(X) = -\mathbb{E}[\,X\,]$ is, e.g., the most common and natural actuarial risk measure.

- Hampel's terminology of qualitative robustness generates a sharp division of risk functionals into those that are called "robust" and others that are called "not robust".



But, e.g., estimating the expected value should be "more robust" than estimating variance.

Can one thus define a **refined** notion of robustness that produces a picture like this one:



By such a refined notion of robustness one can also try to capture the natural tradeoff between robustness and tail-sensitivity

Such a refined version of robustness was introduced in Krätschmer, A.S., and Zähle (2012), and here we apply it to, and explain it at the hand of, law-invariant convex risk measures (Krätschmer, A.S., and Zähle, 2013)

Such a refined version of robustness was introduced in Krätschmer, A.S., and Zähle (2012), and here we apply it to, and explain it at the hand of, law-invariant convex risk measures (Krätschmer, A.S., and Zähle, 2014)

The key is to replace metrics for the weak topology in Hampel's robustness with metrics for the ψ-weak topology on

$$\mathscr{M}_1^{\psi} := \mathscr{M}_1^{\psi}(\mathbb{R}) := \left\{\mu \in \mathscr{M}_1(\mathbb{R}) \,\Big|\, \int \psi \, d\mu < \infty\right\}$$

where $\psi : \mathbb{R} \to [0, \infty)$ is a continuous weight function satisfying $\psi \geq 1$ outside some compact set.

Typical example: $\psi(x) = |x|^p$

We have

$$\mu_n \longrightarrow \mu \ \psi\text{-weakly} \quad :\Longleftrightarrow \quad \int f\,d\mu_n \longrightarrow \int f\,d\mu \ \forall \text{ continuous } f \text{ with } |f| \leq c(1+\psi)$$

$$\Longleftrightarrow \quad \mu_n \longrightarrow \mu \text{ weakly and } \int \psi\,d\mu_n \longrightarrow \int \psi\,d\mu$$

A suitable metric is

$$d_\psi(\mu,\nu) := d_{\mathrm{Proh}}(\mu,\nu) + \left| \int \psi\,d\mu - \int \psi\,d\nu \right|$$

The ψ-weak topology coincides with the weak topology iff ψ is bounded.

2 Preliminaries

The choice $\mathscr{X} := L^\infty := L^\infty(\Omega, \mathscr{F}, \mathbb{P})$ is not suitable when dealing with possibly unbounded risks. Better: Orlicz spaces or Orlicz hearts (S. Biagini and Frittelli (2008), Cheridito and Li (2009)).

A Young function will be a left-continuous, nondecreasing convex function $\Psi : \mathbb{R}_+ \to [0, \infty]$ such that $0 = \Psi(0) = \lim_{x \downarrow 0} \Psi(x)$ and $\lim_{x \uparrow \infty} \Psi(x) = \infty$.

The Orlicz space associated with Ψ is

$$L^\Psi := L^\Psi(\Omega, \mathscr{F}, \mathbb{P}) = \left\{ X \in L^0 \,|\, \mathbb{E}[\, \Psi(c|X|) \,] < \infty \text{ for some } c > 0 \right\}.$$

It is a Banach space when endowed with the Luxemburg norm,

$$\|X\|_\Psi := \inf \{ \lambda > 0 \,|\, \mathbb{E}[\, \Psi(|X|/\lambda) \,] \leq 1 \} \,.$$

The Orlicz heart is defined as

$$H^\Psi := H^\Psi(\Omega, \mathscr{F}, \mathbb{P}) = \left\{ X \in L^0 \,|\, \mathbb{E}[\, \Psi(c|X|) \,] < \infty \textbf{ for all } c > 0 \right\}$$

Cheridito and Li (2009): finite risk measures on H^Ψ are continuous for $\| \cdot \|_\Psi$

For a finite Young function Ψ,

$$L^\infty \subset H^\Psi \subset L^\Psi \subset L^1$$

and these inclusions may all be strict. In fact, the identity $H^\Psi = L^\Psi$ holds if and only if Ψ satisfies the so-called Δ_2-condition

(3) there are C, $x_0 > 0$ such that $\Psi(2x) \leq C\Psi(x)$ for all $x \geq x_0$.

The Δ_2-condition is clearly satisfied when specifically $\Psi(x) = x^p/p$ for some $p \in [1, \infty)$. In this case, $H^\Psi = L^\Psi = L^p$ and $\|Y\|_\Psi = p^{-1/p}\|Y\|_p$.

In the sequel, Ψ will always denote a finite Young function

3 Consistency

For a distortion risk measure ρ, the estimator $\widehat{\rho}_n = \mathscr{R}_\rho(\widehat{m}_n)$ has the form of an *L-statistic* and results by van Zwet (1980), Gilat and Helmers (1997), and Tsukahara (2013) can be applied.

Our following result works for general law-invariant convex risk measures:

Theorem 1. *Suppose that ρ is a law-invariant convex risk measure on H^Ψ and $X_1, X_2, \ldots$ is a stationary and ergodic sequence of random variables with the same law as $X \in H^\Psi$. Then $\widehat{\rho}_n$ is a strongly consistent estimator in the sense that*

$$\widehat{\rho}_n = \mathscr{R}_\rho(\widehat{m}_n) = \mathscr{R}_\rho\Big(\frac{1}{n}\sum_{k=1}^{n}\delta_{X_k}\Big) \longrightarrow \rho(X) \qquad \mathbb{P}\text{-}a.s.$$

It follows from Birkhoff's ergodic theorem that, $\mathbb{P}$-a.s.,

$$\widehat{m}_n \longrightarrow \mu := \mathbb{P} \circ X^{-1} \qquad \Psi(|\cdot|)\text{-weakly}$$

So Theorem 1 would have followed if it were possible to establish the continuity of $\nu \mapsto \mathscr{R}_\rho(\nu)$ with respect to the $\Psi(|\cdot|)$-weak topology. But this is not possible unless Ψ satisfies the Δ_2-condition:

4 Continuity

When Ψ is a Young function, then $\Psi(|\cdot|)$ is a weight function, and we will simply write $\mathscr{M}_1^\Psi$ in place of $\mathscr{M}_1^{\Psi(|\cdot|)}$. We will also use the term Ψ-weak convergence instead of $\Psi(|\cdot|)$-weak convergence etc. We recall the notation

$$\mathscr{M}(H^\Psi) = \left\{\mathbb{P} \circ X^{-1} \,|\, X \in H^\Psi\right\}$$

for the class of all laws of random variables $X \in H^\Psi$.

Remark 1. The identity $\mathscr{M}(H^\Psi) = \mathscr{M}_1^\Psi$ holds if and only if Ψ satisfies the Δ_2-condition (3).

Theorem 2. *For a finite Young function Ψ the following conditions are equivalent.*

(a) *For every law-invariant convex risk measure ρ on H^Ψ, the map $\mathscr{R}_\rho : \mathscr{M}(H^\Psi) \to \mathbb{R}$ is continuous for the Ψ-weak topology.*

(b) *Ψ satisfies the Δ_2-condition* (3).

Theorem 2. *For a finite Young function Ψ the following conditions are equivalent.*

(a) *For every law-invariant convex risk measure ρ on H^Ψ, the map $\mathscr{R}_\rho : \mathscr{M}(H^\Psi) \to \mathbb{R}$ is continuous for the Ψ-weak topology.*

(b) *Ψ satisfies the Δ_2-condition* (3).

Proof is based on the following Skorohod representation result for ψ-weak convergence:

Theorem 3. *For any finite Young function Ψ the following two conditions are equivalent.*

(a) *A sequence (μ_n) in $\mathscr{M}(H^\Psi)$ converges Ψ-weakly to some μ_0 if and only if there exists a sequence $(X_n)_{n\in\mathbb{N}_0}$ in H^Ψ such that X_n has law μ_n for each $n \in \mathbb{N}_0$ and $\|X_n - X_0\|_\Psi \to 0$.*

(b) *Ψ satisfies the Δ_2-condition* (3).

Idea of proof of Theorem 2:

(b)⇒(a): Suppose Ψ satisfies the Δ_2-condition and ρ is a convex risk measure on H^Ψ with associated map $\mathscr{R}_\rho$. Let (μ_n) be a sequence such that $\mu_n \to \mu_0$ Ψ-weakly. By Theorem 3 there exists a sequence $(X_n)_{n\in\mathbb{N}_0}$ in H^Ψ such that each X_n has law μ_n and such that $\|X_n - X_0\|_\Psi \to 0$. But ρ is continuous with respect to $\|\cdot\|_\Psi$ and so

$$\mathscr{R}_\rho(\mu_n) = \rho(X_n) \longrightarrow \rho(X_0) = \mathscr{R}_\rho(\mu_0),$$

which proves the implication (b)⇒(a).

(a)⇒(b): Use utility-based shortfall risk,

$$\rho(X) := \inf\{m \in \mathbb{R} : \mathbb{E}[\,\ell(-X-m)\,] \le x_0\},$$

for

$$\ell(x) = \Psi(8x^+)$$

together with the fact that there exists $Y \ge 0$ such that $\mathbb{E}[\,\Psi(Y)\,] < \infty$ and $\mathbb{E}[\,\Psi(2Y)\,] = \infty$ to construct X_n such that $\mathbb{P}\circ X_n^{-1} \to \delta_0$ but $\rho(X_n) \not\to \rho(0)$. $\square$

One might ask whether $\mathscr{R}_\rho$ is even continuous with respect to a weaker topology. For instance, this would be the case when ρ can be extended to a law-invariant convex risk measure on a larger Orlicz heart $H^\Phi \supset H^\Psi$.

To address this question, let ρ be a law-invariant convex risk measure on L^∞ and let

$$\overline{\rho} : L^1 \longrightarrow \mathbb{R} \cup \{+\infty\} \tag{4}$$

denote the unique extension of ρ that is convex, monotone, cash invariant, and lower semicontinuous with respect to the L^1-norm (Filipovic and Svindland, 2013). When $\overline{\rho}$ is finite on some Orlicz heart H^Ψ, it will be a convex risk measure and hence be continuous on H^Ψ with respect to the corresponding Luxemburg norm

Theorem 4. *Suppose that ρ is a law-invariant convex risk measure on L^∞. Let furthermore Ψ be a Young function satisfying the Δ_2-condition* (3)*. Then the following conditions are equivalent.*

(a) *$\overline{\rho}$ is finite on H^Ψ.*

(b) *The map $\mathscr{R}_{\overline{\rho}} : \mathscr{M}(H^\Psi) \to \mathbb{R}$ is continuous for the Ψ-weak topology.*

(c) *The map $\mathscr{R}_\rho : \mathscr{M}(L^\infty) \to \mathbb{R}$ is continuous for the Ψ-weak topology.*

(d) *If (X_n) is a sequence in L^∞ with $\|X_n\|_\Psi \to 0$, then $\rho(X_n) \to \rho(0)$.*

5 Differentiability, functional detal method, and central limit theorems

When ρ is coherent, one can establish a weak form of Hadamard differentiability of the map $\mathscr{R}_\rho$ (when defined on a suitable subspace of $\mathscr{M}_1^\psi$) and, under additional technical assumptions, obtain results such as the following one. Suppose that $\widehat{\mu}_n$ is a sequence of random measures such that $\widehat{\mu}_n \to \mu$ and

$$\frac{1}{\alpha_n}(\widehat{\mu}_n - \mu) \longrightarrow \beta \qquad \text{weakly}$$

for some random signed measure β and a sequence $\alpha_n \uparrow \infty$. Then

$$\frac{1}{\alpha_n}(\mathscr{R}_\rho(\widehat{\mu}_n) - \mathscr{R}_\rho(\mu)) \longrightarrow \nabla\mathscr{R}_\rho(\mu;\beta) \qquad \text{weakly,}$$

where $\nabla\mathscr{R}_\rho(\mu;\beta)$ is a Hadamard-type derivative of $\mathscr{R}_\rho$ at μ in direction β. See Krätschmer, A.S., and Zähle (2013) for details.

6 Qualitative and comparative robustness

$\Omega = \mathbb{R}^{\mathbb{N}}$, $X_i(\omega) = \omega(i)$ for $\omega \in \Omega$ and $i \in \mathbb{N}$, and $\mathcal{F} := \sigma(X_1, X_2, \dots)$. For any Borel probability measure μ on $\mathbb{R}$, we will denote

$$\mathbb{P}_\mu := \mu^{\otimes \mathbb{N}}$$

Definition 1 (Qualitative robustness)**.** Suppose $\mathscr{N} \subset \mathscr{M}_1$ is a set, d_A is a metric on $\mathscr{N}$, and d_B is a metric on $\mathscr{M}_1$. Then $\mathscr{R}_\rho$ is called robust on $\mathscr{N}$ with respect to d_A and d_B if for all $\mu \in \mathscr{N}$ and $\varepsilon > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that

$$(5) \qquad \nu \in \mathscr{N},\ d_A(\mu, \nu) \le \delta \quad \Longrightarrow \quad d_B\big(\mathbb{P}_\mu \circ \widehat{\rho}_n^{\,-1}, \mathbb{P}_\nu \circ \widehat{\rho}_n^{\,-1}\big) \le \varepsilon \quad \text{for } n \ge n_0.$$

In Hampel's classical notion of qualitative robustness: $\mathscr{N} = \mathscr{M}_1$ and d_A and d_B = Lévy metric or Prohorov metric (Mizera, 2010)

Here we take for d_A:

$$d_\psi(\mu,\nu) \;=\; d_{\mathrm{Proh}}(\mu,\nu) + \Big|\int \psi\,d\mu - \int \psi\,d\nu\Big|$$

and

$$d_B = d_{\mathrm{Proh}}$$

Definition 2. Let ψ be a weight function. A set $\mathscr{N} \subset \mathscr{M}_1^\psi$ is called uniformly ψ-integrating when

$$\lim_{M\to\infty} \sup_{\nu\in\mathscr{N}} \int_{\{\psi\geq M\}} \psi\,d\nu \;=\; 0. \tag{6}$$

When ψ is bounded, every set $\mathscr{N} \subset \mathscr{M}_1$ is uniformly ψ-integrating.

Definition 3. Let ψ be a weight function and $\mathscr{M} \subset \mathscr{M}_1^\psi$. A risk functional $\mathscr{R}_\rho$ is called ψ-robust on $\mathscr{M}$ when $\mathscr{R}_\rho$ is robust with respect to d_ψ and d_{Proh} on every uniformly ψ-integrating set $\mathscr{N} \subset \mathscr{M}$.

Proposition 1. *Let $\mathscr{R}_\rho$ be the risk functional associated with a law-invariant convex risk measure ρ on L^∞. When $\psi : \mathbb{R}_+ \to (0, \infty)$ is a nondecreasing function such that $\mathscr{R}_\rho$ is $\psi(|\cdot|)$-robust on $\mathscr{M}(L^\infty)$, then ψ has at least linear growth:*

$$\liminf_{x\uparrow\infty} \frac{\psi(x)}{x} > 0$$

The preceding proposition is a strengthening of the main result of Cont, Deguest, and Scandolo (2010). It also allows us to essentially limit the analysis of the ψ-robustness of risk functionals to weight functions $\psi(x) = \Psi(|x|)$ arising from a Young function Ψ. In this context, we have the following result.

Theorem 5. *For a finite Young function Ψ, the following conditions are equivalent.*

(a) *For every law-invariant convex risk measure ρ on H^Ψ, $\mathscr{R}_\rho$ is Ψ-robust on $\mathscr{M}(H^\Psi)$.*

(b) *Ψ satisfies the Δ_2-condition* (3).

As in Theorem 4, H^Ψ may not be the "canonical" space for ρ in the sense that ρ can be extended to a larger space. Such a situation has an impact on the robustness of ρ as explained in the next result. By $\overline{\rho}$ we denote again the extension (4).

Theorem 6. *Let Ψ be a Young function satisfying the Δ_2-condition* (3). *For a law-invariant convex risk measure ρ on L^∞, the following conditions are equivalent.*

(a) *$\mathscr{R}_{\overline{\rho}}$ is Ψ-robust on $\mathscr{M}_1^\Psi$.*

(b) *$\mathscr{R}_\rho$ is Ψ-robust on $\mathscr{M}(L^\infty)$.*

(c) *$\overline{\rho}$ is finite on H^Ψ.*

The most important aspect of Theorem 6 is that it allows us to study the robustness properties of a given risk functional on $\mathscr{M}(L^\infty)$ rather than on its full domain. Since any risk functional that arises from a law-invariant convex risk measure is defined on $\mathscr{M}(L^\infty)$, we can thus compare two risk functionals in regard to their degree of robustness.

Definition 4 (Comparative robustness)**.** We will say that ρ_1 is at least as robust as ρ_2 if the following implication holds. When Ψ is a Young function satisfying the Δ_2-condition (3), and $\mathscr{R}_{\rho_2}$ is Ψ-robust on $\mathscr{M}(L^\infty)$, then $\mathscr{R}_{\rho_1}$ is Ψ-robust on $\mathscr{M}(L^\infty)$. When, in addition, there is a Ψ such that $\mathscr{R}_{\rho_1}$ is Ψ-robust on $\mathscr{M}(L^\infty)$ but $\mathscr{R}_{\rho_2}$ is not, then we will say that ρ_1 is more robust than ρ_2.

We immediately get the following corollary.

Corollary 1. *For two law-invariant convex risk measures ρ_1 and ρ_2 on L^∞, the following conditions are equivalent.*

(a) *ρ_1 is at least as robust as ρ_2.*

(b) *When the Young function Ψ satisfies the Δ_2-condition* (3) *and $\overline{\rho}_2$ is finite on H^Ψ, then $\overline{\rho}_1$ is also finite on H^Ψ.*

Definition 5 (Index of qualitative robustness)**.** Let ρ be a law-invariant convex risk measure on L^∞. The associated index of qualitative robustness is defined as

$$\mathrm{iqr}(\rho) = \Big(\inf \big\{ p \in (0,\infty) \,\big|\, \mathscr{R}_\rho \text{ is } |\cdot|^p\text{-robust on } \mathscr{M}(L^\infty) \big\}\Big)^{-1}$$

It follows from Proposition 1 that any law-invariant convex risk measure ρ satisfies $\mathrm{iqr}(\rho) \leq 1$. Thus, Theorem 6 implies that

$$\mathrm{iqr}(\rho) = \Big(\inf \big\{ p \in [1,\infty) \,\big|\, \overline{\rho} \text{ is finite on } L^p \big\}\Big)^{-1}. \tag{7}$$

7 Index of qualitative robustness for distortion risk measures

We now turn to the important example class of *distortion risk measures* defined as

$$\rho_g(X) := \int_{-\infty}^{0} g(F_X(y))\,dy - \int_{0}^{\infty} \big(1 - g(F_X(y))\big)\,dy \tag{8}$$

where F_X denotes the distribution function of X, and g is a nondecreasing function such that $g(0) = 0$ and $g(1) = 1$. Then ρ_g is a law-invariant convex risk measure on L^∞ if and only if g is concave. In this case, ρ_g is even coherent and can be represented as

$$\rho_g(X) = g(0+)\operatorname{ess\,sup}(-X) + \int_0^1 V@R_t(X) g'_+(t)\,dt. \tag{9}$$

Proposition 2. *Suppose that g is concave and continuous. Then, for $p \in [1, \infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$, the extension $\overline{\rho}_g$ of ρ_g is finite on L^p if and only if $g'_+ \in L^q(0,1)$. In particular,*

$$\operatorname{iqr}(\rho_g) = \frac{q^* - 1}{q^*} \qquad \textit{where} \qquad q^* = \sup\left\{ q \geq 1 \,\middle|\, \int_0^1 (g'_+(t))^q \, dt < \infty \right\}.$$

Example 1 (Average Value at Risk)**.** Average Value at Risk at level $\alpha \in (0,1)$, $AV@R_\alpha$, is given in terms of the concave distortion function $g_1(t) = (t/\alpha) \wedge 1$. $AV@R_\alpha$ is also called Expected Shortfall, Conditional Value at Risk, or TailVaR. Since g_1' is bounded, it follows from Proposition 2 that $\operatorname{iqr}(AV@R_\alpha) = 1$.

Example 2 (MINMAXVAR)**.** MINMAXVAR is defined in terms of the concave distortion function

$$g_{\lambda,\gamma}(t) = 1 - (1 - t^{\frac{1}{1+\lambda}})^{1+\gamma},$$

where λ and γ are nonnegative parameters. An easy computation shows that $g'_{\lambda,\gamma}(t) \sim c \cdot t^{-\frac{\lambda}{1+\lambda}}$ as $t \downarrow 0$, and so we have $\operatorname{iqr}(\text{MINMAXVAR}) = \frac{1}{1+\lambda}$.

8 More on qualitative robustness

Definition 6. A set $\mathscr{N} \subset \mathscr{M}_1^\psi$ has the uniform Glivenko–Cantelli (UGC) property if $\widehat{m}_n$ satisfies a weak LLN w.r.t. d_ψ, uniformly over $\mathbb{P}_\nu$ with $\nu \in \mathscr{N}$. That is, for all $\varepsilon > 0$ and $\delta > 0$ there is $n_0 \in \mathbb{N}$ such that

$$\sup_{\nu\in\mathscr{N}} \mathbb{P}_\nu\big[\, d_\psi(\nu, \widehat{m}_n) \geq \delta \,\big] \leq \varepsilon \qquad \text{for } n \geq n_0. \tag{10}$$

Recall that $\mathscr{N} \subset \mathscr{M}_1^\psi$ is called *uniformly ψ-integrating* when

$$\lim_{M\uparrow\infty} \sup_{\nu\in\mathscr{N}} \int_{\{\psi\geq M\}} \psi\, d\nu = 0.$$

Theorem 7. *If $\mathscr{N}$ is uniformly ψ-integrating, then it has the UGC property. Conversely, if $\mathscr{N}$ has the UGC property and the medians of ψ under $\nu \in \mathscr{N}$ are bounded, then $\mathscr{N}$ is uniformly ψ-integrating.*

Proof: It is known that $\mathscr{M}_1$ has the UGC property for the Prohorov metric (Mizera, 2010). All one therefore needs to show is

$$\sup_{\nu\in\mathscr{N}} \mathbb{P}_\nu\Big[\Big|\frac{1}{n}\sum_{i=1}^{n}\psi(X_i) - \int \psi\, d\nu\Big| \geq \delta\Big] \leq \varepsilon \qquad \text{for } n \geq n_0. \tag{11}$$

Chung (1951) showed that such a LLN holds if $\mathscr{N}$ is uniformly ψ-integrating. He showed moreover that the validity of (11) requires that $\mathscr{N}$ is uniformly ψ-integrating in the case the medians of ψ under $\nu \in \mathscr{N}$ are bounded.

Let us give a new argument for

$$\mathscr{N} \text{ is uniformly } \psi\text{-integrating} \implies (11)$$

By Markov's inequality, (11) will follow once we have

$$\lim_{n\to\infty} \sup_{\nu\in\mathscr{N}} \mathbb{E}_\nu\Big[\Big|\frac{1}{n}\sum_{i=1}^{n}\psi(X_i) - \mathbb{E}_\nu\big[\psi(X_1)\big]\Big|\Big] = 0. \tag{12}$$

The fact that $\mathscr{N}$ is uniformly ψ-integrating implies compactness of

$$\mathscr{K} := \{\nu \circ \psi^{-1} \,|\, \nu \in \mathscr{N}\}$$

w.r.t. the $|\cdot|$-weak topology in $\mathscr{M}_1^{|\cdot|}([0,\infty))$.

Further, notice that (12) is implied by

$$\lim_{n\to\infty} F_n(\pi) = 0 \text{ uniformly in } \pi \in \mathscr{K} \tag{13}$$

where

$$F_n(\pi) := \mathbb{E}_\pi\Big[\Big|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_\pi[X_1]\Big|\Big]$$

In sequel, we will show that (13) holds. By Dini's lemma and the compactness of $\mathscr{K}$, for (13) to be true it suffices to show that

(a) $F_n \searrow 0$ pointwise and

(b) F_n is continuous for the $|\cdot|$-weak topology.

As for (a), $\frac{1}{n}\sum_{i=1}^{n} X_i$ is a backwards (reversed) $\mathbb{P}_\pi$-martingale and so

$$\Big|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_\pi[X_1]\Big|$$

is a backwards $\mathbb{P}_\pi$-submartingale converging to 0. This implies (a).

For (b), we need to show now that each F_n is continuous for the $|\cdot|$-weak topology. Let $(\pi_k)_{k\in\mathbb{N}_0}$ such that $\pi_n \to \pi_0$ $|\cdot|$-weakly; in particular

$$\int x\,\pi_k(dx) \to \int x\,\pi_0(dx). \tag{14}$$

We will show $F_n(\pi_k) \to F_n(\pi_0)$ as $k \to \infty$. To prove the latter convergence, we define functions $f_n : [0,\infty) \times [0,\infty)^n \to [0,\infty)$ by

$$f_n(t, x_1, \ldots, x_n) := \Big|\frac{1}{n}\sum_{i=1}^{n} x_i\ - t\Big|.$$

For each $t \in [0,\infty)$ fixed, the function $f_n(t,\cdot)$ can be bounded by $C(1+\widetilde{\psi})$ where $\widetilde{\psi}_n(x_1,\ldots,x_n) := \frac{1}{n}\sum_{i=1}^{n} x_i$. Hence, for each $t \in [0,\infty)$ the map

$$\pi \longmapsto \int f_n(t,\cdot)\, d\pi^{\otimes n}\ = \mathbb{E}_\pi\Big[\Big|\frac{1}{n}\sum_{i=1}^{n} X_i - t\Big|\Big]$$

is continuous for the $|\cdot|$-weak topology. Thus, for $t := \mathbb{E}_{\pi_0}\big[X_1\big]$ we obtain

$$\mathbb{E}_{\pi_k}\Big[\Big|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_{\pi_0}\big[X_1\big]\Big|\Big] \longrightarrow \mathbb{E}_{\pi_0}\Big[\Big|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_{\pi_0}\big[X_1\big]\Big|\Big]. \tag{15}$$

Now, combining the inequality

$$
\begin{aligned}
&\left|F_n(\pi_k) - F_n(\pi_0)\right| \\
&\quad \le \left|F_n(\pi_k) - \mathbb{E}_{\pi_k}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_{\pi_0}\left[X_1\right]\right|\right]\right| + \left|\mathbb{E}_{\pi_k}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_{\pi_0}\left[X_1\right]\right|\right] - F_n(\pi_0)\right| \\
&\quad \le \left|\int x\,\pi_k(dx) - \int x\,\pi_0(dx)\right| + \left|\mathbb{E}_{\pi_k}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_{\pi_0}\left[X_1\right]\right|\right] - F_n(\pi_0)\right|
\end{aligned}
$$

with (14) and (15) yields $F_n(\pi_k) \to F_n(\pi_0)$. $\square$

Consider robustness for a general estimator of the form

$$\widehat{T}_n = T(\widehat{m}_n)$$

for some statistical functional $T : \mathscr{M} \to \mathbb{R}$.

Definition 7 (ψ-robustness)**.** Let $\mathscr{M}$ be a subset of $\mathscr{M}_1^\psi$. T is called ψ-robust at μ in $\mathscr{M}$ if for each $\varepsilon > 0$ and every uniformly ψ-integrating set $\mathscr{N} \subset \mathscr{M}$ with $\mu \in \mathscr{N}$ there are $\delta > 0$ and $n_0 \in \mathbb{N}$ such that

$$\nu \in \mathscr{N},\ d_\psi(\mu, \nu) \leq \delta \quad \Longrightarrow \quad d_{\mathrm{Proh}}(\mathbb{P}_\mu \circ \widehat{T}_n^{-1}, \mathbb{P}_\nu \circ \widehat{T}_n^{-1}) \leq \varepsilon \quad \text{for } n \geq n_0.$$

Theorem 8 (Hampel's theorem for the ψ-weak topology)**.** *When $T : \mathscr{M} \to \mathbb{R}$ is ψ-weakly continuous at $\mu \in \mathscr{M}$, then T is ψ-robust at μ in $\mathscr{M}$.*

Proof: Must show that for every $\varepsilon > 0$ there are some $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $\nu \in \mathscr{N}$ and $n \geq n_0$,

$$d_\psi(\mu,\nu) \leq \delta \quad \Longrightarrow \quad d_{\text{Proh}}(\mathbb{P}_\mu \circ \hat{T}_n^{-1}, \mathbb{P}_\nu \circ \hat{T}_n^{-1}) \leq \varepsilon$$

So, let $\varepsilon > 0$ be fixed. We look for $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$d_\psi(\mu,\nu) \leq \delta \quad \Longrightarrow \quad d_{\text{Proh}}(\delta_{T(\mu)}, \mathbb{P}_\nu \circ \hat{T}_n^{-1}) \leq \frac{\varepsilon}{2},$$

where $\delta_{T(\mu)}$ is the Dirac measure at $T(\mu)$. Strassen's theorem implies

$$\mathbb{P}_\nu\Big[\big|T(\mu) - T(\widehat{m}_n)\big| \leq \frac{\varepsilon}{2}\Big] \geq 1 - \frac{\varepsilon}{2} \quad \Longrightarrow \quad d_{\text{Proh}}(\delta_{T(\mu)}, \mathbb{P}_\nu \circ \hat{T}_n^{-1}) \leq \frac{\varepsilon}{2}$$

Thus, it suffices to find some $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $\nu \in \mathscr{N}$ and $n \geq n_0$

$$d_\psi(\mu,\nu) \leq \delta \quad \Longrightarrow \quad \mathbb{P}_\nu\Big[\big|T(\mu) - T(\widehat{m}_n)\big| \leq \frac{\varepsilon}{2}\Big] \geq 1 - \frac{\varepsilon}{2}. \tag{16}$$

Since T is ψ-continuous at μ there is $\delta > 0$ such that $d_\psi(\mu, \widehat{m}_n) \leq 2\delta$ implies $|T(\mu) - T(\widehat{m}_n)| \leq \frac{\varepsilon}{2}$.

Thus, in order to obtain (16), let us fix any $\nu \in \mathscr{N}$ satisfying $d_\psi(\mu,\nu) \leq \delta$. In view of the triangular inequality $d_\psi(\mu,\widehat{m}_n) \leq d_\psi(\nu,\widehat{m}_n) + d_\psi(\mu,\nu)$ we have

$$\begin{aligned}
&\mathbb{P}_\nu\Big[d_\psi(\nu,\widehat{m}_n) \leq \delta \Big] \\
&\quad \leq \quad \mathbb{P}_\nu\Big[d_\psi(\mu,\widehat{m}_n) \leq \delta + d_\psi(\mu,\nu) \Big] \\
&\quad \leq \quad \mathbb{P}_\nu\Big[d_\psi(\mu,\widehat{m}_n) \leq 2\delta \Big] \\
&\quad \leq \quad \mathbb{P}_\nu\Big[|T(\mu) - T(\widehat{m}_n)| \leq \frac{\varepsilon}{2} \Big].
\end{aligned}$$

Now, (16) is an immediate consequence of the UGC property of $(\mathscr{N}, d)$. $\square$

Thus, in order to obtain (16), let us fix any $\nu \in \mathscr{N}$ satisfying $d_\psi(\mu, \nu) \le \delta$. In view of the triangular inequality $d_\psi(\mu, \widehat{m}_n) \le d_\psi(\nu, \widehat{m}_n) + d_\psi(\mu, \nu)$ we have

$$\begin{aligned}
&\mathbb{P}_\nu\Big[d_\psi(\nu, \widehat{m}_n) \le \delta\Big] \\
&\quad\le \quad \mathbb{P}_\nu\Big[d_\psi(\mu, \widehat{m}_n) \le \delta + d_\psi(\mu, \nu)\Big] \\
&\quad\le \quad \mathbb{P}_\nu\Big[d_\psi(\mu, \widehat{m}_n) \le 2\delta\Big] \\
&\quad\le \quad \mathbb{P}_\nu\Big[|T(\mu) - T(\widehat{m}_n)| \le \frac{\varepsilon}{2}\Big].
\end{aligned}$$

Now, (16) is an immediate consequence of the UGC property of $(\mathscr{N}, d)$. $\square$

Theorem 9 (Converse of Hampel's theorem for the ψ-weak topology)**.** *Suppose that $T : \mathscr{M} \to \mathbb{R}$ is a statistical functional. Let $\mu \in \mathscr{M}$ and $\delta_0 > 0$ be given, and suppose that T is weakly consistent at each ν in $\mathscr{M}$ with $d_\psi(\nu, \mu) \le \delta_0$. When T is ψ-robust at μ in $\mathscr{M}$, then $T : \mathscr{M} \to \mathbb{R}$ is ψ-weakly continuous at μ.*

9 Conclusion

We have introduced a refined version of qualitative robustness of statistical functionals. For the risk functional associated with a law-invariant convex risk measure, robustness can be formulated within the context of Young functions Ψ and it holds if and only if the risk measure is finite on the corresponding Orlicz space—provided that Ψ satisfies the Δ_2-condition. There is still work to be done for the case in which Ψ does not satisfy the Δ_2-condition. Also, how can one better describe the tradeoff between robustness and tail sensitivity?

Thank you

References

Cont, R., Deguest, R. & Scandolo, G. (2010), 'Robustness and sensitivity analysis of risk measurement procedures', *Quant. Finance* **10**(6), 593–606.
URL: *http://dx.doi.org/10.1080/14697681003685597*

Filipović, D. & Svindland, G. (2012), 'The canonical model space for law-invariant convex risk measures is L^1', *Math. Finance* **22**(3), 585–589.
URL: *http://dx.doi.org/10.1111/j.1467-9965.2012.00534.x*

Hampel, F. R. (1971), 'A general qualitative definition of robustness', *Ann. Math. Statist.* **42**, 1887–1896.

Krätschmer, V., Schied, A. & Zähle, H. (2012), 'Qualitative and infinitesimal robustness of tail-dependent statistical functionals', *J. Multivariate Anal.* **103**, 35–47.
URL: *http://dx.doi.org/10.1016/j.jmva.2011.06.005*

Krätschmer, V., Schied, A. & Zähle, H. (2013), 'Quasi-Hadamard differentiability of general risk functionals and its application', *arxiv preprint 401.3167* .

Krätschmer, V., Schied, A. & Zähle, H. (2014), 'Comparative and qualitative robustness for law-invariant risk measures', *to appear in Finance and Stochastics* .

Mizera, I. (2010), Qualitative robustness and weak continuity: the extreme unction?, *in* 'Nonparametrics and robustness in modern statistical inference and time series analysis: a Festschrift in honor of Professor Jana Jurečková', Vol. 7 of *Inst. Math. Stat. Collect.*, Inst. Math. Statist., Beachwood, OH, pp. 169–181.

Weber, S. (2006), 'Distribution-invariant risk measures, information, and dynamic consistency', *Math. Finance* **16**(2), 419–441.
URL: *http://dx.doi.org/10.1111/j.1467-9965.2006.00277.x*