Model uncertainty and applications in insurance design

Fangda Liu

Department of Statistics and Actuarial Science
University of Waterloo

Actuarial, Finance, Risk and Insurance Congress
July 25, 2023, Victoria Falls, Zimbabwe

Table of contents

Insurance 101

Insurance is an effective risk management tool used to protect against contingent losses of market participants.

$$X$$

$$\underbrace{I(X)}_{\text{Reimbursement}} \qquad \underbrace{X - I(X)}_{\text{Insured's retained loss}}$$

$$\underbrace{I(X) - \underbrace{\pi(I(X))}_{\text{premium}}}_{\text{Insurer's loss or benefit}} \qquad \underbrace{X - I(X) + \underbrace{\pi(I(X))}_{\text{premium}}}_{\text{Insured's total loss}}$$

where $I \in \mathcal{I}$ is an admissible indemnity function, and π is a premium principle.

Classical optimization problems in insurance

Popular optimal (re-)insurance design problems:

1. Maximize expected utility:

$$\max_{I \in \mathcal{I}} \mathbb{E}\left[v(w - X + I(X) - \pi(I(X)))\right].$$

   - Arrow (1963): optimality of a stop-loss contract.
   - Gerber(1979), Young (1999), Kaluszka (2001,2005), etc.

2. Minimize risk measure:

$$\min_{I \in \mathcal{I}} \rho\left(X - I(X) + \pi(I(X))\right).$$

   - Cai et al. (2008), Kaluszka and Okolewki (2008), Bernard and Tian (2009), Cheung (2010), etc.

All problems are considered under the assumption that **the distribution of X is known.** Can we take this assumption for granted?

Uncertainty

From data to models

- Parameter uncertainty
  Estimation error, simulation error, etc
- Model uncertainty
  Choice of models, complexity of models, etc.

Distributional uncertainty

- Only partial information about the true distribution are observed from the historical data.
- Changes of the underlying risks
- In a conservative decision, the **worst-case distribution** is important

Worst-case scenario

- Suppose an agent faces an underlying risk X
  - ℓ is the loss function/strategy the agent adopts.
  - ρ is the risk measure used to quantify the agent's risk exposure
  - $\mathcal{S}$ is the uncertainty set includes all distributions of alternative risks considered
- From the perspective of risk management, the **worst-case scenario** in which the agent has the largest risk exposure is of special interests.
- The agent's optimization problem with model uncertainty can be formulated as

$$\min_{\ell} \underbrace{\sup_{F \in \mathcal{S}} \rho(\ell(X^F)),}_{\text{worst-case scenario}} \qquad X^F \sim F.$$

Literature

In a financial market, under the mean-variance constraints

- Theorem 1 in El Ghaoui et al. (2003) solves the worst-case VaR where $\mathrm{VaR}_u(X^F) = F^{-1}(u)$
- Theorem 2.9 in Chen et al. (2011) solves the worst-case ES where $\mathrm{ES}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(X)\mathrm{d}u$
- Li (2018) determines the closed-from solutions for worst-case law invariant coherent risk measures

Under both the mean-variance and Wasserstein distance constraints

- Bernard et al. (2020b) Consider both the worst-case and the best-case scenarios:

$$\sup_{F\in\mathcal{S}} \rho(X^F), \qquad \text{and} \qquad \inf_{F\in\mathcal{S}} \rho(X^F)$$

  for a distortion risk measure ρ.

Literature

In the literature of insurance

- Asimit et al. (2017): for $\rho = \mathrm{VaR}, \mathrm{ES}$,

$$\begin{cases} \min_{(I,P)\in\mathcal{I}\times\mathbb{R}} \max_{k\in\mathcal{M}} \{\rho_{\mathcal{P}_k}(X - I(X) + P)\}, \\ \text{s.t. } \omega_0 + (1+\theta)\mathbb{H}_{\mathcal{P}_k}(I(X)) \le P \le \bar{P}, \forall k \in \mathcal{M}. \end{cases}$$

where $\mathcal{P}_k$, $k \in \mathcal{M}$ includes finite many probability measures.

- Birghila and Pflug (2019)

$$\min_{I\in\mathcal{I}} \max_{F\in\mathcal{C}} \{\rho(X^F - I(X^F) + \pi(I(X^F))\}, \quad \text{s.t. } \pi(I(X^F)) \le B$$

where $\mathcal{C}$ is the convex cone of n reference distributions.

- Liu and Mao (2021): for $\rho = \mathrm{VaR}, \mathrm{ES}$,

$$\min_{d\ge 0} \sup_{F\in\mathcal{S}(\mu,\sigma)} \rho(X^F \wedge d + (1+\theta)\mathbb{E}^F[(X^F - d)_+]).$$

where $\mathcal{S}(\mu, \sigma)$ gives first & second moments constraints.

In this talk, we focus on the **worst-case scenario** for an agent

$$\sup_{F \in \mathcal{S}} \rho_h(\ell(X^F)), \qquad X^F \sim F$$

where

- ρ_h is a **distortion risk measure** (e.g. Dhaene et al. (2012)):

$$\rho_h(X^F) = -\int_{-\infty}^{0} h(F(x))\mathrm{d}x + \int_{0}^{\infty} 1 - h(F(x))\mathrm{d}x = \int_{0}^{1} \gamma(u) F^{-1}(u)\mathrm{d}u,$$

  where $h : [0,1] \mapsto [0,1]$ is non-decreasing (convex) with $h(0) = 0$ and $h(1) = 1$, and $\gamma(u) = h'(u)$, $0 < u < 1$
- $\mathcal{S}$ is the **uncertainty set** defined by Wasserstein distance constraints
- ℓ is the **loss function/strategy** the agent adopts.

Uncertainty set with Wasserstein distance constraint

- For $X \sim F$ and $Y \sim G$, for $k \geq 1$, the **Wasserstein distance** is

$$W_k(X,Y) = W_k(F,G) = \left(\int_0^1 \left| F^{-1}(x) - G^{-1}(x) \right|^k \right)^{1/k}.$$

- The uncertainty set with Wasserstein distance constraint

$$\mathcal{S} = \{\text{r. v. } Y : W_k(Y,X) \leq \varepsilon\} = \{\text{distribution } G : W_k(G,F) \leq \varepsilon\}$$

where

  - $X \sim F$ is the reference distribution
  - ε is the tolerant bound for the Wasserstein distance

- Consider **worst-case scenario**

$$\begin{aligned}\sup_{G \in \mathcal{S}} \rho_h(X^G) &= \sup \left\{ \rho_h(X^G) : W_k(G,F) \leq \varepsilon \right\} \\ &= \sup \left\{ \int_0^1 \gamma(u) G^{-1}(u) \mathrm{d}u : W_k(G,F) \leq \varepsilon \right\}\end{aligned}$$

Uncertainty set with Wasserstein distance constraint

Theorem (Proposition 4 in Liu et al. (2022))

For a continuous and convex distortion function h,

$$\sup\left\{\rho_h(X^G) : W_k(G,F) \le \varepsilon\right\} = \rho_h(X^F) + \varepsilon\|\gamma\|_q,$$

where $q = (1 - 1/k)^{-1}$ *with the convention* $0^{-1} = \infty$, *and* $||\cdot||_q$ *is the* $\mathcal{L}_q$*-norm.*

For $k > 1$, *the above maximum value is attained by the worst-case distribution*

$$G^{-1}(t) = F^{-1}(t) + \varepsilon\frac{(\gamma(t))^{q-1}}{\|\gamma\|_q^{q/k}}, \qquad 0 < t < 1.$$

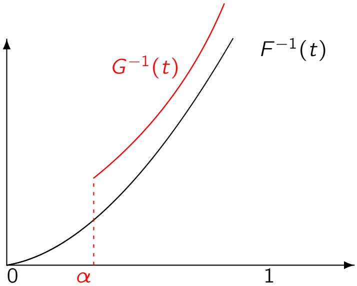
Example – Expected shortfall (ES)

Take $\rho = \mathrm{ES}_\alpha$ for $\alpha \in (0,1)$, then $\rho(X) = \int_0^1 \mathrm{VaR}_t(X)\mathrm{d}h(t)$, where

$$h(t) = \frac{1}{1-\alpha}(t-\alpha)^+ \quad \text{and} \quad \gamma(t) = \frac{1}{1-\alpha}\mathbb{1}_{[\alpha,1]}.$$

The worst-case value is

$$\sup\left\{\mathrm{ES}_\alpha(X^G) : W_k(G,F) \le \varepsilon\right\} = \mathrm{ES}_\alpha(X^F) + \varepsilon \cdot (1-\alpha)^{-1/k}.$$



Uncertainty set with Wasserstein distance constraint

- Uncertainty set is

$$\mathcal{S} = \{G : W_k(G, F) \leq \varepsilon\}$$

where $X^F \sim F$ is considered as a reference distribution, and ε is the tolerant bound for the Wasserstein distance.

- Consider the worst-case scenario:

$$\sup_{G \in \mathcal{S}} \rho_h(\ell(X^G)) = \sup \left\{\rho_h(\ell(X^G)), W_k(G, F) \leq \varepsilon\right\},$$

with two types of **loss functions**:

  - Stop-loss function: (optimal to the utility maximization)

$$\ell(x) = (x - d)^+$$

  - Limited-loss function: (optimal to the VaR minimization)

$$\ell(x) = \min\{x, M\}$$

Stop-loss function

- Take $\ell_1(x) = (x-d)^+$ for $d > \text{ess-inf}(X)$
- Worst-case risk measure

$$\sup\left\{\rho_h((X^G - d)^+) : W_k(G,F) \leq \varepsilon\right\}$$

- For $\beta \in [0,1]$, define $\gamma_{1,\beta} := \gamma \cdot \mathbb{I}_{[\beta,1]}$ which is again a non-negative and increasing function.

$$\begin{aligned}
\sup_{G\in\mathcal{S}} \rho_h\left((X^G - d)^+\right) &= \sup_{G\in\mathcal{S}} \int_{G(0)}^1 \gamma(u)\left(G^{-1}(u) - d\right)\mathrm{d}u \\
&= \sup_{G\in\mathcal{S}} \max_{\beta\in[0,1]} \int_{\beta}^1 \gamma(u)\left(G^{-1}(u) - d\right)\mathrm{d}u \\
&= \sup_{\beta\in[0,1]} \underbrace{\sup_{G\in\mathcal{S}} \int_0^1 \gamma_{1,\beta}(u)\left(G^{-1}(u) - d\right)\mathrm{d}u}_{\text{worst-case without transform}},
\end{aligned}$$

Wasserstein distance constraint and stop-loss transform

Theorem (Cai et al. (2022b))

Take $k \geq 1$ and $q = (1 - 1/k)^{-1}$.

(i) *The worst-case risk measures value is*

$$\sup\left\{\rho_h((X^G - d)^+) : W_k(G, F) \leq \varepsilon\right\}$$
$$= \max_{\beta \in [0,1]} \left\{\int_0^1 \gamma_{1,\beta}(u) F^{-1}(u) \mathrm{d}u + \varepsilon \|\gamma_{1,\beta}\|_q - d\|\gamma_{1,\beta}\|_1\right\}.$$

(ii) *The worst-case distribution is given by*

$$G^{-1}(t) = F^{-1}(t) + \varepsilon \cdot \frac{(\gamma_{1,\beta^*}(t))^{q-1}}{\|\gamma_{1,\beta^*}\|_q^{q/k}}, \quad 0 < t < 1.$$

where β^* is the maximizer in (i).

Example - Expected shortfall

Take $\rho = \mathrm{ES}_\alpha$ for some $\alpha \in (0,1)$.

(i) The worst-case value is

$$\begin{aligned} &\sup\left\{\mathrm{ES}_\alpha((X^G - d)^+) : W_k(G,F) \le \varepsilon\right\} \\ &= \frac{1}{1-\alpha} \max_{\beta\in[\alpha,1]} \left\{(1-\beta)\left(\mathrm{ES}_\beta(X^{\hat{F}}) - d\right) + \varepsilon(1-\beta)^{1/\bar{k}}\right\}. \end{aligned}$$

(ii) The worst-case distribution is

$$G^{-1}(t) = F^{-1}(t) + \varepsilon \cdot \frac{(\gamma_{1,\beta^*}(t))^{q-1}}{\|\gamma_{1,\beta^*}\|_q^{q/k}}$$

where $\gamma_{1,\beta^*} = \frac{1}{1-\alpha}\mathbb{I}_{[\alpha\vee\beta^*,1]}$ and β^* is the solution to the maximization problem in (i).

Example - Wang's premium

- **Wang's premium:**

$$\rho_h(X) = \int_0^\infty (1 - h(F(x)))\,\mathrm{d}u$$

where

$$h(u) = 1 - \Phi(\Phi^{-1}(1-u) + 0.5), \quad 0 \le u \le 1$$

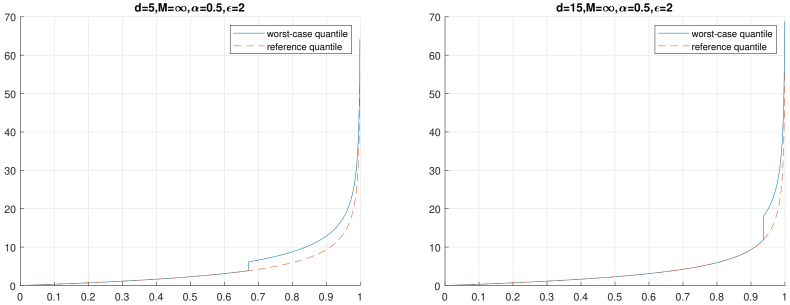
- Take Pareto loss distribution: (heavy tail, non-negative, etc)

$$F(x) = 1 - \left(\frac{12}{x+12}\right)^4, \quad x \ge 0$$

- $\varepsilon = 2$ and $k = 2$, i.e., $\sup\left\{\rho_h(\ell(X^G)), W_2(G,F) \le 2\right\}$

Example - Wang's premium

Figure: Worst-case distributions with stop-loss function.



Limited-loss function

- Take $\ell_2(x) = \max\{x, M\} = x \wedge M$ for $M < \text{ess-sup}(X)$
- Worst-case risk measure

$$\sup\left\{\rho_h(X^G \wedge M) : W_k(G, F) \leq \varepsilon\right\}$$

- Define $\gamma_{2,\beta} = \gamma \cdot \mathbb{I}_{[0,\beta]}$ which is not an increasing function

$$\begin{aligned}\sup_{G\in\mathcal{S}} \rho_h(X^G \wedge M) &= M + \sup_{G\in\mathcal{S}} \int_0^{G(d)} \gamma(u)\left(G^{-1}(u) - M\right)\mathrm{d}u \\ &= M + \sup_{G\in\mathcal{S}} \min_{\beta\in[0,1]} \int_0^1 \gamma_{2,\beta}(u)\left(G^{-1}(u) - M\right)\mathrm{d}u \\ &= M + \min_{\beta\in[0,1]} \sup_{G\in\mathcal{S}} \int_0^1 \gamma_{2,\beta}(u)\left(G^{-1}(u) - M\right)\mathrm{d}u,\end{aligned}$$

by the Min-Max theorem (e.g., Sion et al. (1958))

Wasserstein distance constraint and limited-loss transform

Theorem (Cai et al. (2022b))

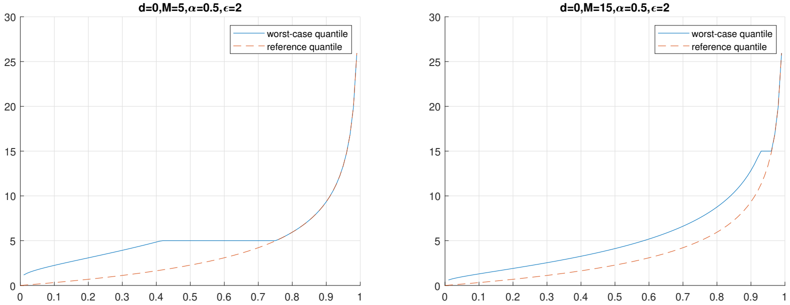
Let $k = 2$. The worst-case distribution is given by

$$(F^*)^{-1}(u) = \begin{cases} F^{-1}(u) + \lambda^* \gamma(u), & \text{for } 0 < u \leq \theta^*, \\ M, & \text{for } \theta^* < u \leq F(M), \\ F^{-1}(u), & \text{for } F(M) < u < 1 \end{cases}$$

where $\lambda^* > 0$ and $\theta^* \in (0, F(M))$ satisfies $W_2(F^*, F) = \varepsilon$.

Example - Wang's premium (cont')

Figure: Worst-case distributions with limited loss function.



Wasserstein distance constraint and limited stop-loss transform

- Wang's premium ρ_h with $h(u) = 1 - \Phi(\Phi^{-1}(1-u) + 0.5)$.
- Exponential reference $F_1(x) = 1 - e^{-x/4}$, $x \geq 0$
- Pareto reference $F_2(x) = 1 - \left(\frac{12}{x+12}\right)^4$
- Limited stop-loss function

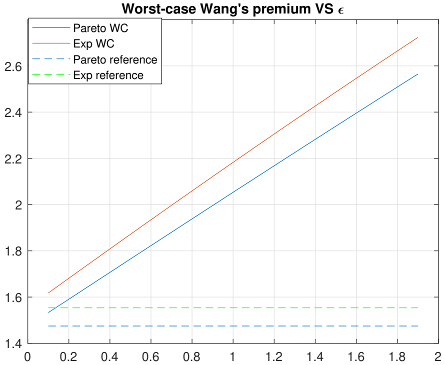
$$\ell(x) = \max\left\{(x-d)^+, M\right\}$$

- Wang's premium in the worst-case:

$$\sup\left\{\rho_h\left(\max\left\{(X^G - d)^+, M\right\}\right), W_2(G, F_i) \leq \varepsilon\right\}, \quad i = 1, 2.$$

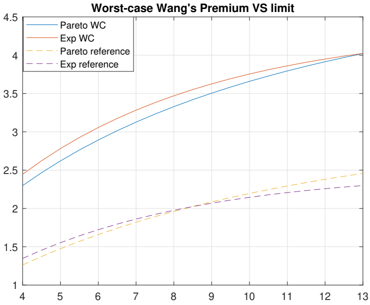
Example - Worst-case Wang's premium VS ε

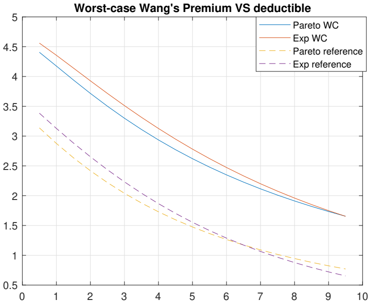
- Fix $d = 5$ and $M = 5$



Example - Worst-case Wang's premium VS deductible and limit

- Fix $\varepsilon = 2$





Uncertainty set with Wasserstein distance and moments constraints

- Take 2nd-order Wasserstein distance. Let $X \sim F$ with $\mathbb{E}[X] = \mu_F$ and $\mathrm{var}(X) = \sigma_F^2$. The uncertainty set is

$$\mathcal{S} = \left\{ G : W_2(F, G) \leq \varepsilon,\ \mathbb{E}[Y^G] = \mu,\ \mathrm{var}(Y^G) = \sigma^2 \right\},$$

- It is not necessary to assume $\mu_F = \mu$ and $\sigma_F^2 = \sigma^2$. Note

$$\varepsilon^2 \geq (\mu - \mu_F)^2 + (\sigma - \sigma_F)^2 \quad \Rightarrow \quad \mathcal{S} \neq \emptyset.$$

Isotonic Projection: For $h \in \mathcal{L}^2(0,1)$, let

$$h^{\uparrow} = \arg\min_{k \in \mathcal{K}} ||h - k||^2,$$

$$\text{where } \mathcal{K} = \left\{ k : (0,1) \mapsto \mathbb{R} \,\middle|\, \int_0^1 k(u)^2 \mathrm{d}u < \infty, k \text{ non-decreasing} \right\}.$$

Notation

- Denote $\gamma_{1,\beta}(u) := \gamma(u)\,\mathbb{1}_{[\beta,1]}(u)$, for $u \in [0,1]$, and the isotonic Projection for $\gamma_{1,\beta} + \lambda F^{-1}$ for some $\lambda \geq 0$ as

$$h^{\uparrow}_{1,\beta,\lambda} = \arg\min_{h \in \mathcal{K}} ||h - \gamma_{1,\beta} - \lambda F^{-1}||_2.$$

- Denote $\gamma_{2,\beta}(u) := \gamma(u)\,\mathbb{1}_{[0,\beta]}(u)$, for $u \in [0,1]$, and the isotonic Projection for $\gamma_{2,\beta} + \lambda F^{-1}$ for some $\lambda \geq 0$ as

$$h^{\uparrow}_{2,\beta,\lambda} = \arg\min_{h \in \mathcal{K}} ||h - \gamma_{2,\beta} - \lambda F^{-1}||_2.$$

Wasserstein distance plus moments constraints and stop-loss transform

Theorem (Cai et al. (2022a))

Consider the worst-case problem $\sup_{G\in\mathcal{S}} \rho_h\left((Y^G - d)_+\right)$.
The quantile function of the worst-case distribution is

$$G_{\beta^*}^{-1}(u) = \mu + \sigma\left(\frac{h^{\uparrow}_{1,\beta^*,\lambda}(u) - a_{\beta^*,\lambda}}{b_{\beta^*,\lambda}}\right), \quad 0 < u < 1,$$

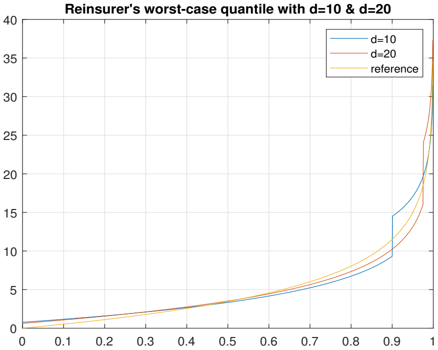
where $a_{\beta^*,\lambda} = \mathbb{E}[h^{\uparrow}_{1,\beta^*,\lambda}(U)]$, $b_{\beta^*,\lambda} = \sqrt{\mathrm{var}(h^{\uparrow}_{1,\beta^*,\lambda}(U))}$, $\lambda > 0$ *is determined uniquely by the distance constraint* $W_2(F, G_{\beta^*}) = \varepsilon$, *and*

$$\beta^* = \underset{\beta\in[0,1]}{\arg\max} \int_0^1 \gamma_{1,\beta}(u)\left(G_\beta^{-1}(u) - d\right)\mathrm{d}u.$$

Example – Expected shortfall

Assume the reference distribution is $F(x) = 1 - e^{-x/5}$, $\mu = \sigma = 5$, $\varepsilon = 1$, and $\rho_h = \mathrm{ES}_{0.9}$.

d	β^*
10	$[0, 0.9]$
20	0.9751



Wasserstein distance plus moments constraints and limited-loss transform

Theorem (Cai et al. (2022a))

Consider the worst-case problem $\sup_{G\in\mathcal{S}} \rho_h\left(Y^G \wedge M\right)$.
The quantile function of the worst-case distribution is

$$F_{\beta^*}^{-1}(u) = \mu + \sigma\left(\frac{h^{\uparrow}_{2,\beta^*,\lambda}(u) - a_{\beta^*,\lambda}}{b_{\beta^*,\lambda}}\right), \quad 0 < u < 1,$$

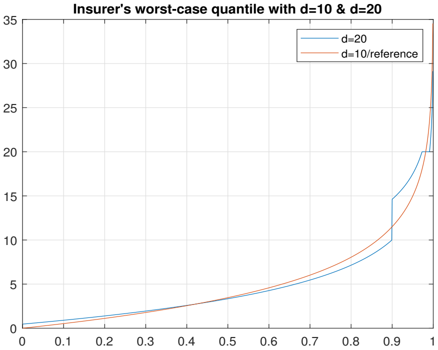
where $a_{\beta^*,\lambda} = \mathbb{E}[h^{\uparrow}_{2,\beta^*,\lambda}(U)]$, $b_{\beta^*,\lambda} = \sqrt{\mathrm{var}(h^{\uparrow}_{2,\beta^*,\lambda}(U))}$, $\lambda > 0$ *is determined uniquely by the distance constraint* $W_2(F, F_{\beta^*}) = \varepsilon$, *and*

$$\beta^* = \underset{\beta\in[0,1]}{\arg\min} \int_0^\beta \gamma_{2,\beta}(u)\left(F_\beta^{-1}(u) - d\right)\mathrm{d}u.$$

Example – Expected shortfall

Assume the reference distribution is $F(x) = 1 - e^{-x/5}$, $\mu = \sigma = 5$, $\varepsilon = 1$, and $\rho_h = \mathrm{ES}_{0.9}$.

d	β^*
10	$[0, 0.9]$
20	0.9835



Summary

In this talk we discuss multiple model uncertainty models

- Distortion risk measure
- With or without transform
  - Stop-loss, limited-loss
- Wasserstein distance, moments contraints

Future works

- Other risk measures
- General transformation
- Various uncertainty sets: likelihood ratio, KL-divergent, etc.
- Novel techniques to characterize worst-case distribution and worst-case risk measure value

Reference I

Liu, F., Cai, J., Lemieux, C. and Wang, R. (2020). Convex risk functionals: Representation and applications. *Insurance: Mathematics and Economics*, 90, 66–79.

Liu, F., Mao, T., Wei, L. and Wang, R. (2022). Inf-Convolution, Optimal Allocations, and Model Uncertainty for Tail Risk Measures. *Mathematics of Operations Research*, 47(3), 1707–2545.

Cai, J., Liu, F. and Yin, M. (2022). Worst-case risk measures of stop-loss and limited loss random variables under distribution uncertainty, *working paper*.

Cai, J., Liu, F. and Yin, M. (2022). Worst-case risk measures of limited stop-loss transformation under Wasserstein-type uncertainty, *working paper*.

Asimit, A. V., Bignozzi, V., Cheung, K. C., Hu, J., and Kim, E.-S. (2017). Robust and Pareto optimality of insurance contracts. *European Journal of Operational Research*, 262(2):720–732.

Bernard, C., Pesenti, S. M., and Vanduffel, S. (2020b). Robust distortion risk measures. *Available at SSRN*

Corina Birghila and Georg Ch. Pflug. (2019). Optimal XL-insurance under Wasserstein-type ambiguity. *Insurance: Mathematics and Economics*, 88: 30-43.

Cai, J. and Tan, K. (2007). Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. *ASTIN Bulletin*, 37(1):93–112.

Cai, J., Li, J., and Mao, T. (2020). Distributional robust optimization under distorted expectations. *Available at SSRN 3566708*

Chen L, He S, Zhang S (2011) Tight bounds for some risk measures, with applications to robust portfolio selection. Oper. Res. 59(4):847–865.

Reference II

Dhaene, J., Kukush, A., Linders, D., and Tang, Q. (2012). Remarks on quantiles and distortion risk measures. *European Actuarial Journal*, 2(2):319–328.

El Ghaoui L, Oks M, Oustry F (2003) Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. Oper. Res. 51(4):543–556.

Li, J. Y.-M. (2018). Closed-form solutions for worst-case law invariant risk measures with application to robust portfolio optimization. *Operations Research*, 66(6):1533–1541.

Liu, H. and Mao, T. (2021). Distributional robust reinsurance with value-at-risk and conditional value-at-risk. *Available at SSRN 3849078*

Liu, F. and Wang, R. (2021). A theory for measures of tail risk. *Mathematics of Operations Research*, 46(3), 1109–1128.

Hu, X., Yang, H., and Zhang, L. (2015). Optimal retention for a stop-loss reinsurance with incomplete information. *Insurance: Mathematics and Economics*, 65:15–21.

Sion, M. (1958). On general minimax theorems, *Pacific Journal of mathematics*, 8(1), 171–176.