What Is a Good Risk Measure: Bridging the Gaps between Robustness, Subadditivity, Prospect Theory, and Insurance Risk Measures

C.C. Heyde\textsuperscript{1}  S.G. Kou\textsuperscript{2}  X.H. Peng\textsuperscript{2}

\textsuperscript{1}Department of Statistics
Columbia University

\textsuperscript{2}Department of IEOR
Columbia University
Outline

1. Introduction
   - Review of Risk Measures

2. Motivation and Examples
   - What is an Axiom
   - Some Concepts of the Law
   - Examples and Tail Conditional Median

3. Reasons to Relax Subadditivity
   - Review of Existing Reasons
   - Two New Reasons: Prospect Theory and Robustness

4. Natural Risk Statistic
   - Axioms and Characterization of Natural Risk Statistic
   - Comparing the Natural Risk Statistic, Coherent Risk Measure and Insurance Risk Measure
1 Introduction
   - Review of Risk Measures

2 Motivation and Examples
   - What is an Axiom
   - Some Concepts of the Law
   - Examples and Tail Conditional Median

3 Reasons to Relax Subadditivity
   - Review of Existing Reasons
   - Two New Reasons: Prospect Theory and Robustness

4 Natural Risk Statistic
   - Axioms and Characterization of Natural Risk Statistic
   - Comparing the Natural Risk Statistic, Coherent Risk Measure and Insurance Risk Measure
Risk: future loss of a position—random variable $X$

Measure of risk: mapping from a set of risks to the real line

$$X \to \rho(X)$$

Examples:
- margin requirement for financial trading
- capital requirement against market risk
- insurance risk premium
The First Measure: Coherent Measure of Risk (Artzner et al. 1999, Huber 1981)

- **Translation invariance:** $\rho(X + a) = \rho(X) + a$
- **Positive homogeneity:** $\rho(\lambda X) = \lambda \rho(X), \lambda \geq 0$
- **Monotonicity:** $\rho(X) \leq \rho(Y)$, if $X \leq Y$
- **Subadditivity:** $\rho(X + Y) \leq \rho(X) + \rho(Y)$

Characterization of coherent measure of risk

- **Maximum expected loss**

$$\rho(X) = \sup_{P \in \mathcal{P}} E^P[X]$$

- **Acceptance sets**

$$\rho(X) = \min\{m : X - m \in \text{acceptance set}\}$$
Coherent Measure of Risk

The First Measure: Coherent Measure of Risk (Artzner et al. 1999, Huber 1981)

- Translation invariance: \( \rho(X + a) = \rho(X) + a \)
- Positive homogeneity: \( \rho(\lambda X) = \lambda \rho(X), \lambda \geq 0 \)
- Monotonicity: \( \rho(X) \leq \rho(Y), \text{ if } X \leq Y \)
- Subadditivity: \( \rho(X + Y) \leq \rho(X) + \rho(Y) \)

Characterization of coherent measure of risk

- Maximum expected loss

\[
\rho(X) = \sup_{P \in \mathcal{P}} E^P[X]
\]

- Acceptance sets

\[
\rho(X) = \min\{m : X - m \in \text{acceptance set}\}
\]
Insurance Measure of Risk

\(X\) and \(Y\) are comonotonic if and only if

\[
\forall \omega_1, \omega_2 \in \Omega, (X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0
\]

The Second Measure: Insurance Measure of Risk (Wang et al. 1997)

- Law invariance: \(\rho(X) = \rho(Y)\), if \(X\) and \(Y\) have the same distribution
- Monotonicity: \(\rho(X) \leq \rho(Y)\), if \(X \leq Y\)
- Comonotonic additivity:
  \(\rho(X + Y) = \rho(X) + \rho(Y)\), if \(X\) and \(Y\) are comonotonic
- Continuity:
  \[
  \lim_{d \to 0} \rho(\max(X - d, 0)) = \rho(X^+), \quad \lim_{d \to \infty} \rho(\min(X, d)) = \rho(X)
  \]
  \[
  \lim_{d \to -\infty} \rho(\max(X, d)) = \rho(X)
  \]
- Scale normalization: \(\rho(1) = 1\)
Insurance Measure of Risk

\(X\) and \(Y\) are comonotonic if and only if

\[
\forall \omega_1, \omega_2 \in \Omega, (X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0
\]

The Second Measure: Insurance Measure of Risk (Wang et al. 1997)

- **Law invariance:** \(\rho(X) = \rho(Y)\), if \(X\) and \(Y\) have the same distribution
- **Monotonicity:** \(\rho(X) \leq \rho(Y)\), if \(X \leq Y\)
- **Comonotonic additivity:**
  \(\rho(X + Y) = \rho(X) + \rho(Y)\), if \(X\) and \(Y\) are comonotonic
- **Continuity:**
  \[
  \lim_{d \to 0} \rho(\max(X - d, 0)) = \rho(X^+), \quad \lim_{d \to \infty} \rho(\min(X, d)) = \rho(X)
  \]
  \[
  \lim_{d \to -\infty} \rho(\max(X, d)) = \rho(X)
  \]
- **Scale normalization:** \(\rho(1) = 1\)
It does not require subadditivity

\( \rho(\cdot) \) is an insurance measure of risk if and only if \( \rho(\cdot) \) has a Choquet integral representation:

\[
\rho(X) = \int Xd(g \circ P) = \int_{-\infty}^{0} (g(P(X > t)) - 1)dt + \int_{0}^{\infty} g(P(X > t))dt
\]

where \( g(\cdot) \) is increasing, \( g(0) = 0, g(1) = 1 \)

It does not generate scenarios, unlike the coherent risk measure
Comonotonic Additivity Does Not Hold in Two Scenarios

- Probability space: \((\Omega = \{\omega_1, \omega_2, \omega_3\}, \mathcal{F}, P), P(\omega_i) = \frac{1}{3}\)
- Two comonotonic random variables
  \[
  (Z(\omega_1), Z(\omega_2), Z(\omega_3)) = (3, 2, 4), \ Y = Z^2
  \]
- Two distortion function
  \[
  g_1\left(\frac{1}{3}\right) = 0.5, \ g_1\left(\frac{2}{3}\right) = 1.0; \quad g_2\left(\frac{1}{3}\right) = 0.72, \ g_2\left(\frac{2}{3}\right) = 0.8
  \]
- Risk measure incorporating two scenarios:
  \[
  \rho(X) = \max \left\{ \int X \, d(g_1 \circ P), \int X \, d(g_2 \circ P) \right\}
  \]
- Only strict comonotonic subadditivity holds!
  \[
  \rho(Z + Y) = 9.28 < \rho(Z) + \rho(Y) = 2.5 + 6.8 = 9.3
  \]
Two Definitions in Cambridge English Dictionary:
(1) “FORMAL: a statement or principle which is generally accepted to be true, but is not necessarily so.”
(2) “SPECIALIZED: a formal statement or principle in mathematics, science, etc., from which other statements can be obtained.”

Axioms are not theorems. By definition, they cannot be proved right or wrong.

It is useful to provide alternative axioms so that people can have a choice.

Axioms should be as general as possible; otherwise, there is almost no difference between axioms and assumptions.
Large literature on the subject. See, e.g., the textbook by H. Hart (1994) *The Concept of Law*.

1. “Legal Realism”. A law is only a guideline for judges and enforcement officers, i.e. a law is only intended to be the average of what the judges and officers will decide.
   - This requests the robustness of the law.
2. “Legal Positivism”. A law should reflect the society norm.
   - This requests consistency to the social behavior of the society.
3. “The Separability Thesis”. Law and internal standards (e.g. morality) are conceptually distinct. We should expect laws to be less restrictive than internal standards.
   - Here we talk about the risk measures for law and regulation, not internal risk measures within firms.
   - For the external risk measures, it has to be robust and consistent with what the society practices.
An Example of Traffic Limit

- Not by studying physics of road conditions
- “85 rule”
- Recognizing that people constantly drive 5 to 10 miles beyond the traffic limit.
Example: Value-at-risk (VaR)

- Value-at-risk (VaR) at level $\alpha$:
  
  $$\text{VaR}_\alpha(X) = \alpha\text{-quantile of } X, \; \alpha \in [95\%, 99.9\%]$$

- Value-at-risk does not satisfy subadditivity
  
  $$\text{VaR}_\alpha(X + Y) \nless \text{ VaR}_\alpha(X) + \text{VaR}_\alpha(Y), \text{ in general}$$

- **Inconsistency**: 99% VaR is the primary measure of risk imposed by Basel Accord, but it is excluded by coherent measure of risk

- Value-at-risk satisfies axioms for the insurance risk measure
Tail conditional expectation (TCE) at level $\alpha$ (Artzner et.al, 1999):

$$TCE_\alpha(X) = E[X | X \geq VaR_\alpha(X)]$$

Tail conditional expectation is also called conditional value-at-risk (Rockafellar & Uryasev, 2002) and expected shortfall (Acerbi, 2002)

Tail conditional expectation is a coherent measure of risk, at least for continuous random variables
Tail conditional median (TCM) at level $\alpha$:

$$TCM_\alpha(X) = \text{median}[X | X \geq \text{VaR}_\alpha(X)] = \text{VaR}_{\frac{1+\alpha}{2}}(X)$$

- Tail conditional median is equal to VaR at a higher level.
- Tail conditional median does NOT satisfy subadditivity.
- Tail conditional median is more robust than tail conditional expectation.
Tail conditional median (TCM) at level $\alpha$:

$$TCM_\alpha(X) = \text{median}[X | X \geq \text{VaR}_\alpha(X)] = \text{VaR}_{1+\alpha}^\frac{1}{2}(X)$$

- Tail conditional median is equal to VaR at a higher level
- Tail conditional median does NOT satisfy subadditivity
- Tail conditional median is more robust than tail conditional expectation
Tail conditional median (TCM) at level $\alpha$:

$$TCM_\alpha(X) = median[X \mid X \geq VaR_\alpha(X)] = VaR_{\frac{1+\alpha}{2}}(X)$$

- Tail conditional median is equal to VaR at a higher level
- Tail conditional median does NOT satisfy subadditivity
- Tail conditional median is more robust than tail conditional expectation
## TCM and TCE for Auto Insurance Claim

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$TCE_\alpha$</th>
<th>$TCM_\alpha$</th>
<th>$TCE_\alpha - TCM_\alpha$</th>
<th>$\frac{TCE_\alpha - TCM_\alpha}{TCE_\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.0%</td>
<td>6390627.05</td>
<td>4489416.38</td>
<td>1901210.66</td>
<td>29.75%</td>
</tr>
<tr>
<td>98.5%</td>
<td>4454513.70</td>
<td>1682970.01</td>
<td>2771543.68</td>
<td>62.22%</td>
</tr>
<tr>
<td>98.0%</td>
<td>3681944.04</td>
<td>1384060.89</td>
<td>2297883.14</td>
<td>62.41%</td>
</tr>
<tr>
<td>97.5%</td>
<td>3014237.87</td>
<td>1039186.87</td>
<td>1975051.00</td>
<td>65.52%</td>
</tr>
<tr>
<td>97.0%</td>
<td>2579508.48</td>
<td>962778.28</td>
<td>1616730.20</td>
<td>62.68%</td>
</tr>
<tr>
<td>96.5%</td>
<td>2333814.60</td>
<td>851033.85</td>
<td>1482780.74</td>
<td>63.53%</td>
</tr>
<tr>
<td>96.0%</td>
<td>2073066.45</td>
<td>705136.33</td>
<td>1367930.11</td>
<td>65.99%</td>
</tr>
<tr>
<td>95.5%</td>
<td>1865231.51</td>
<td>676514.44</td>
<td>1188717.07</td>
<td>63.73%</td>
</tr>
<tr>
<td>95.0%</td>
<td>1736077.53</td>
<td>662045.27</td>
<td>1074032.25</td>
<td>61.87%</td>
</tr>
</tbody>
</table>

Which one is better, Tail Conditional Median or Tail Conditional Expectation?
Outline

1. Introduction
   - Review of Risk Measures

2. Motivation and Examples
   - What is an Axiom
   - Some Concepts of the Law
   - Examples and Tail Conditional Median

3. Reasons to Relax Subadditivity
   - Review of Existing Reasons
   - Two New Reasons: Prospect Theory and Robustness

4. Natural Risk Statistic
   - Axioms and Characterization of Natural Risk Statistic
   - Comparing the Natural Risk Statistic, Coherent Risk Measure and Insurance Risk Measure
Reason 1: Tail Subadditivity of VaR

- If $X$ and $Y$ have elliptical distributions, then (Emberchts, et. al, 1998)
  \[ \text{VaR}_\alpha(X + Y) \leq \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) \]

- VaR has subadditivity in the tail region (Daníelsson et al. 2005)
  If $X$ and $Y$ have jointly regularly varying tails with tail index bigger than 1, then there exists $\alpha_0 \in (0, 1)$, such that
  \[ \text{VaR}_\alpha(X + Y) \leq \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y), \forall \alpha \in (\alpha_0, 1) \]

- VaR has subadditivity in simulations when $\alpha \in [95\%, 99\%]$. (Daníelsson et al. 2005)

- The counterexamples for subadditivity may be too narrow and may be pathological.
Reason 2: A Merger May Create Extra Risk

- Merger of two firms removes the fire walls between them and hence increase the risk of complete breakdown of the two firms as a whole. (Dhaene et al. 2005)
- By splitting a risky trading business into separate sub-firms, the parent firm can reduce the risk of bankruptcy.
- Example: collapse of Britain’s Barings Bank is caused by a single trader, Nick Leeson, in Singapore.
Rational people systematically violate the axioms of expected utility theory, e.g., the Ellsberg paradoxes.

Prospect theory (Kahneman & Tversky, 1979, 1992, 1993) is able to explain the major violations of expected utility theory

- Four patterns of risk attitudes: People are risk seeking for losses of moderate or high probability
- It is better to postulate preference axioms on comonotonic random variables rather than on arbitrary random variables

Non-comonotonic random loss can violate subadditivity
An Example Violating Subadditivity

A fair coin is flipped

- Three positions:
  - H: losing $10,000 if it comes up “head”
  - T: losing $10,000 if it comes up “tail”
  - S: losing $5,000 for sure

- $\rho(S) = 5,000$, $\rho(H) = \rho(T)$
- $\rho(H) < \rho(S)$ because people are risk seeking in face of loss
- $\rho(H + T) = \rho($sure loss of 10,000$) = 10,000$

Subadditivity does not hold:

$$\rho(H) + \rho(T) < 10,000 = \rho(H + T)!$$
An Example Violating Subadditivity

A fair coin is flipped

- Three positions:
  - H: losing $10,000 if it comes up “head”
  - T: losing $10,000 if it comes up “tail”
  - S: losing $5,000 for sure

- $\rho(S) = 5,000$, $\rho(H) = \rho(T)$
- $\rho(H) < \rho(S)$ because people are risk seeking in face of loss
- $\rho(H + T) = \rho($sure loss of 10,000$) = 10,000$

Subadditivity does not hold:

$$\rho(H) + \rho(T) < 10,000 = \rho(H + T)!$$
Reason 4: from Robustness Viewpoint

Enforcement of risk measure:

- A regulator imposes a risk measure
- Each firm calculates the level of risk and reports to the regulator
- Each firm has the freedom to choose its own internal models
- For example, Basel Accord imposes 99%-VaR as the risk measure, banks can use their own internal models in the calculation

The risk measure should be robust with respect to underlying models, otherwise the firms can choose a model which produces a low level of risk
Another issue: How to specify a correct model?

- Difficulty: heaviness of tail distribution is hard to determine
- With 5,000 observations one cannot distinguish between the Laplace distribution and the T-distributions. (Heyde & Kou, 2004)
- Since we cannot specify the correct model to some extent, the risk measure should have robustness with respect to models.
- The risk measure should also be robust with respect to data

Tail Conditional Expectation (TCE) is sensitive to the model and outliers in the data.
Another issue: How to specify a correct model?

- Difficulty: heaviness of tail distribution is hard to determine
- With 5,000 observations one can not distinguish between the Laplace distribution and the T-distributions. (Heyde & Kou, 2004)
- Since we can not specify the correct model to some extent, the risk measure should have robustness with respect to models.
- The risk measure should also be robust with respect to data

Tail Conditional Expectation (TCE) is sensitive to the model and outliers in the data
TCE is Sensitive to the Heaviness of Tail Distribution

TCM and TCE for Laplace and T-distributions, $\alpha \in [99\%, 99.9\%]$
TCE is Sensitive to the Heaviness of Tail Distribution (continued)

TCM and TCE for Laplace and T-distributions, $\alpha \in [95\%, 99\%]$
TCM is More Robust Than TCE

- Influence function of an estimator $T(F)$:

$$IF(x, T, F) = \lim_{\varepsilon \downarrow 0} \frac{T((1 - \varepsilon)F + \varepsilon \delta_x) - T(F)}{\varepsilon}.$$ 

- TCM has **bounded** influence function:

$$\sup_u IF(u, TCM_\alpha, X) < \infty$$

- TCE has **unbounded** influence function:

$$\sup_u IF(u, TCE_\alpha, X) = \infty$$
1. **Introduction**
   - Review of Risk Measures

2. **Motivation and Examples**
   - What is an Axiom
   - Some Concepts of the Law
   - Examples and Tail Conditional Median

3. **Reasons to Relax Subadditivity**
   - Review of Existing Reasons
   - Two New Reasons: Prospect Theory and Robustness

4. **Natural Risk Statistic**
   - Axioms and Characterization of Natural Risk Statistic
   - Comparing the Natural Risk Statistic, Coherent Risk Measure and Insurance Risk Measure
Axioms for Natural Risk Statistic

Data $\tilde{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$
Risk statistic: a mapping from the data set to the real line:

$$\hat{\rho} : \mathbb{R}^n \rightarrow \mathbb{R}$$
$$\tilde{x} \rightarrow \hat{\rho}(\tilde{x})$$

Axioms for natural risk statistic:

- Positive homogeneity and translation invariance:
  $$\hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \ \forall \tilde{x} \in \mathbb{R}^n, \ a \geq 0, \ b \in \mathbb{R}$$

- Monotonicity: $\hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y})$, if $\tilde{x} \leq \tilde{y}$, i.e., $x_i \leq y_i, i = 1, \ldots, n$.

- Comonotonic subadditivity: $\hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y})$, if $\tilde{x}$ and $\tilde{y}$ are comonotonic, i.e., $(x_i - x_j)(y_i - y_j) \geq 0$ for any $i \neq j$.

- Empirical law invariance:
  $$\hat{\rho}((x_1, \ldots, x_n)) = \hat{\rho}((x_{i_1}, \ldots, x_{i_n})), \ for \ any \ permutation \ (i_1, \ldots, i_n).$$
Axioms for Natural Risk Statistic

Data \( \tilde{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \)

Risk statistic: a mapping from the data set to the real line:

\[
\hat{\rho} : \mathbb{R}^n \to \mathbb{R} \\
\tilde{x} \mapsto \hat{\rho}(\tilde{x})
\]

Axioms for natural risk statistic:

- Positive homogeneity and translation invariance:
  \( \hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \forall \tilde{x} \in \mathbb{R}^n, a \geq 0, b \in \mathbb{R} \)

- Monotonicity: \( \hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \leq \tilde{y}, \text{ i.e., } x_i \leq y_i, i = 1, \ldots, n. \)

- Comonotonic subadditivity: \( \hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \text{ and } \tilde{y} \text{ are comonotonic, i.e., } (x_i - x_j)(y_i - y_j) \geq 0 \text{ for any } i \neq j. \)

- Empirical law invariance:
  \( \hat{\rho}((x_1, \ldots, x_n)) = \hat{\rho}((x_{i_1}, \ldots, x_{i_n})), \text{ for any permutation } (i_1, \ldots, i_n). \)
Axioms for Natural Risk Statistic

Data $\tilde{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$

Risk statistic: a mapping from the data set to the real line:

$$\hat{\rho} : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$\tilde{x} \mapsto \hat{\rho}(\tilde{x})$$

Axioms for natural risk statistic:

- **Positive homogeneity and translation invariance:**
  $$\hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \; \forall \tilde{x} \in \mathbb{R}^n, \; a \geq 0, \; b \in \mathbb{R}$$

- **Monotonicity:**
  $$\hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \leq \tilde{y}, \text{ i.e., } x_i \leq y_i, \; i = 1, \ldots, n.$$

- **Comonotonic subadditivity:**
  $$\hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \text{ and } \tilde{y} \text{ are comonotonic, i.e., } (x_i - x_j)(y_i - y_j) \geq 0 \text{ for any } i \neq j.$$  

- **Empirical law invariance:**
  $$\hat{\rho}((x_1, \ldots, x_n)) = \hat{\rho}((x_{i_1}, \ldots, x_{i_n})), \text{ for any permutation } (i_1, \ldots, i_n).$$
Axioms for Natural Risk Statistic

Data $\tilde{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$

Risk statistic: a mapping from the data set to the real line:

$$\hat{\rho} : \mathbb{R}^n \rightarrow \mathbb{R}$$

$\tilde{x} \mapsto \hat{\rho}(\tilde{x})$

Axioms for natural risk statistic:

- **Positive homogeneity and translation invariance:**
  $$\hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \ \forall \tilde{x} \in \mathbb{R}^n, \ a \geq 0, \ b \in \mathbb{R}$$

- **Monotonicity:**
  $$\hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}), \ \text{if} \ \tilde{x} \leq \tilde{y}, \ \text{i.e.,} \ x_i \leq y_i, \ i = 1, \ldots, n.$$

- **Comonotonic subadditivity:**
  $$\hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}), \ \text{if} \ \tilde{x} \ \text{and} \ \tilde{y} \ \text{are comonotonic, i.e.,} \ (x_i - x_j)(y_i - y_j) \geq 0 \ \text{for any} \ i \neq j.$$

- **Empirical law invariance:**
  $$\hat{\rho}((x_1, \ldots, x_n)) = \hat{\rho}((x_{i_1}, \ldots, x_{i_n})), \ \text{for any permutation} \ (i_1, \ldots, i_n).$$
Representation Theorem for Natural Risk Statistic

Theorem

Let $x(1), \ldots, x(n)$ be the order statistics of the data $\tilde{x} = (x_1, \ldots, x_n)$ with $x(n)$ being the largest. Then $\hat{\rho}$ is a natural risk statistic if and only if there exists a set of weights $\mathcal{W} = \{\tilde{w} = (w_1, \ldots, w_n)\} \subset \mathbb{R}^n$ with each $\tilde{w} \in \mathcal{W}$ satisfying $\sum_{i=1}^{n} w_i = 1$ and $w_i \geq 0, \forall 1 \leq i \leq n$, such that

$$\hat{\rho}(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \left\{ \sum_{i=1}^{n} w_i x(i) \right\}, \forall \tilde{x} \in \mathbb{R}^n.$$
Theorem

Let \( x(1), \ldots, x(n) \) be the order statistics of the data \( \tilde{x} = (x_1, \ldots, x_n) \) with \( x(n) \) being the largest. Then \( \hat{\rho} \) is a natural risk statistic if and only if there exists a set of weights \( \mathcal{W} = \{ \tilde{w} = (w_1, \ldots, w_n) \} \subset \mathbb{R}^n \) with each \( \tilde{w} \in \mathcal{W} \) satisfying \( \sum_{i=1}^{n} w_i = 1 \) and \( w_i \geq 0, \forall 1 \leq i \leq n \), such that

\[
\hat{\rho}(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \left\{ \sum_{i=1}^{n} w_i x(i) \right\}, \; \forall \tilde{x} \in \mathbb{R}^n.
\]

Natural risk statistic includes VaR and TCM
Another Representation via Acceptance Sets

Representation of natural risk statistic via acceptance sets:

\[ \hat{\rho}(\tilde{x}) = \inf\{m \mid \tilde{x} - m \in \mathcal{A}\}, \quad \forall \tilde{x} \in \mathbb{R}^n. \]

A statistical acceptance set is a subset of \( \mathbb{R}^n \):

- The acceptance set \( \mathcal{A} \) contains \( \mathbb{R}^n_- \) where
  \[ \mathbb{R}^n_- = \{\tilde{x} \in \mathbb{R}^n \mid x_i \leq 0, i = 1, \ldots, n\}. \]

- The acceptance set \( \mathcal{A} \) does not intersect the set \( \mathbb{R}^n_{++} \) where
  \[ \mathbb{R}^n_{++} = \{\tilde{x} \in \mathbb{R}^n \mid x_i > 0, i = 1, \ldots, n\}. \]

- If \( \tilde{x} \) and \( \tilde{y} \) are comonotonic and \( \tilde{x} \in \mathcal{A} \), \( \tilde{y} \in \mathcal{A} \), then
  \[ \lambda \tilde{x} + (1 - \lambda)\tilde{y} \in \mathcal{A}, \quad \forall \lambda \in [0, 1]. \]

- The acceptance set \( \mathcal{A} \) is positively homogeneous, i.e., if \( \tilde{x} \in \mathcal{A} \), then
  \[ \lambda \tilde{x} \in \mathcal{A} \text{ for all } \lambda \geq 0. \]

- If \( \tilde{x} \leq \tilde{y} \) and \( \tilde{y} \in \mathcal{A} \), then \( \tilde{x} \in \mathcal{A} \).

- If \( \tilde{x} \in \mathcal{A} \), then \((x_{i_1}, \ldots, x_{i_n}) \in \mathcal{A}\) for any permutation \((i_1, \ldots, i_n)\).
Another Representation via Acceptance Sets

Representation of natural risk statistic via acceptance sets:

\[ \hat{\rho}(\tilde{x}) = \inf \{ m \mid \tilde{x} - m \in \mathcal{A} \}, \quad \forall \tilde{x} \in \mathbb{R}^n. \]

A statistical acceptance set is a subset of \( \mathbb{R}^n \):

- The acceptance set \( \mathcal{A} \) contains \( \mathbb{R}^n \) where \( \mathbb{R}^n_\leq = \{ \tilde{x} \in \mathbb{R}^n \mid x_i \leq 0, i = 1, \ldots, n \} \).
- The acceptance set \( \mathcal{A} \) does not intersect the set \( \mathbb{R}^n_{++} \) where \( \mathbb{R}^n_{++} = \{ \tilde{x} \in \mathbb{R}^n \mid x_i > 0, i = 1, \ldots, n \} \).
- If \( \tilde{x} \) and \( \tilde{y} \) are comonotonic and \( \tilde{x} \in \mathcal{A}, \tilde{y} \in \mathcal{A} \), then \( \lambda \tilde{x} + (1 - \lambda)\tilde{y} \in \mathcal{A} \), for \( \forall \lambda \in [0, 1] \).
- The acceptance set \( \mathcal{A} \) is positively homogeneous, i.e., if \( \tilde{x} \in \mathcal{A} \), then \( \lambda \tilde{x} \in \mathcal{A} \) for all \( \lambda \geq 0 \).
- If \( \tilde{x} \leq \tilde{y} \) and \( \tilde{y} \in \mathcal{A} \), then \( \tilde{x} \in \mathcal{A} \).
- If \( \tilde{x} \in \mathcal{A} \), then \( (x_{i_1}, \ldots, x_{i_n}) \in \mathcal{A} \) for any permutation \( (i_1, \ldots, i_n) \).
Axioms for coherent risk statistic

- Positive homogeneity and translation invariance:
  \[ \hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \quad \forall \tilde{x} \in \mathbb{R}^n, \quad a \geq 0, \quad b \in \mathbb{R} \]

- Monotonicity: \( \hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}), \) if \( \tilde{x} \leq \tilde{y}, \) i.e., \( x_i \leq y_i, \) \( i = 1, \ldots, n. \)

- Subadditivity: \( \hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}), \) for \( \forall \tilde{x} \) and \( \tilde{y}. \)

Proposition: \( \hat{\rho} \) is a coherent risk statistic if and only if there exists a set of weights \( \mathcal{W} = \{ \tilde{w} = (w_1, \ldots, w_n) \} \subset \mathbb{R}^n, \) such that

\[ \hat{\rho}(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \left\{ \sum_{i=1}^{n} w_i x_i \right\}, \quad \forall \tilde{x} \in \mathbb{R}^n. \]
Axioms for coherent risk statistic

- **Positive homogeneity and translation invariance:** 
  \[ \hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \quad \forall \tilde{x} \in \mathbb{R}^n, \quad a \geq 0, \quad b \in \mathbb{R} \]
- **Monotonicity:** 
  \[ \hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \leq \tilde{y}, \text{ i.e., } x_i \leq y_i, i = 1, \ldots, n. \]
- **Subadditivity:** 
  \[ \hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}), \text{ for } \forall \tilde{x} \text{ and } \tilde{y}. \]

**Proposition:** \( \hat{\rho} \) is a coherent risk statistic if and only if there exists a set of weights \( \mathcal{W} = \{ \tilde{w} = (w_1, \ldots, w_n) \} \subset \mathbb{R}^n \), such that

\[
\hat{\rho}(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \{ \sum_{i=1}^{n} w_i x_i \}, \quad \forall \tilde{x} \in \mathbb{R}^n.
\]
Axioms for law invariant coherent risk statistic

- Positive homogeneity and translation invariance:
  \[ \hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \quad \forall \tilde{x} \in \mathbb{R}^n, \quad a \geq 0, \quad b \in \mathbb{R} \]

- Monotonicity: \( \hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}) \), if \( \tilde{x} \leq \tilde{y} \), i.e., \( x_i \leq y_i, \quad i = 1, \ldots, n \).

- Subadditivity: \( \hat{\rho}(\tilde{x} + \tilde{y}) \leq \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}) \), for \( \forall \tilde{x} \) and \( \tilde{y} \).

- Empirical law invariance:
  \[ \hat{\rho}((x_1, \ldots, x_n)) = \hat{\rho}((x_{i_1}, \ldots, x_{i_n})), \quad \text{for any permutation } (i_1, \ldots, i_n). \]
Theorem

\( \hat{\rho} \) is a law invariant coherent risk statistic if and only if there exists a set of weights \( \mathcal{W} = \{ \tilde{w} = (w_1, \ldots, w_n) \} \subset \mathbb{R}^n \) with each \( \tilde{w} \in \mathcal{W} \) satisfying \( \sum_{i=1}^{n} w_i = 1 \) and \( w_i \geq 0, \forall 1 \leq i \leq n \) and \( w_1 \leq w_2 \leq \ldots \leq w_n \), such that

\[
\hat{\rho}(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \left\{ \sum_{i=1}^{n} w_i x(i) \right\}, \quad \forall \tilde{x} \in \mathbb{R}^n.
\]

- Assigning larger weights to larger observations leads to less robust risk statistics
- Natural risk statistic is more robust
Representation Theorem for Law Invariant Coherent Risk Statistic

Theorem

\( \hat{\rho} \) is a law invariant coherent risk statistic if and only if there exists a set of weights \( \mathcal{W} = \{ \tilde{w} = (w_1, \ldots, w_n) \} \subset \mathbb{R}^n \) with each \( \tilde{w} \in \mathcal{W} \) satisfying \( \sum_{i=1}^{n} w_i = 1 \) and \( w_i \geq 0 \), \( \forall 1 \leq i \leq n \) and \( w_1 \leq w_2 \leq \ldots \leq w_n \), such that

\[
\hat{\rho}(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \left\{ \sum_{i=1}^{n} w_i x(i) \right\}, \forall \tilde{x} \in \mathbb{R}^n.
\]

- Assigning larger weights to larger observations leads to less robust risk statistics
- Natural risk statistic is more robust
Axioms for insurance risk statistic

- **Monotonicity**: \( \hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}) \), if \( \tilde{x} \leq \tilde{y} \), i.e., \( x_i \leq y_i \), \( i = 1, \ldots, n \).
- **Comonotonic additivity**: \( \hat{\rho}(\tilde{x} + \tilde{y}) = \hat{\rho}(\tilde{x}) + \hat{\rho}(\tilde{y}) \), if \( \tilde{x} \) and \( \tilde{y} \) are comonotonic.
- **Scale normalization**: \( \hat{\rho}(1) = 1 \).
- **Empirical law invariance**: 
  \[ \hat{\rho}((x_1, \ldots, x_n)) = \hat{\rho}((x_{i_1}, \ldots, x_{i_n})), \text{ for any permutation } (i_1, \ldots, i_n). \]
Theorem

\( \hat{\rho} \) is an insurance risk statistic if and only if there exists a single weight \( \tilde{w} = (w_1, \ldots, w_n) \) with \( w_i \geq 0 \) for \( i = 1, \ldots, n \) and \( \sum_{i=1}^{n} w_i = 1 \), such that

\[
\hat{\rho}(\tilde{x}) = \sum_{i=1}^{n} w_i x_{(i)}, \quad \forall \tilde{x} \in \mathbb{R}^n.
\]

- Insurance risk statistic is just one L-statistic.
- Insurance risk statistic cannot incorporate different scenarios.
We propose a new risk measure, the natural risk statistic

- It requires **comonotonic subadditivity** instead of subadditivity
- It provides an axiomatic justification for VaR
- It includes tail conditional median (TCM) which is more robust than TCE.
- It is able to incorporate scenario analysis.
Thank you!