Elicitability and Identifiability of Measures of Systemic Risk

Tobias Fissler

Imperial College London

based on joint work with
Jana Hlavinová and Birgit Rudloff

IMS–FIPS London

10 – 11 September 2018

Briefing: Risk Measures

Let the random variable Y model the gains and losses of a financial position.
A risk measure ρ maps Y to the real value $\rho(Y) \in \mathbb{R}$ which stands for the money one has to add to Y in order to make it acceptable. That is

$$\rho(Y + \rho(Y)) = 0.$$

Briefing: Risk Measures

Let the random variable Y model the gains and losses of a financial position.
A risk measure ρ maps Y to the real value $\rho(Y) \in \mathbb{R}$ which stands for the money one has to add to Y in order to make it acceptable. That is

$$\rho(Y + \rho(Y)) = 0.$$

Properties

Let Y, X be random variables.

Cash-invariance For any $m \in \mathbb{R}$ $\rho(X + m) = \rho(X) - m$. $\rightsquigarrow \rho(0) = 0$.

Briefing: Risk Measures

Let the random variable Y model the gains and losses of a financial position.
A risk measure ρ maps Y to the real value $\rho(Y) \in \mathbb{R}$ which stands for the money one has to add to Y in order to make it acceptable. That is

$$\rho(Y + \rho(Y)) = 0.$$

Properties

Let Y, X be random variables.

Cash-invariance For any $m \in \mathbb{R}$ $\rho(X + m) = \rho(X) - m$. $\rightsquigarrow \rho(0) = 0$.

Homogeneity For any $c > 0$ $\rho(cX) = c\rho(X)$.

Briefing: Risk Measures

Let the random variable Y model the gains and losses of a financial position.
A risk measure ρ maps Y to the real value $\rho(Y) \in \mathbb{R}$ which stands for the money one has to add to Y in order to make it acceptable. That is

$$\rho(Y + \rho(Y)) = 0.$$

Properties

Let Y, X be random variables.

Cash-invariance For any $m \in \mathbb{R}$ $\rho(X + m) = \rho(X) - m$. $\rightsquigarrow \rho(0) = 0$.

Homogeneity For any $c > 0$ $\rho(cX) = c\rho(X)$.

Monotonicity If $X \leqslant Y$ a.s. then $\rho(X) \geqslant \rho(Y)$.

Briefing: Risk Measures

Let the random variable Y model the gains and losses of a financial position.
A risk measure ρ maps Y to the real value $\rho(Y) \in \mathbb{R}$ which stands for the money one has to add to Y in order to make it acceptable. That is

$$\rho(Y + \rho(Y)) = 0.$$

Properties

Let Y, X be random variables.

Cash-invariance For any $m \in \mathbb{R}$ $\rho(X + m) = \rho(X) - m$. $\rightsquigarrow \rho(0) = 0$.

Homogeneity For any $c > 0$ $\rho(cX) = c\rho(X)$.

Monotonicity If $X \leqslant Y$ a.s. then $\rho(X) \geqslant \rho(Y)$.

(Sub-additivity) $\rho(X + Y) \leqslant \rho(X) + \rho(Y)$.

Briefing: Risk Measures

Let the random variable Y model the gains and losses of a financial position.
A risk measure ρ maps Y to the real value $\rho(Y) \in \mathbb{R}$ which stands for the money one has to add to Y in order to make it acceptable. That is

$$\rho(Y + \rho(Y)) = 0.$$

Properties

Let Y, X be random variables.

Cash-invariance For any $m \in \mathbb{R}$ $\rho(X + m) = \rho(X) - m$. $\rightsquigarrow \rho(0) = 0$.

Homogeneity For any $c > 0$ $\rho(cX) = c\rho(X)$.

Monotonicity If $X \leqslant Y$ a.s. then $\rho(X) \geqslant \rho(Y)$.

(Sub-additivity) $\rho(X + Y) \leqslant \rho(X) + \rho(Y)$.

(Law-invariance) If $X \stackrel{d}{=} Y$ then $\rho(X) = \rho(Y)$.

Briefing: Risk Measures (Examples)

Value-at-Risk

Let $Y \sim F$ and $\alpha \in (0,1)$ (close to 0). Then

$$\mathrm{VaR}_\alpha(Y) = -q_\alpha^-(F) = -\inf\{x \in \mathbb{R} \mid F(x) \geqslant \alpha\}.$$

Briefing: Risk Measures (Examples)

Value-at-Risk

Let $Y \sim F$ and $\alpha \in (0,1)$ (close to 0). Then

$$\mathrm{VaR}_\alpha(Y) = -q_\alpha^-(F) = -\inf\{x \in \mathbb{R} \mid F(x) \geqslant \alpha\}.$$

Expected Shortfall

Let $Y \sim F$ and $\alpha \in (0,1)$ (close to 0). Then (if $F(q_\alpha^-(F)) = \alpha$)

$$\mathrm{ES}_\alpha(Y) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_\beta(Y)\mathrm{d}\beta \;\big(= -\mathbf{E}_F[Y \mid Y \leqslant q_\alpha^-(F)]\big).$$

Measures of Systemic Risk

- Suppose you have some financial system consisting of n firms. The system can be represented as a random vector $Y = (Y_1, \ldots, Y_n)$. How to measure the risk of the entire system Y?

Measures of Systemic Risk

- Suppose you have some financial system consisting of n firms. The system can be represented as a random vector $Y = (Y_1, \dots, Y_n)$. How to measure the risk of the entire system Y?
- Apply some scalar risk measure to each component:

$$\rho(Y) = (\rho(Y_1), \dots, \rho(Y_n))$$

Caveat: Ignores the dependence structure! (Usually high correlation in the tails!)

Measures of Systemic Risk

- Suppose you have some financial system consisting of n firms. The system can be represented as a random vector $Y = (Y_1, \dots, Y_n)$.
  How to measure the risk of the entire system Y?
- Apply some scalar risk measure to each component:

  $$\rho(Y) = (\rho(Y_1), \dots, \rho(Y_n))$$

  Caveat: Ignores the dependence structure! (Usually high correlation in the tails!)
- Use some kind of generalisation of quantiles to replace VaR (this will be set-valued).

Measures of Systemic Risk

- Suppose you have some financial system consisting of n firms. The system can be represented as a random vector $Y = (Y_1, \dots, Y_n)$. How to measure the risk of the entire system Y?
- Apply some scalar risk measure to each component:

$$\rho(Y) = (\rho(Y_1), \dots, \rho(Y_n))$$

  Caveat: Ignores the dependence structure! (Usually high correlation in the tails!)
- Use some kind of generalisation of quantiles to replace VaR (this will be set-valued).
- Aggregate the system with some monotone aggregation function $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$. Measure the risk via

$$\rho(\Lambda(Y)).$$

  $\rightsquigarrow$ Bail-out costs. This is insensitive with respect to capital allocations and thus ignores transaction costs.

Measures of Systemic Risk

Feinstein, Rudloff, Weber (2017)

Take an ex ante point of view: How do we need to allocate additional money $k \in \mathbb{R}^n$ in order to make the aggregate system $\Lambda(Y + k)$ acceptable under ρ?

$$R(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) \leqslant 0\}.$$

Measures of Systemic Risk

Feinstein, Rudloff, Weber (2017)

Take an ex ante point of view: How do we need to allocate additional money $k \in \mathbb{R}^n$ in order to make the aggregate system $\Lambda(Y+k)$ acceptable under ρ?

$$R(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y+k)) \leqslant 0\}.$$

Example 1

Examples for the aggregation $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$

$$\Lambda(x) = \sum_{i=1}^n x_i, \qquad \Lambda(x) = \sum_{i=1}^n -x_i^-,$$

$$\Lambda(x) = \sum_{i=1}^n [\alpha_i (x_i - v_i)^+ - \beta_i (x_i - v_i)^-], \qquad \Lambda(x) = \sum_{i=1}^n [1 - \exp(2x_i^-)].$$

Measures of Systemic Risk

Properties I of $R(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) \leqslant 0\}$

- The values of R are subsets of $\mathbb{R}^n$.

Measures of Systemic Risk

Properties I of $R(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) \leqslant 0\}$

- The values of R are subsets of $\mathbb{R}^n$.
- Due to the monotonicity of Λ, they are upper sets. So

$$R(Y) = R(Y) + \mathbb{R}^n_+.$$

Measures of Systemic Risk

Properties I of $R(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) \leqslant 0\}$

- The values of R are subsets of $\mathbb{R}^n$.
- Due to the monotonicity of Λ, they are upper sets. So

$$R(Y) = R(Y) + \mathbb{R}^n_+.$$

- If Λ is continuous, then $R(Y)$ is closed.

Measures of Systemic Risk

Properties I of $R(Y) = \{k \in \mathbb{R}^n \,|\, \rho(\Lambda(Y+k)) \leqslant 0\}$

- The values of R are subsets of $\mathbb{R}^n$.
- Due to the monotonicity of Λ, they are upper sets. So

$$R(Y) = R(Y) + \mathbb{R}^n_+.$$

- If Λ is continuous, then $R(Y)$ is closed.
- If Λ is concave and ρ convex, then the set $R(Y)$ is convex.

Measures of Systemic Risk – illlustration

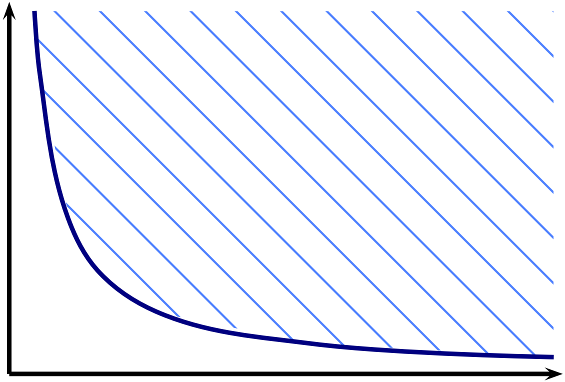


Figure: Illustration of a systemic risk measure $R(Y) = \{k \in \mathbb{R}^n \,|\, \rho(\Lambda(Y + k)) \leqslant 0\}$.

Measures of Systemic Risk – Properties II

Properties II

Let Y, X be random vectors.

Cash-invariance For any $m \in \mathbb{R}^n$: $R(Y + m) = R(Y) - m$.

Homogeneity If Λ is homogeneous, then R is homogeneous:

$$R(cY) = c\,R(Y), \quad \forall c > 0.$$

Monotonicity If $X \leqslant Y$ a.s. then $R(X) \subseteq R(Y)$.

(Law-invariance) If ρ is law-invariant, then R is law-invariant. That is, if $X \overset{d}{=} Y$ then $R(X) = R(Y)$.

Statistical Properties

Possible tasks:

(i) M-estimation of $R(Y)$, using realisations $Y_1, \ldots, Y_N$.

(ii) Fit a parametric model for $R(Y)$ with regression.

(iii) Compare and rank competing forecasts for R.

(iv) Validate forecasts / estimates for R.

(v) Z-estimation or GMM.

Statistical Properties

Possible tasks:

(i) M-estimation of $R(Y)$, using realisations $Y_1, \ldots, Y_N$.

(ii) Fit a parametric model for $R(Y)$ with regression.

(iii) Compare and rank competing forecasts for R.

(iv) Validate forecasts / estimates for R.

(v) Z-estimation or GMM.

For (i) – (iii) we need loss functions of the form

$$L\colon 2^{\mathbb{R}^n} \times \mathbb{R}^n \to \mathbb{R}.$$

They should incentivise truthful and honest forecasts.
In regression, we'd like to have a consistent estimator for the true parameter.

Statistical Properties

Possible tasks:

(i) M-estimation of $R(Y)$, using realisations $Y_1, \ldots, Y_N$.
(ii) Fit a parametric model for $R(Y)$ with regression.
(iii) Compare and rank competing forecasts for R.
(iv) Validate forecasts / estimates for R.
(v) Z-estimation or GMM.

For (i) – (iii) we need loss functions of the form

$$L\colon 2^{\mathbb{R}^n} \times \mathbb{R}^n \to \mathbb{R}.$$

They should incentivise truthful and honest forecasts.
In regression, we'd like to have a consistent estimator for the true parameter.
$\rightsquigarrow$ This calls for the notion of elicitability!

(iv) and (v) need identifiability.

Strictly consistent loss functions

- Suppose we have some observations $Y_1, \ldots, Y_N$ of a random quantity with values in some observation domain O.

Strictly consistent loss functions

- Suppose we have some observations $Y_1, \ldots, Y_N$ of a random quantity with values in some observation domain O.
- We want to compare competing forecasts

$$x_1^{(1)}, \ldots, x_N^{(1)} \in \mathsf{A}, \qquad x_1^{(2)}, \ldots, x_N^{(2)} \in \mathsf{A},$$

for some functional T (mean, quantile, risk measure, probability) of the (conditional) distributions of $Y_1, \ldots, Y_N$, taking values in some action domain A.

Strictly consistent loss functions

- Suppose we have some observations $Y_1, \ldots, Y_N$ of a random quantity with values in some observation domain O.
- We want to compare competing forecasts

$$x_1^{(1)}, \ldots, x_N^{(1)} \in \mathsf{A}, \qquad x_1^{(2)}, \ldots, x_N^{(2)} \in \mathsf{A},$$

  for some functional T (mean, quantile, risk measure, probability) of the (conditional) distributions of $Y_1, \ldots, Y_N$, taking values in some action domain A.
- Using a loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ we compare and rank the competing forecasts in terms of their realized losses:

$$\mathbf{L}_N^{(1)} = \frac{1}{N}\sum_{t=1}^{N} L\big(x_t^{(1)}, Y_t\big) \overset{?}{\lesseqgtr} \mathbf{L}_N^{(2)} = \frac{1}{N}\sum_{t=1}^{N} L\big(x_t^{(2)}, Y_t\big)$$

Strictly consistent loss functions

- Suppose we have some observations $Y_1, \dots, Y_N$ of a random quantity with values in some observation domain O.
- We want to compare competing forecasts

$$x_1^{(1)}, \dots, x_N^{(1)} \in \mathsf{A}, \qquad x_1^{(2)}, \dots, x_N^{(2)} \in \mathsf{A},$$

  for some functional T (mean, quantile, risk measure, probability) of the (conditional) distributions of $Y_1, \dots, Y_N$, taking values in some action domain A.
- Using a loss function $L \colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ we compare and rank the competing forecasts in terms of their realized losses:

$$\mathbf{L}_N^{(1)} = \frac{1}{N} \sum_{t=1}^{N} L\big(x_t^{(1)}, Y_t\big) \overset{?}{\lessgtr} \mathbf{L}_N^{(2)} = \frac{1}{N} \sum_{t=1}^{N} L\big(x_t^{(2)}, Y_t\big)$$

- Ranking depends on the choice of the loss function!

Strictly consistent loss functions

- Suppose we have some observations $Y_1, \dots, Y_N$ of a random quantity with values in some observation domain O.
- We want to compare competing forecasts

$$x_1^{(1)}, \dots, x_N^{(1)} \in \mathsf{A}, \qquad x_1^{(2)}, \dots, x_N^{(2)} \in \mathsf{A},$$

  for some functional T (mean, quantile, risk measure, probability) of the (conditional) distributions of $Y_1, \dots, Y_N$, taking values in some action domain A.
- Using a loss function $L \colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ we compare and rank the competing forecasts in terms of their realized losses:

$$\mathbf{L}_N^{(1)} = \frac{1}{N} \sum_{t=1}^{N} L\big(x_t^{(1)}, Y_t\big) \overset{?}{\lesseqgtr} \mathbf{L}_N^{(2)} = \frac{1}{N} \sum_{t=1}^{N} L\big(x_t^{(2)}, Y_t\big)$$

- Ranking depends on the choice of the loss function!
- The loss function should incentivise truthful and honest forecasts!

Strictly consistent loss functions

Definition 2 (Consistency)

A loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ is strictly $\mathcal{F}$-consistent for some functional $T\colon \mathcal{F} \to \mathsf{A}$ if

$$\mathbf{E}_F[L(T(F), Y)] < \mathbf{E}_F[L(x, Y)]$$

for any $F \in \mathcal{F}$ and any $x \in \mathsf{A}$, $x \neq T(F)$.

Strictly consistent loss functions

Definition 2 (Consistency)

A loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ is strictly $\mathcal{F}$-consistent for some functional $T\colon \mathcal{F} \to \mathsf{A}$ if

$$\mathbf{E}_F[L(T(F), Y)] < \mathbf{E}_F[L(x, Y)]$$

for any $F \in \mathcal{F}$ and any $x \in \mathsf{A}$, $x \neq T(F)$.

Definition 3 (Elicitability)

A functional $T\colon \mathcal{F} \to \mathsf{A}$ is elicitable if there is a strictly $\mathcal{F}$-consistent loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ for T. Then

$$T(F) = \arg\min_{x \in \mathsf{A}} \mathbf{E}_F[L(x, Y)].$$

Strictly consistent loss functions

Definition 2 (Consistency)

A loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ is strictly $\mathcal{F}$-consistent for some functional $T\colon \mathcal{F} \to \mathsf{A}$ if

$$\mathbf{E}_F[L(T(F), Y)] < \mathbf{E}_F[L(x, Y)]$$

for any $F \in \mathcal{F}$ and any $x \in \mathsf{A}$, $x \neq T(F)$.

Definition 3 (Elicitability)

A functional $T\colon \mathcal{F} \to \mathsf{A}$ is elicitable if there is a strictly $\mathcal{F}$-consistent loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ for T. Then

$$T(F) = \arg\min_{x \in \mathsf{A}} \mathbf{E}_F[L(x, Y)].$$

Applications:

- M-estimation
- Regression
- (Meaningful) forecast comparison; forecast ranking; model selection.

Regression

Classic situation: There is some parametric model $m\colon \Theta \times \mathbb{R} \to \mathbb{R}$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$Y = m_{\theta*}(X) + \varepsilon, \qquad \text{where} \quad \mathbf{E}[\varepsilon|X] = 0. \tag{1}$$

Regression

Classic situation: There is some parametric model $m\colon \Theta \times \mathbb{R} \to \mathbb{R}$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$Y = m_{\theta*}(X) + \varepsilon, \qquad \text{where} \quad \mathbf{E}[\varepsilon|X] = 0. \tag{1}$$

Equivalent form of (1):

$$\mathbf{E}[Y|X] = m_{\theta*}(X).$$

Regression

Classic situation: There is some parametric model $m\colon \Theta \times \mathbb{R} \to \mathbb{R}$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$Y = m_{\theta^*}(X) + \varepsilon, \qquad \text{where} \quad \mathbf{E}[\varepsilon|X] = 0. \tag{1}$$

Equivalent form of (1):

$$\mathbf{E}[Y|X] = m_{\theta^*}(X).$$

Find an estimator $\hat{\theta}_n$ for θ^* by

$$\hat{\theta}_n = \arg\min_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} (m_\theta(X_i) - Y_i)^2.$$

Regression

Classic situation: There is some parametric model $m\colon \Theta \times \mathbb{R} \to \mathbb{R}$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$Y = m_{\theta*}(X) + \varepsilon, \qquad \text{where} \quad \mathbf{E}[\varepsilon|X] = 0. \tag{1}$$

Equivalent form of (1):

$$\mathbf{E}[Y|X] = m_{\theta*}(X).$$

Find an estimator $\hat{\theta}_n$ for θ^* by

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n}(m_\theta(X_i) - Y_i)^2.$$

Relying in the fact that

$$\theta^* \in \arg\min_{\theta\in\Theta} \mathbf{E}(m_\theta(X) - Y)^2 \qquad \left\{\theta^* \in \arg\min_{\theta\in\Theta} \mathbf{E}\left[(m_\theta(X) - Y)^2|X\right]\right\}$$

Regression

Classic situation: There is some parametric model $m\colon \Theta \times \mathbb{R} \to \mathbb{R}$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$Y = m_{\theta*}(X) + \varepsilon, \qquad \text{where} \quad \mathbf{E}[\varepsilon|X] = 0. \tag{1}$$

Equivalent form of (1):

$$\mathbf{E}[Y|X] = m_{\theta*}(X).$$

Find an estimator $\hat{\theta}_n$ for θ^* by

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n}(m_\theta(X_i) - Y_i)^2.$$

Relying in the fact that

$$\theta^* \in \arg\min_{\theta\in\Theta} \mathbf{E}(m_\theta(X) - Y)^2 \qquad \left\{\theta^* \in \arg\min_{\theta\in\Theta} \mathbf{E}\left[(m_\theta(X) - Y)^2|X\right]\right\}$$

However, instead of squared loss, we could use any strictly consistent loss function for the mean functional.

Regression II

General situation: There is some parametric model $m\colon \Theta \times \mathbb{R}^\ell \to \mathbb{R}^k$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$T(\mathcal{L}(Y|\,X)) = m_{\theta*}(X).$$

Regression II

General situation: There is some parametric model $m\colon \Theta \times \mathbb{R}^\ell \to \mathbb{R}^k$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$T(\mathcal{L}(Y|\,X)) = m_{\theta*}(X).$$

Assume that $L\colon \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}$ is a strictly consistent loss function for T.

Regression II

General situation: There is some parametric model $m\colon \Theta \times \mathbb{R}^\ell \to \mathbb{R}^k$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$T(\mathcal{L}(Y|\,X)) = m_{\theta*}(X).$$

Assume that $L\colon \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}$ is a strictly consistent loss function for T.

Find an estimator $\hat{\theta}_n$ for θ^* by

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n} L(m_\theta(X_i), Y_i).$$

Regression II

General situation: There is some parametric model $m \colon \Theta \times \mathbb{R}^\ell \to \mathbb{R}^k$ and we assume that there is some true parameter $\theta^* \in \Theta$ such that

$$T(\mathcal{L}(Y|\, X)) = m_{\theta*}(X).$$

Assume that $L \colon \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}$ is a strictly consistent loss function for T.

Find an estimator $\hat{\theta}_n$ for θ^* by

$$\hat{\theta}_n = \operatorname*{arg\,min}_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} L(m_\theta(X_i), Y_i).$$

Relying in the fact that

$$\theta^* \in \operatorname*{arg\,min}_{\theta \in \Theta} \mathbf{E}\big[L(m_\theta(X), Y)\big] \qquad \Big\{\theta^* \in \operatorname*{arg\,min}_{\theta \in \Theta} \mathbf{E}\big[L(m_\theta(X), Y)|X\big]\Big\}$$

Examples

T	$L(x, y)$
mean	$(x-y)^2$
median	$\lvert x-y\rvert$
α-quantile	$(\mathbb{1}\{y \leqslant x\} - \alpha)(x-y)$
τ-expectile	$\lvert\mathbb{1}\{y \leqslant x\} - \tau\rvert(x-y)^2$

Examples

T	$L(x, y)$
mean	$(x-y)^2$
median	$\lvert x-y\rvert$
α-quantile	$(\mathbb{1}\{y \leqslant x\} - \alpha)(x-y)$
τ-expectile	$\lvert \mathbb{1}\{y \leqslant x\} - \tau\rvert(x-y)^2$
variance	$\times$
Expected Shortfall	$\times$

Examples

T	$L(x, y)$
mean	$(x - y)^2$
median	$\|x - y\|$
α-quantile	$(\mathbb{1}\{y \leqslant x\} - \alpha)(x - y)$
τ-expectile	$\|\mathbb{1}\{y \leqslant x\} - \tau\|(x - y)^2$
variance	×
Expected Shortfall	×
(mean, variance)	✓
(quantile, Expected Shortfall)	✓

Examples

T	$L(x, y)$
mean	$(x-y)^2$
median	$\lvert x-y \rvert$
α-quantile	$(\mathbb{1}\{y \leqslant x\} - \alpha)(x-y)$
τ-expectile	$\lvert \mathbb{1}\{y \leqslant x\} - \tau \rvert (x-y)^2$
variance	×
Expected Shortfall	×
(mean, variance)	✓
(quantile, Expected Shortfall)	✓
identity (probabilistic forecast)	$L(F, y) = -\log(f(y))$

Examples

T	$L(x, y)$
mean	$(x-y)^2$
median	$\lvert x-y\rvert$
α-quantile	$(\mathbb{1}\{y \leqslant x\} - \alpha)(x-y)$
τ-expectile	$\lvert\mathbb{1}\{y \leqslant x\} - \tau\rvert(x-y)^2$
variance	×
Expected Shortfall	×
(mean, variance)	✓
(quantile, Expected Shortfall)	✓
identity (probabilistic forecast)	$L(F, y) = -\log(f(y))$
	$L(F, y) = \int \big(F(x) - \mathbb{1}\{y \leqslant x\}\big)^2 \mathrm{d}x$

Two modes of elicitability

- Many functionals such as moments, variance, or continuous densities are unique, taking a single point in the action domain A.

Two modes of elicitability

- Many functionals such as moments, variance, or continuous densities are unique, taking a single point in the action domain A.
- Other functionals, such as quantiles, are naturally set-valued

$$q_\alpha(F) = \{x \in \mathbb{R} \mid \lim_{t \uparrow x} F(t) \leqslant \alpha \leqslant F(x)\} \subset \mathbb{R}.$$

Two modes of elicitability

- Many functionals such as moments, variance, or continuous densities are unique, taking a single point in the action domain A.
- Other functionals, such as quantiles, are naturally set-valued

$$q_\alpha(F) = \{x \in \mathbb{R} \mid \lim_{t \uparrow x} F(t) \leqslant \alpha \leqslant F(x)\} \subset \mathbb{R}.$$

- Choice of the action domain A:

Two modes of elicitability

- Many functionals such as moments, variance, or continuous densities are unique, taking a single point in the action domain A.
- Other functionals, such as quantiles, are naturally set-valued

$$q_\alpha(F) = \{x \in \mathbb{R} \mid \lim_{t \uparrow x} F(t) \leqslant \alpha \leqslant F(x)\} \subset \mathbb{R}.$$

- Choice of the action domain A:

  $\mathsf{A} = \mathbb{R}$: The forecasts are points in $\mathbb{R}$. There are multiple best actions, namely every $x \in q_\alpha(F)$.

  $\rightsquigarrow$ The functional T is set-valued, that is

$$T\colon \mathcal{F} \to 2^{\mathsf{A}}.$$

Two modes of elicitability

- Many functionals such as moments, variance, or continuous densities are unique, taking a single point in the action domain A.
- Other functionals, such as quantiles, are naturally set-valued

$$q_\alpha(F) = \{x \in \mathbb{R} \mid \lim_{t \uparrow x} F(t) \leqslant \alpha \leqslant F(x)\} \subset \mathbb{R}.$$

- Choice of the action domain A:

$\mathsf{A} = \mathbb{R}$: The forecasts are points in $\mathbb{R}$. There are multiple best actions, namely every $x \in q_\alpha(F)$.
$\rightsquigarrow$ The functional T is set-valued, that is

$$T\colon \mathcal{F} \to 2^{\mathsf{A}}.$$

$\mathsf{A} \subseteq 2^{\mathbb{R}}$: The forecasts are subsets of $\mathbb{R}$. These are points in the power set $\mathsf{A} \subseteq 2^{\mathbb{R}}$. There is a unique best action namely $x = q_\alpha(F)$.
$\rightsquigarrow$ The functional T is point-valued in some space $\mathsf{A} \subseteq 2^{\mathbb{R}}$, that is,

$$T\colon \mathcal{F} \to \mathsf{A}.$$

Two modes of elicitability

To unify the framework, we can consider all functionals as set-valued, possibly identifying them with singletons. E.g., we consider the mean functional as

$$F \mapsto T(F) = \left\{ \int x \, \mathrm{d}F(x) \right\} \in 2^{\mathbb{R}}.$$

Definition 4

(a) A functional $T \colon \mathcal{F} \to 2^{\mathsf{A}}$ is selectively elicitable if there is a loss function $L \colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ such that

$$\mathbf{E}_F[L(t, Y)] < \mathbf{E}_F[L(x, Y)]$$

for all $F \in \mathcal{F}$ and for all $t \in T(F)$ and for all $x \in \mathsf{A} \setminus T(F)$.

Two modes of elicitability

To unify the framework, we can consider all functionals as set-valued, possibly identifying them with singletons. E.g., we consider the mean functional as

$$F \mapsto T(F) = \left\{ \int x \, \mathrm{d}F(x) \right\} \in 2^{\mathbb{R}}.$$

Definition 4

(a) A functional $T\colon \mathcal{F} \to 2^{\mathsf{A}}$ is **selectively elicitable** if there is a loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ such that

$$\mathbf{E}_F[L(t, Y)] < \mathbf{E}_F[L(x, Y)]$$

for all $F \in \mathcal{F}$ and for all $t \in T(F)$ and for all $x \in \mathsf{A} \setminus T(F)$.

(b) A functional $T\colon \mathcal{F} \to \mathsf{A}$ is **exhaustively elicitable** if there is a loss function $L\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$ such that

$$\mathbf{E}_F[L(T(F), Y)] < \mathbf{E}_F[L(x, Y)]$$

for all $F \in \mathcal{F}$ and for all $x \in \mathsf{A}$, $x \neq T(F)$.

Two modes of elicitability

Remarks:

- For single-valued functionals such as the mean, the notions of selective and exhaustive elicitability are equivalent.

Two modes of elicitability

Remarks:

- For single-valued functionals such as the mean, the notions of selective and exhaustive elicitability are equivalent.
- Forecasting / regression in the exhaustive sense is more ambitious than in the selective sense!

Two modes of elicitability

Remarks:

- For single-valued functionals such as the mean, the notions of selective and exhaustive elicitability are equivalent.
- Forecasting / regression in the exhaustive sense is more ambitious than in the selective sense!
- Quantiles are selectively elicitable. Strictly consistent selective loss functions are given by $S\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

$$L(x, y) = (\mathbb{1}\{y \leqslant x\} - \alpha)(g(x) - g(y)),$$

where g is strictly increasing.

Two modes of elicitability

Remarks:

- For single-valued functionals such as the mean, the notions of selective and exhaustive elicitability are equivalent.
- Forecasting / regression in the exhaustive sense is more ambitious than in the selective sense!
- Quantiles are selectively elicitable. Strictly consistent selective loss functions are given by $S\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

  $$L(x, y) = (\mathbb{1}\{y \leqslant x\} - \alpha)(g(x) - g(y)),$$

  where g is strictly increasing.
- What about their exhaustive elicitability?

Mutual Exclusivity

Theorem 5 (F, Hlavinová, Rudloff (2018))

Under weak regularity conditions, a set-valued functional is

- *either selectively elicitable*
- *or exhaustively elicitable*
- *or not elicitable at all.*

Mutual Exclusivity

Theorem 5 (F, Hlavinová, Rudloff (2018))

Under weak regularity conditions, a set-valued functional is

- *either selectively elicitable*
- *or exhaustively elicitable*
- *or not elicitable at all.*

Novel structural insight of its own!

Mutual Exclusivity

Theorem 5 (F, Hlavinová, Rudloff (2018))

Under weak regularity conditions, a set-valued functional is

- *either selectively elicitable*
- *or exhaustively elicitable*
- *or not elicitable at all.*

Novel structural insight of its own!

Implications:

- Quantiles are generally not exhaustively elicitable!

Mutual Exclusivity

Theorem 5 (F, Hlavinová, Rudloff (2018))

Under weak regularity conditions, a set-valued functional is

- *either selectively elicitable*
- *or exhaustively elicitable*
- *or not elicitable at all.*

Novel structural insight of its own!

Implications:

- Quantiles are generally not exhaustively elicitable!
- What about systemic risk measures?

Identifiability

- An identification function (moment function in Econometrics) is a function $V\colon \mathsf{A} \times \mathsf{O} \to \mathbb{R}$.
- V selectively identifies $T\colon \mathcal{F} \to 2^{\mathsf{A}}$ if

$$\mathbf{E}_F[V(x, Y)] = 0 \quad \Longleftrightarrow \quad x \in T(F)$$

  for all $F \in \mathcal{F}$ and for all $x \in \mathsf{A}$.
- V exhaustively identifies $T\colon \mathcal{F} \to \mathsf{A}$ if

$$\mathbf{E}_F[V(x, Y)] = 0 \quad \Longleftrightarrow \quad x = T(F)$$

  for all $F \in \mathcal{F}$ and for all $x \in \mathsf{A}$.

Identifiability results

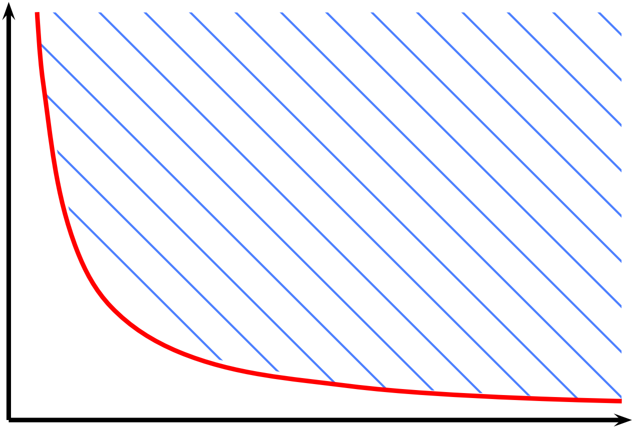
Consider the boundary of R

$$R_0(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) = 0\}.$$

Identifiability results

Consider the boundary of R

$$R_0(Y) = \{k \in \mathbb{R}^n \,|\, \rho(\Lambda(Y+k)) = 0\}.$$



Proposition 6 (F, Hlavinová, Rudloff (2018+))

Let $V_\rho \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be an oriented identification function for* ρ*. That is*

$$\mathbf{E}_F[V_\rho(x, Z)] \begin{cases} < 0, & x < \rho(Z) \\ = 0, & x = \rho(Z) \\ > 0, & x > \rho(Z). \end{cases}$$

Proposition 6 (F, Hlavinová, Rudloff (2018+))

Let $V_\rho \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be an oriented identification function for* ρ. *That is*

$$\mathbf{E}_F[V_\rho(x, Z)] \begin{cases} < 0, & x < \rho(Z) \\ = 0, & x = \rho(Z) \\ > 0, & x > \rho(Z). \end{cases}$$

Then $R_0 = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) = 0\}$ *is selectively identifiable with the identification function*

$$V_{R_0} \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad V_{R_0}(k, y) = V_\rho(0, \Lambda(y + k)).$$

Proposition 6 (F, Hlavinová, Rudloff (2018+))

Let $V_\rho \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be an oriented identification function for ρ. That is

$$\mathbf{E}_F[V_\rho(x, Z)] \begin{cases} < 0, & x < \rho(Z) \\ = 0, & x = \rho(Z) \\ > 0, & x > \rho(Z). \end{cases}$$

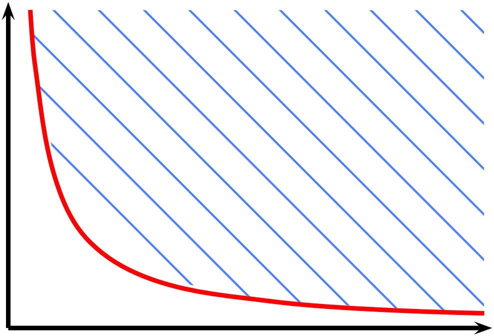
Then $R_0 = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) = 0\}$ is selectively identifiable with the identification function

$$V_{R_0} \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad V_{R_0}(k, y) = V_\rho(0, \Lambda(y + k)).$$

V_{R_0} is oriented in the sense that

$$\mathbf{E}_F[V_{R_0}(k, Y)] \begin{cases} < 0, & k \notin R(Y) \\ = 0, & k \in R_0(Y) \\ > 0, & k \in R(Y) \setminus R_0(Y). \end{cases}$$

Illustration



$$\mathbf{E}_F[V_{R_0}(k, Y)] \begin{cases} < 0, & k \notin R(Y) \\ = 0, & k \in R_0(Y) \\ > 0, & k \in R(Y) \backslash R_0(Y). \end{cases}$$

Strong elicitability of R

Theorem 7 (F, Hlavinová, Rudloff (2018+))

- *Let* $V_{R_0}: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *be an oriented selective identification function for* R_0.
- *Let* π *be a measure on* $\hat{\mathcal{B}}(\mathbb{R}^n)$ *that assigns positive mass to any open, non-empty set.*

Under some integrability conditions, the loss function

$$L_R \colon \mathsf{A} \times \mathbb{R}^n \to \mathbb{R}, \qquad L_R(K, y) = -\int_K V_{R_0}(k, y)\,\pi(\mathrm{d}k)$$

is a strictly consistent exhaustive loss function for R, *where*

$$\mathsf{A} \subset \left\{K \in 2^{\mathbb{R}^n} \mid K = K + \mathbb{R}^n_+\right\}$$

is the collection of closed upper subsets of $\mathbb{R}^n$.

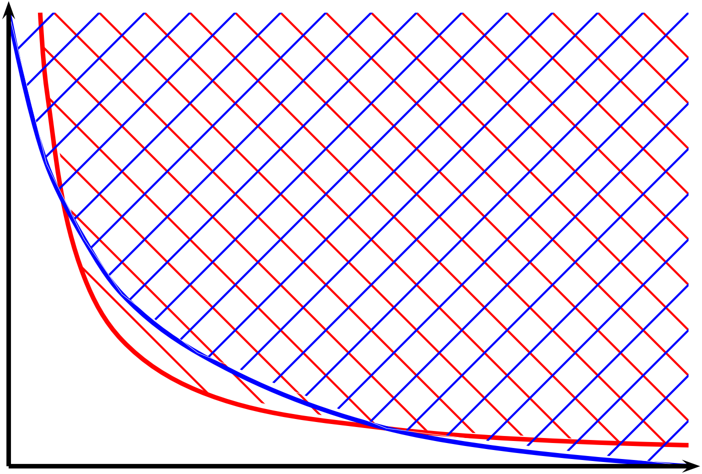


Figure: Illustration of a systemic risk measure $R(Y) = \{k \in \mathbb{R}^n \mid \rho(\Lambda(Y + k)) \leqslant 0\}$ and some misspecified forecast K.

Order-Sensitivity

Can we compare two misspecified forecasts?

Order-Sensitivity

Can we compare two misspecified forecasts?

Proposition 8

The loss functions are order-sensitive with respect to $\subseteq$. That is, for any $K_1, K_2 \in \mathsf{A}$

$$R(Y) \subseteq K_1 \subseteq K_2 \text{ or } K_2 \subseteq K_1 \subseteq R(Y) \implies \mathbf{E}[L_R(K_1, Y)] \leqslant \mathbf{E}[L_R(K_2, Y)].$$

Order-Sensitivity

Can we compare two misspecified forecasts?

Proposition 8

The loss functions are order-sensitive *with respect to* $\subseteq$*. That is, for any* $K_1, K_2 \in \mathsf{A}$

$$R(Y) \subseteq K_1 \subseteq K_2 \text{ or } K_2 \subseteq K_1 \subseteq R(Y) \implies \mathbf{E}[L_R(K_1, Y)] \leqslant \mathbf{E}[L_R(K_2, Y)].$$

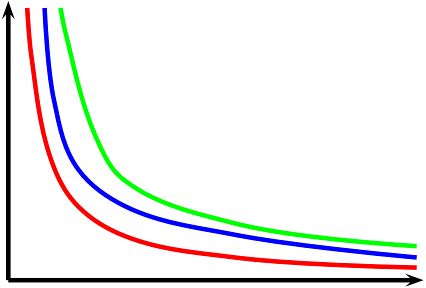


Figure: Illustration of $R(Y) \subseteq K_1 \subseteq K_2$.

Remarks

- We have identifiability results for *Efficient Cash-Invariant Allocation Rules (EARs)*. The corresponding identification functions are *functional-valued*.

Remarks

- We have identifiability results for *Efficient Cash-Invariant Allocation Rules (EARs)*. The corresponding identification functions are *functional-valued*.

- $$R_{\mathrm{ES}_\alpha}(Y) = \{k \in \mathbb{R}^n \mid \mathrm{ES}_\alpha(\Lambda(Y + k)) \leqslant 0\}$$

  is jointly elicitable with the functional-valued risk measure

  $$\mathbb{R}^n \ni k \mapsto \mathrm{VaR}_\alpha(\Lambda(Y + k))\,.$$

Questions & Problems

- Characterisation of the class of strictly consistent exhaustive loss functions for R.

Questions & Problems

- Characterisation of the class of strictly consistent exhaustive loss functions for R.
- What are "nice choices" of π, leading to desirable properties (translation invariance, homogeneity, ...).

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.
- Two modes of elicitability allow for a rigorous treatment of set-valued functionals such as quantiles or systemic risk measures.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.
- Two modes of elicitability allow for a rigorous treatment of set-valued functionals such as quantiles or systemic risk measures.
- Structural insight: Selective and exhaustive elicitability are mutually exclusive.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.
- Two modes of elicitability allow for a rigorous treatment of set-valued functionals such as quantiles or systemic risk measures.
- Structural insight: Selective and exhaustive elicitability are mutually exclusive.
- First (interesting) case of strictly consistent loss functions taking sets as arguments.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.
- Two modes of elicitability allow for a rigorous treatment of set-valued functionals such as quantiles or systemic risk measures.
- Structural insight: Selective and exhaustive elicitability are mutually exclusive.
- First (interesting) case of strictly consistent loss functions taking sets as arguments.
- Possibility to do comparative backtests of Diebold-Mariano type.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.
- Two modes of elicitability allow for a rigorous treatment of set-valued functionals such as quantiles or systemic risk measures.
- Structural insight: Selective and exhaustive elicitability are mutually exclusive.
- First (interesting) case of strictly consistent loss functions taking sets as arguments.
- Possibility to do comparative backtests of Diebold-Mariano type.
- M-estimation where one minimises over sets.

Summary

- The risk measures we consider
  - capture the dependence structure;
  - are sensitive with respect to capital allocations;
  - take an *ex ante* view specifying the capital allocations that prevent a crises.
- They are inherently set-valued functionals.
- Two modes of elicitability allow for a rigorous treatment of set-valued functionals such as quantiles or systemic risk measures.
- Structural insight: Selective and exhaustive elicitability are mutually exclusive.
- First (interesting) case of strictly consistent loss functions taking sets as arguments.
- Possibility to do comparative backtests of Diebold-Mariano type.
- M-estimation where one minimises over sets.
- Regression with set-valued models.

Further Reading

- Main reference for this talk:

  T. Fissler, J. Hlavinová, and B. Rudloff. Elicitability and identifiability of systemic risk measures.

  In preparation, 2018

- Measures of Systemic Risk:

  Z. Feinstein, B. Rudloff, and S. Weber. Measures of Systemic Risk.

  SIAMJ. Financial Math., 8:672–708, 2017

- Good introduction to elicitability:

  T. Gneiting. Making and evaluating point forecasts.

  Journal of the American Statistical Association, 106:746–762, 2011

- Elicitability of vector-valued functionals and elicitability of (VaR, ES):

  T. Fissler and J. F. Ziegel. Higher order elicitability and Osband's principle.

  Annals of Statistics, 44:1680–1707, 2016

Thank you for your attention!