Risk Averse Stochastic Programming

Shabbir Ahmed
sahmed@isye.gatech.edu

Outline

- Classical stochastic programming
- Modeling issues
  - Risk aversion
  - Distribution robustness
- Algorithmic issues
  - Sampling
  - Optimization

Based on work with J. Luedtke, A. Shapiro, and W. Wang.

Classical Stochastic Programming

Stochastic Programming

$$SP: \quad \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x,\xi)] \right\}$$

- x is the decision vector,
- X is the set of feasible solutions,
- ξ is a random vector with known distribution P,
- F is a "cost" function, and
- we want to minimize expected cost.

Examples

- Newsvendor
  - x: order quanity; ξ: demand
  - $F(x, \xi) := q_+(x - \xi)_+ + q_-(\xi - x)_+$
  - $X := \mathbb{R}_+$.
- Portfolio selection
  - x: investment proportions; ξ: asset returns
  - $F(x, \xi)$: portfolio loss function, e.g. $F(x, \xi) := -\xi^\top x$
  - $X := \{x \in \mathbb{R}^n : \ e^\top x = 1, \ x \geq 0\}$
- Two-stage stochastic programs
  - x: first stage decisions; ξ: uncertain parameters; y: second stage decisions
  - $F(x, \xi := (q, h, T)) = \min\{q^\top y : \ Wy \geq h - Tx\}$
  - $X := \{x \in \mathbb{R}^n : \ Ax \geq b\}$

Examples

- Newsvendor
  - x: order quanity; ξ: demand
  - $F(x,\xi) := q_+(x-\xi)_+ + q_-(\xi - x)_+$
  - $X := \mathbb{R}_+$.
- Portfolio selection
  - x: investment proportions; ξ: asset returns
  - $F(x,\xi)$: portfolio loss function, e.g. $F(x,\xi) := -\xi^\top x$
  - $X := \{x \in \mathbb{R}^n : \ e^\top x = 1, \ x \geq 0\}$
- Two-stage stochastic programs
  - x: first stage decisions; ξ: uncertain parameters; y: second stage decisions
  - $F(x, \xi := (q,h,T)) = \min\{q^\top y : \ Wy \geq h - Tx\}$
  - $X := \{x \in \mathbb{R}^n : \ Ax \geq b\}$

Examples

- Newsvendor
  - x: order quanity; ξ: demand
  - $F(x, \xi) := q_+(x - \xi)_+ + q_-(\xi - x)_+$
  - $X := \mathbb{R}_+$.

- Portfolio selection
  - x: investment proportions; ξ: asset returns
  - $F(x, \xi)$: portfolio loss function, e.g. $F(x, \xi) := -\xi^\top x$
  - $X := \{x \in \mathbb{R}^n : \ e^\top x = 1, \ x \geq 0\}$

- Two-stage stochastic programs
  - x: first stage decisions; ξ: uncertain parameters; y: second stage decisions
  - $F(x, \xi := (q, h, T)) = \min\{q^\top y : \ Wy \geq h - Tx\}$
  - $X := \{x \in \mathbb{R}^n : \ Ax \geq b\}$

Examples

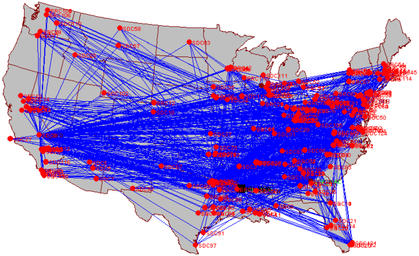
- Newsvendor
  - x: order quanity; ξ: demand
  - $F(x,\xi) := q_+(x-\xi)_+ + q_-(\xi - x)_+$
  - $X := \mathbb{R}_+$.

- Portfolio selection
  - x: investment proportions; ξ: asset returns
  - $F(x,\xi)$: portfolio loss function, e.g. $F(x,\xi) := -\xi^\top x$
  - $X := \{x \in \mathbb{R}^n : \; e^\top x = 1, \; x \geq 0\}$

- Two-stage stochastic programs
  - x: first stage decisions; ξ: uncertain parameters; y: second stage decisions
  - $F(x, \xi := (q,h,T)) = \min\{q^\top y : \; Wy \geq h - Tx\}$
  - $X := \{x \in \mathbb{R}^n : \; Ax \geq b\}$

2S-SP Example: Supply Chain Network Design

- Strategic decisions: Locate DCs and warehouses
- Operational decisions: Shipments through the network to satisfy customer demands
- Locate → observe demand → ship.



Issues

$$SP: \quad \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x,\xi)] \right\}$$

- (Typically) evaluating $f(\hat{x}) = \mathbb{E}_P[F(\hat{x},\xi)]$ exactly is impossible.
- Large-scale optimization problem.

Stochastic Programming $\approx$ Sampling + (Deterministic) Optimization

Issues

$$SP: \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x,\xi)] \right\}$$

- (Typically) evaluating $f(\hat{x}) = \mathbb{E}_P[F(\hat{x},\xi)]$ exactly is impossible.
- Large-scale optimization problem.

Stochastic Programming $\approx$ Sampling + (Deterministic) Optimization

Issues

$$SP: \min_{x \in X} \{ f(x) := \mathbb{E}_P[F(x,\xi)] \}$$

- (Typically) evaluating $f(\hat{x}) = \mathbb{E}_P[F(\hat{x},\xi)]$ exactly is impossible.
- Large-scale optimization problem.

Stochastic Programming $\approx$ Sampling + (Deterministic) Optimization

Issues

$$SP: \quad \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x,\xi)] \right\}$$

- (Typically) evaluating $f(\hat{x}) = \mathbb{E}_P[F(\hat{x},\xi)]$ exactly is impossible.
- Large-scale optimization problem.

Stochastic Programming $\approx$ Sampling + (Deterministic) Optimization

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \ldots, \xi^N)$ from P.
- Solve

$$SAA_N : \min_{x \in X} \left\{ f_N(x) := N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).
- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?
- How to solve SAA_N?

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \ldots, \xi^N)$ from P.
- Solve

$$SAA_N: \quad \min_{x \in X} \left\{ f_N(x) := N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).
- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?
- How to solve SAA_N?

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \dots, \xi^N)$ from P.
- Solve

$$SAA_N: \quad \min_{x \in X} \left\{ f_N(x) := N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).
- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?
- How to solve SAA_N?

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \ldots, \xi^N)$ from P.
- Solve

$$SAA_N : \quad \min_{x \in X} \left\{ f_N(x) := \; N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).
- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?
- How to solve SAA_N?

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \ldots, \xi^N)$ from P.
- Solve

$$SAA_N: \quad \min_{x \in X} \left\{ f_N(x) := N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).
- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?
- How to solve SAA_N?

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \ldots, \xi^N)$ from P.
- Solve

$$SAA_N: \quad \min_{x \in X} \left\{ f_N(x) := N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).
- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?
- How to solve SAA_N?

Sample Average Approximation

- Generate i.i.d sample $(\xi^1, \ldots, \xi^N)$ from P.
- Solve

$$SAA_N : \quad \min_{x \in X} \left\{ f_N(x) := \; N^{-1} \sum_{i=1}^{N} [F(x, \xi^i)] \right\}$$

- Let v_N be the optimal value and X_N be the set of optimal solutions.
- Let v^* and X^* be the optimal value and optimal solution set for SP (assume these exist).

- What is the relation between v_N and v^* and between X_N and X^* w.r.t sample size N?

- How to solve SAA_N?

Convergence

Theorem (Shapiro, Wets, Birge etc.)

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^*
... exponentially fast!

- Implication: For problem with n variables, with

$$N = O\left(\frac{n}{\epsilon^2}\right)$$

  an optimal solution to SAA_N is an ϵ-optimal solution to SP with very high probability.

- SAA_N also serves in a simple statistical procedure to validate the quality of a candidate solution.

Convergence

Theorem (Shapiro, Wets, Birge etc.)

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^*
... exponentially fast!

- Implication: For problem with n variables, with

$$N = O\left(\frac{n}{\epsilon^2}\right)$$

  an optimal solution to SAA_N is an ϵ-optimal solution to SP with very high probability.

- SAA_N also serves in a simple statistical procedure to validate the quality of a candidate solution.

Convergence

Theorem (Shapiro, Wets, Birge etc.)

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^* ... exponentially fast!

- Implication: For problem with n variables, with

$$N = O\left(\frac{n}{\epsilon^2}\right)$$

  an optimal solution to SAA_N is an ϵ-optimal solution to SP with very high probability.

- SAA_N also serves in a simple statistical procedure to validate the quality of a candidate solution.

Convergence

Theorem (Shapiro, Wets, Birge etc.)

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^* ... exponentially fast!

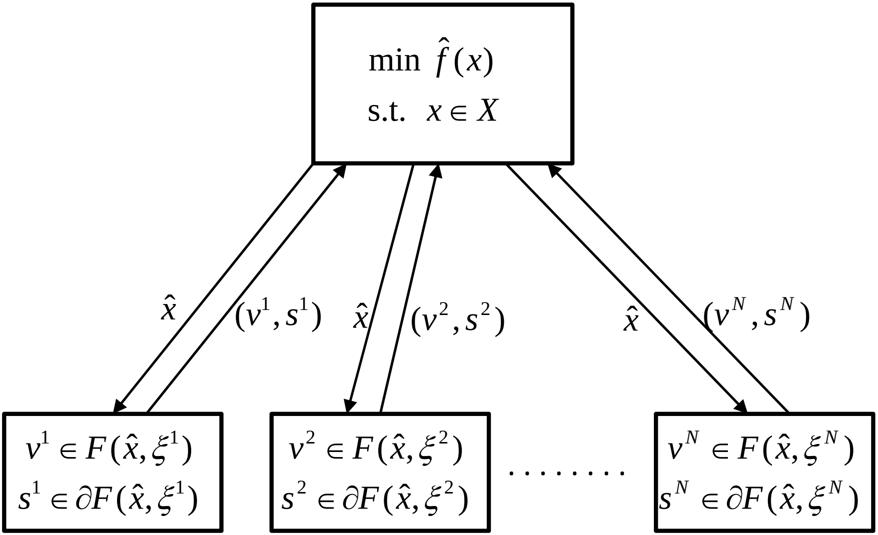
- Implication: For problem with n variables, with

$$N = O\left(\frac{n}{\epsilon^2}\right)$$

  an optimal solution to SAA_N is an ϵ-optimal solution to SP with very high probability.

- SAA_N also serves in a simple statistical procedure to validate the quality of a candidate solution.

Optimization via Decomposition

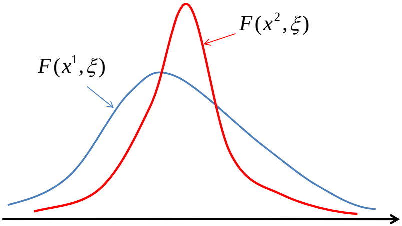


Risk Aversion

Risk Aversion

$$SP: \quad \min_{x \in X} \{f(x) := \mathbb{E}[F(x, \xi)]\}$$

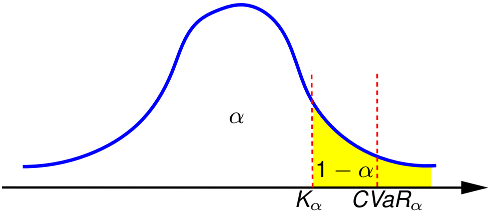
- Why *expected* costs?
- Given x, the objective (cost) is a random variable $F(x, \xi)$.



- We need a scalar measure to compare solutions x^1 and x^2.

Risk functions ρ

- A risk function is a mapping that assigns a real number $\rho[Z]$ to a random variable Z.
- Examples:
  - Expected value: $\rho[Z] = \mathbb{E}[Z]$
  - Expected (dis)utility: $\rho[Z] = \mathbb{E}[u(Z)]$
  - Mean-Variance: $\rho[Z] = \mathbb{E}[Z] + \lambda\mathbb{V}[Z]$
  - α-quantile or α-VaR: $\rho[Z] = K_\alpha[Z] = \min\{t : \Pr(Z \leq t) \geq \alpha\}$
  - α-Conditional-VaR: $\rho[Z] = \text{CVaR}_\alpha = \mathbb{E}[Z|Z \geq K_\alpha[Z]]$



Risk Averse SP

$$RASP: \quad \min_{x \in X} \left\{ f(x) := \ \rho[F(x,\xi)] \right\}$$

- Choice of ρ is (mostly) a modeling issue.
- How to solve *RASP*? (Sampling + Optimization)
- Expected (dis)utility $\rightarrow$ straight-forward (Rutenberg '73).
- Utility function hard to elicit $\rightarrow$ Dispersion statistics preferred.

Risk Averse SP

$$\boxed{RASP: \quad \min_{x \in X} \left\{ f(x) := \ \rho[F(x,\xi)] \right\}}$$

- Choice of ρ is (mostly) a modeling issue.
- How to solve *RASP*? (Sampling + Optimization)
- Expected (dis)utility $\rightarrow$ straight-forward (Rutenberg '73).
- Utility function hard to elicit $\rightarrow$ Dispersion statistics preferred.

Risk Averse SP

$$\boxed{RASP: \quad \min_{x \in X} \left\{ f(x) := \ \rho[F(x,\xi)] \right\}}$$

- Choice of ρ is (mostly) a modeling issue.
- How to solve *RASP*? (Sampling + Optimization)
- Expected (dis)utility $\rightarrow$ straight-forward (Rutenberg '73).
- Utility function hard to elicit $\rightarrow$ Dispersion statistics preferred.

Risk Averse SP

$$\boxed{RASP: \quad \min_{x\in X} \left\{ f(x) := \ \rho[F(x,\xi)] \right\}}$$

- Choice of ρ is (mostly) a modeling issue.
- How to solve *RASP*? (Sampling + Optimization)
- Expected (dis)utility $\rightarrow$ straight-forward (Rutenberg '73).
- Utility function hard to elicit $\rightarrow$ Dispersion statistics preferred.

Risk Averse SP

$$\boxed{RASP: \quad \min_{x \in X} \left\{ f(x) := \ \rho[F(x,\xi)] \right\}}$$

- Choice of ρ is (mostly) a modeling issue.
- How to solve *RASP*? (Sampling + Optimization)
- Expected (dis)utility $\rightarrow$ straight-forward (Rutenberg '73).
- Utility function hard to elicit $\rightarrow$ Dispersion statistics preferred.

Mean-Risk Optimization

- We will consider risk functions of the form

$$\rho[Z] = \gamma\mathbb{E}[Z] + \lambda\mathbb{D}[Z]$$

  where $\mathbb{D}$ is a dispersion measure, and γ and λ are trade-off weights.
- Can analyze risk return tradeoff.
- What $\mathbb{D}$, γ, and λ "makes sense"?
- How do we solve (sampling + optimization) the corresponding stochastic programs?

Mean-Risk Optimization

- We will consider risk functions of the form

$$\rho[Z] = \gamma \mathbb{E}[Z] + \lambda \mathbb{D}[Z]$$

  where $\mathbb{D}$ is a dispersion measure, and γ and λ are trade-off weights.
- Can analyze risk return tradeoff.
- What $\mathbb{D}$, γ, and λ "makes sense"?
- How do we solve (sampling + optimization) the corresponding stochastic programs?

Mean-Risk Optimization

- We will consider risk functions of the form

$$\rho[Z] = \gamma\mathbb{E}[Z] + \lambda\mathbb{D}[Z]$$

  where $\mathbb{D}$ is a dispersion measure, and γ and λ are trade-off weights.
- Can analyze risk return tradeoff.
- What $\mathbb{D}$, γ, and λ "makes sense"?
- How do we solve (sampling + optimization) the corresponding stochastic programs?

Mean-Risk Optimization

- We will consider risk functions of the form

$$\rho[Z] = \gamma\mathbb{E}[Z] + \lambda\mathbb{D}[Z]$$

  where $\mathbb{D}$ is a dispersion measure, and γ and λ are trade-off weights.
- Can analyze risk return tradeoff.
- What $\mathbb{D}$, γ, and λ "makes sense"?
- How do we solve (sampling + optimization) the corresponding stochastic programs?

Mean-Variance

- Markowitz approach: $\rho[Z] = \mathbb{E}[Z] + \lambda\mathbb{V}[Z]$.
- $\mathbb{V}[F(x,\xi)]$ is often non-convex.
- Provably NP-hard.
- Sampling theory hard to extend.

Mean-Variance

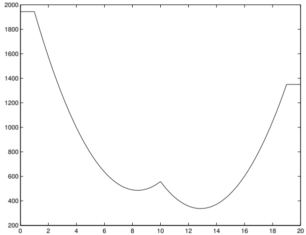
- Markowitz approach:
  $\rho[Z] = \mathbb{E}[Z] + \lambda\mathbb{V}[Z]$.
- $\mathbb{V}[F(x, \xi)]$ is often non-convex.
- Provably NP-hard.
- Sampling theory hard to extend.

Mean-Variance

- Markowitz approach: $\rho[Z] = \mathbb{E}[Z] + \lambda\mathbb{V}[Z]$.
- $\mathbb{V}[F(x,\xi)]$ is often non-convex.
- Provably NP-hard.
- Sampling theory hard to extend.

Mean-Variance

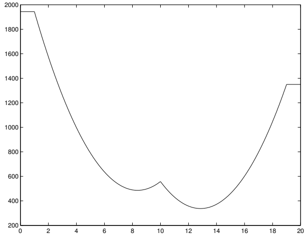
- Markowitz approach: $\rho[Z] = \mathbb{E}[Z] + \lambda\mathbb{V}[Z]$.
- $\mathbb{V}[F(x,\xi)]$ is often non-convex.
- Provably NP-hard.
- Sampling theory hard to extend.



$$F(x,\xi) = q_+(x-\xi)_+ + q_-(\xi - x)_+$$

Mean-Variance

- Markowitz approach: $\rho[Z] = \mathbb{E}[Z] + \lambda\mathbb{V}[Z]$.
- $\mathbb{V}[F(x,\xi)]$ is often non-convex.
- Provably NP-hard.
- Sampling theory hard to extend.



$$F(x,\xi) = q_+(x-\xi)_+ + q_-(\xi - x)_+$$

- VaR Stochastic Programming:

$$\min\{f(x) := K_\alpha[F(x,\xi)] : \ x \in X\}.$$

- Equivalent to a chance-constrained stochastic program:

$$\min\{t : \ \ \Pr[F(x,\xi) - t \leq 0] \geq \alpha, \ x \in X\}$$

- Non-convex even in the linear setting $F(x,\xi) = -\xi^\top x$.
- Sampling theory extended (Luedtke + A.'08)
- SAA_N is NP-hard, solve via integer programming

Value-at-Risk Optimization

- VaR Stochastic Programming:

$$\min\{f(x) := K_\alpha[F(x,\xi)] : \; x \in X\}.$$

- Equivalent to a chance-constrained stochastic program:

$$\min\{t : \;\; \Pr[F(x,\xi) - t \leq 0] \geq \alpha, \; x \in X\}$$

- Non-convex even in the linear setting $F(x,\xi) = -\xi^\top x$.
- Sampling theory extended (Luedtke + A.'08)
- SAA_N is NP-hard, solve via integer programming

Value-at-Risk Optimization

- VaR Stochastic Programming:

$$\min\{f(x) := K_\alpha[F(x,\xi)] : \; x \in X\}.$$

- Equivalent to a chance-constrained stochastic program:

$$\min\{t : \; \Pr[F(x,\xi) - t \leq 0] \geq \alpha, \; x \in X\}$$

- Non-convex even in the linear setting $F(x,\xi) = -\xi^\top x$.
- Sampling theory extended (Luedtke + A.'08)
- SAA_N is NP-hard, solve via integer programming

Value-at-Risk Optimization

- VaR Stochastic Programming:

$$\min\{f(x) := K_\alpha[F(x,\xi)] : \; x \in X\}.$$

- Equivalent to a chance-constrained stochastic program:

$$\min\{t : \;\; \Pr[F(x,\xi) - t \leq 0] \geq \alpha, \; x \in X\}$$

- Non-convex even in the linear setting $F(x,\xi) = -\xi^\top x$.
- Sampling theory extended (Luedtke + A.'08)
- SAA_N is NP-hard, solve via integer programming

Value-at-Risk Optimization

- VaR Stochastic Programming:

$$\min\{f(x) := K_\alpha[F(x,\xi)] : \ x \in X\}.$$

- Equivalent to a chance-constrained stochastic program:

$$\min\{t : \ \Pr[F(x,\xi) - t \leq 0] \geq \alpha, \ x \in X\}$$

- Non-convex even in the linear setting $F(x,\xi) = -\xi^\top x$.
- Sampling theory extended (Luedtke + A.'08)
- SAA_N is NP-hard, solve via integer programming

Distribution Robustness

Distribution Robustness

$$SP: \quad \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x,\xi)] \right\}$$

- *SP* assumes precise knowledge of underlying distribution *P*.
- In the SAA framework, a sampling oracle for *P* is needed.
- How to handle imprecision of the distribution?

Distribution Robustness

$$SP: \quad \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x,\xi)] \right\}$$

- *SP* assumes precise knowledge of underlying distribution P.
- In the SAA framework, a sampling oracle for P is needed.
- How to handle imprecision of the distribution?

Distribution Robustness

$$SP: \quad \min_{x \in X} \left\{ f(x) := \mathbb{E}_P[F(x, \xi)] \right\}$$

- *SP* assumes precise knowledge of underlying distribution P.
- In the SAA framework, a sampling oracle for P is needed.
- How to handle imprecision of the distribution?

Distribution Families

- Instead of knowing P exactly, suppose we only know a family $\mathbb{P}$ of likely distributions.
- E.g. We have estimated a nominal distribution P^0 and to account for estimation errors we can consider

$$\mathbb{P} = \{P : \ (1-\epsilon_1)P^0 \preceq P \preceq (1+\epsilon_2)P^0, \ \mathbb{E}_P[1] = 1\},$$

where $0 \leq \epsilon_1 < 1$ and $0 \leq \epsilon_2$.

Distribution Robust SP

$$DRSP: \quad \min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x, \xi)]$$

- Convex problem in x, but evaluation may be difficult.
- Sampling theory does not immediately extend.
- Q. Can we extend classical SP methodology (sampling and optimization) to *DRSP*?
- A. Yes (for some $\mathbb{P}$) but indirectly.

Distribution Robust SP

$$\boxed{DRSP: \quad \min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x,\xi)]}$$

- Convex problem in x, but evaluation may be difficult.
- Sampling theory does not immediately extend.
- Q. Can we extend classical SP methodology (sampling and optimization) to *DRSP*?
- A. Yes (for some $\mathbb{P}$) but indirectly.

Distribution Robust SP

$$DRSP: \quad \min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x, \xi)]$$

- Convex problem in x, but evaluation may be difficult.
- Sampling theory does not immediately extend.
- Q. Can we extend classical SP methodology (sampling and optimization) to *DRSP*?
- A. Yes (for some $\mathbb{P}$) but indirectly.

Distribution Robust SP

$$\boxed{DRSP: \quad \min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x,\xi)]}$$

- Convex problem in x, but evaluation may be difficult.
- Sampling theory does not immediately extend.
- Q. Can we extend classical SP methodology (sampling and optimization) to *DRSP*?
- A. Yes (for some $\mathbb{P}$) but indirectly.

Distribution Robust SP

$$DRSP: \quad \min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x, \xi)]$$

- Convex problem in x, but evaluation may be difficult.
- Sampling theory does not immediately extend.
- Q. Can we extend classical SP methodology (sampling and optimization) to *DRSP*?
- A. Yes (for some $\mathbb{P}$) but indirectly.

Distribution robustness and risk aversion

Theorem (Artzner et al.'99)

If $\mathbb{P}$ is a closed convex family of distributions then there exists a (convex + ...) risk function ρ such that

$$\max_{P \in \mathbb{P}} \mathbb{E}_P[Z] = \rho[Z],$$

and vice versa.

- Follows from conjugate duality.
- The associated risk functions are "consistent" with rational choice (e.g. stochastic ordering, risk aversion, coherence etc.).

Distribution robustness and risk aversion

Theorem (Artzner et al.'99)

If $\mathbb{P}$ is a closed convex family of distributions then there exists a (convex + ...) risk function ρ such that

$$\max_{P \in \mathbb{P}} \mathbb{E}_P[Z] = \rho[Z],$$

and vice versa.

- Follows from conjugate duality.
- The associated risk functions are "consistent" with rational choice (e.g. stochastic ordering, risk aversion, coherence etc.).

Distribution robustness and risk aversion

Theorem (Artzner et al.'99)

If $\mathbb{P}$ is a closed convex family of distributions then there exists a (convex + ...) risk function ρ such that

$$\max_{P \in \mathbb{P}} \mathbb{E}_P[Z] = \rho[Z],$$

and vice versa.

- Follows from conjugate duality.
- The associated risk functions are "consistent" with rational choice (e.g. stochastic ordering, risk aversion, coherence etc.).

Distribution robustness and risk aversion

Theorem (Artzner et al.'99)

If $\mathbb{P}$ is a closed convex family of distributions then there exists a (convex + ...) risk function ρ such that

$$\max_{P \in \mathbb{P}} \mathbb{E}_P[Z] = \rho[Z],$$

and vice versa.

- Follows from conjugate duality.
- The associated risk functions are "consistent" with rational choice (e.g. stochastic ordering, risk aversion, coherence etc.).

Example

Consider a setting with K scenarios:

$$\begin{aligned}
&\max_{p}\left\{\sum_{k=1}^{K} p_k F(x,\xi^k) : \sum_{k=1}^{K} p_k = 1,\ 0 \le p_k \le (1+\epsilon)p_k^0\right\}\\
=\ &\min_{\lambda}\max_{p}\left\{\lambda + \sum_{k=1}^{K} p_k (F(x,\xi^k)-\lambda) : \ 0 \le p_k \le (1+\epsilon)p_k^0\right\}\\
=\ &\min_{\lambda}\left\{\lambda + (1+\epsilon)\sum_{k=1}^{K} p_k^0 (F(x,\xi^k)-\lambda)_+\right\}\\
=\ &\min_{\lambda}\left\{\lambda + (1+\epsilon)\mathbb{E}(F(x,\xi)-\lambda)_+\right\}\\
=\ &\mathrm{CVaR}[F(x,\xi)]
\end{aligned}$$

Example

Consider a setting with K scenarios:

$$\max_p \left\{ \sum_{k=1}^{K} p_k F(x, \xi^k) : \sum_{k=1}^{K} p_k = 1,\ 0 \leq p_k \leq (1+\epsilon) p_k^0 \right\}$$

$$= \min_\lambda \max_p \left\{ \lambda + \sum_{k=1}^{K} p_k (F(x,\xi^k) - \lambda) :\ 0 \leq p_k \leq (1+\epsilon) p_k^0 \right\}$$

$$= \min_\lambda \left\{ \lambda + (1+\epsilon) \sum_{k=1}^{K} p_k^0 (F(x,\xi^k) - \lambda)_+ \right\}$$

$$= \min_\lambda \{ \lambda + (1+\epsilon) \mathbb{E}(F(x,\xi) - \lambda)_+ \}$$

$$= \mathrm{CVaR}[F(x,\xi)]$$

Example

Consider a setting with K scenarios:

$$\max_p \left\{ \sum_{k=1}^K p_k F(x, \xi^k) : \sum_{k=1}^K p_k = 1,\ 0 \le p_k \le (1+\epsilon) p_k^0 \right\}$$

$$= \min_\lambda \max_p \left\{ \lambda + \sum_{k=1}^K p_k (F(x, \xi^k) - \lambda) : \ 0 \le p_k \le (1+\epsilon) p_k^0 \right\}$$

$$= \min_\lambda \left\{ \lambda + (1+\epsilon) \sum_{k=1}^K p_k^0 (F(x, \xi^k) - \lambda)_+ \right\}$$

$$= \min_\lambda \{ \lambda + (1+\epsilon) \mathbb{E}(F(x, \xi) - \lambda)_+ \}$$

$$= \mathrm{CVaR}[F(x, \xi)]$$

Example

Consider a setting with K scenarios:

$$\max_p \left\{ \sum_{k=1}^K p_k F(x,\xi^k) : \sum_{k=1}^K p_k = 1,\ 0 \le p_k \le (1+\epsilon)p_k^0 \right\}$$

$$= \min_\lambda \max_p \left\{ \lambda + \sum_{k=1}^K p_k (F(x,\xi^k) - \lambda) : 0 \le p_k \le (1+\epsilon)p_k^0 \right\}$$

$$= \min_\lambda \left\{ \lambda + (1+\epsilon) \sum_{k=1}^K p_k^0 (F(x,\xi^k) - \lambda)_+ \right\}$$

$$= \min_\lambda \{\lambda + (1+\epsilon)\mathbb{E}(F(x,\xi) - \lambda)_+\}$$

$$= \mathrm{CVaR}[F(x,\xi)]$$

Example

Consider a setting with K scenarios:

$$\max_p \left\{ \sum_{k=1}^K p_k F(x, \xi^k) : \sum_{k=1}^K p_k = 1, \; 0 \le p_k \le (1+\epsilon) p_k^0 \right\}$$

$$= \min_\lambda \max_p \left\{ \lambda + \sum_{k=1}^K p_k (F(x,\xi^k) - \lambda) : \; 0 \le p_k \le (1+\epsilon) p_k^0 \right\}$$

$$= \min_\lambda \left\{ \lambda + (1+\epsilon) \sum_{k=1}^K p_k^0 (F(x,\xi^k) - \lambda)_+ \right\}$$

$$= \min_\lambda \{ \lambda + (1+\epsilon) \mathbb{E}(F(x,\xi) - \lambda)_+ \}$$

$$= \mathrm{CVaR}[F(x,\xi)]$$

Example

Consider a setting with K scenarios:

$$\begin{aligned}
&\max_p \left\{ \sum_{k=1}^K p_k F(x,\xi^k) : \sum_{k=1}^K p_k = 1,\ 0 \le p_k \le (1+\epsilon)p_k^0 \right\} \\
=\ &\min_\lambda \max_p \left\{ \lambda + \sum_{k=1}^K p_k (F(x,\xi^k) - \lambda) :\ 0 \le p_k \le (1+\epsilon)p_k^0 \right\} \\
=\ &\min_\lambda \left\{ \lambda + (1+\epsilon) \sum_{k=1}^K p_k^0 (F(x,\xi^k) - \lambda)_+ \right\} \\
=\ &\min_\lambda \{\lambda + (1+\epsilon)\mathbb{E}(F(x,\xi) - \lambda)_+\} \\
=\ &\mathrm{CVaR}[F(x,\xi)]
\end{aligned}$$

Equivalence

- When ρ is a convex risk function and $\mathbb{P}$ is a closed convex set,

$$\min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x,\xi)] \Leftrightarrow \min_{x \in X} \rho[F(x,\xi)]$$

- Unified way of treating both distribution robustness and risk aversion.
- We focus on solving the risk model.

Equivalence

- When ρ is a convex risk function and $\mathbb{P}$ is a closed convex set,

$$\min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x,\xi)] \Leftrightarrow \min_{x \in X} \rho[F(x,\xi)]$$

- Unified way of treating both distribution robustness and risk aversion.
- We focus on solving the risk model.

Equivalence

- When ρ is a convex risk function and $\mathbb{P}$ is a closed convex set,

$$\boxed{\min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x,\xi)] \Leftrightarrow \min_{x \in X} \rho[F(x,\xi)]}$$

- Unified way of treating both distribution robustness and risk aversion.
- We focus on solving the risk model.

Equivalence

- When ρ is a convex risk function and $\mathbb{P}$ is a closed convex set,

$$\boxed{\min_{x \in X} \max_{P \in \mathbb{P}} \mathbb{E}_P[F(x,\xi)] \Leftrightarrow \min_{x \in X} \rho[F(x,\xi)]}$$

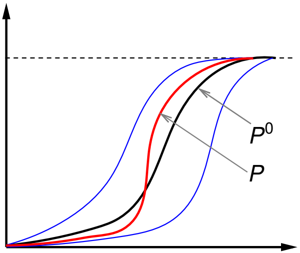
- Unified way of treating both distribution robustness and risk aversion.
- We focus on solving the risk model.

Band distribution family

Given a nominal distribution P^0 consider the "band" distribution family

$$\mathbb{P} = \{P : \ (1 - \epsilon_1)P^0 \preceq P \preceq (1 + \epsilon_2)P^0, \ \mathbb{E}_P[1] = 1\}$$

with $0 \leq \epsilon_1 \leq 1$ and $0 \leq \epsilon_2$.



The Mean-QDEV Risk Function

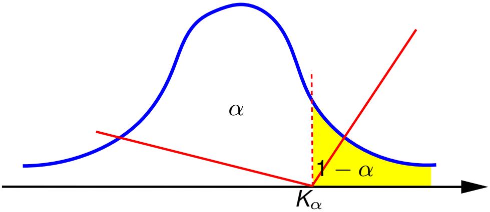
Theorem

The risk function corresponding to the band family is

$$\rho[Z] = \mathbb{E}_{P^0}[Z] + \lambda QDEV_\alpha[Z]$$

where $\lambda = (\epsilon_1 + \epsilon_2)$ *and* $\alpha = \epsilon_2/(\epsilon_1 + \epsilon_2)$*, and*

$$QDEV_\alpha[Z] = \mathbb{E}_{P^0}[\alpha(Z - K_\alpha[Z])_+ + (1-\alpha)(K_\alpha[Z] - Z)_+].$$



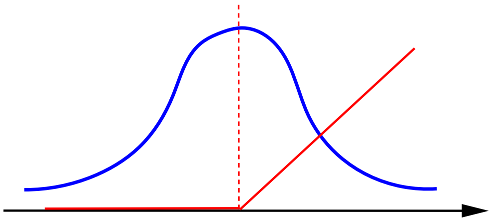
Mean-ASD

- Note that $\epsilon_1 + \epsilon_2$ controls the "width" of the "band" and $\epsilon_2/(\epsilon_1 + \epsilon_2)$ controls its "position."
- If the worst position is allowed then the corresponding coherent risk measure is

$$\boxed{\rho[Z] = \mathbb{E}_{P0}[Z] + \lambda \mathbb{E}_{P0}[(Z - \mathbb{E}Z)_+]}$$

with $\lambda = (\epsilon_1 + \epsilon_2)$.

- $MASD[Z] := \mathbb{E}_{P0}[(Z - \mathbb{E}Z)_+]$.



Two Mean-Risk SP Models

- We consider mean-risk models

$$\min_{x \in X}\{\mathbb{E}[F(x,\xi)] + \lambda \mathbb{D}[F(x,\xi)]\}$$

  where $\mathbb{D}$ is

  (i) $QDEV_\alpha$
  (ii) $MASD$

- Equivalent to RSP models corresponding to the "band" distribution family.
- Can we extend SAA? Can we optimize SAA_N efficiently?

Two Mean-Risk SP Models

- We consider mean-risk models

$$\min_{x \in X} \{\mathbb{E}[F(x,\xi)] + \lambda \mathbb{D}[F(x,\xi)]\}$$

  where $\mathbb{D}$ is

  (i) $QDEV_\alpha$
  (ii) *MASD*

- Equivalent to RSP models corresponding to the "band" distribution family.
- Can we extend SAA? Can we optimize SAA_N efficiently?

Two Mean-Risk SP Models

- We consider mean-risk models

$$\min_{x \in X}\{\mathbb{E}[F(x,\xi)] + \lambda \mathbb{D}[F(x,\xi)]\}$$

  where $\mathbb{D}$ is

  (i) $QDEV_\alpha$

  (ii) $MASD$

- Equivalent to RSP models corresponding to the “band” distribution family.
- Can we extend SAA? Can we optimize SAA_N efficiently?

Two Mean-Risk SP Models

- We consider mean-risk models

$$\min_{x \in X} \{\mathbb{E}[F(x,\xi)] + \lambda \mathbb{D}[F(x,\xi)]\}$$

  where $\mathbb{D}$ is

  (i) $QDEV_\alpha$
  (ii) $MASD$

- Equivalent to RSP models corresponding to the “band” distribution family.
- Can we extend SAA? Can we optimize SAA_N efficiently?

Two Mean-Risk SP Models

- We consider mean-risk models

$$\min_{x \in X} \{\mathbb{E}[F(x,\xi)] + \lambda \mathbb{D}[F(x,\xi)]\}$$

  where $\mathbb{D}$ is

  (i) $QDEV_\alpha$
  (ii) *MASD*

- Equivalent to RSP models corresponding to the "band" distribution family.
- Can we extend SAA? Can we optimize SAA_N efficiently?

Two Mean-Risk SP Models

- We consider mean-risk models

$$\min_{x \in X} \{\mathbb{E}[F(x,\xi)] + \lambda \mathbb{D}[F(x,\xi)]\}$$

where $\mathbb{D}$ is

(i) $QDEV_\alpha$
(ii) *MASD*

- Equivalent to RSP models corresponding to the “band” distribution family.
- Can we extend SAA? Can we optimize SAA_N efficiently?

Mean-QDEV

Theorem

A Mean-QDEV SP is equivalent to a standard (expectation minimization) SP.

$$\boxed{QDEV_\alpha[Z] = \min_{y\in\mathbb{R}} \mathbb{E}[\alpha(Z-y)_+ + (1-\alpha)(y-Z)_+]}$$

$$\min_{x\in X} \mathbb{E}[F(x,\xi)] + \lambda QDEV_\alpha[F(x,\xi)]$$

$$\Updownarrow$$

$$\min_{x\in X, y\in\mathbb{R}} \mathbb{E}[\underbrace{F(x,\xi)] + \lambda\alpha(F(x,\xi)-y)_+ + \lambda(1-\alpha)(y-F(x,\xi))_+}_{\phi(x,y,\xi)}]$$

$$\Updownarrow$$

$$\min_{x\in X, y\in\mathbb{R}} \mathbb{E}[\phi(x,y,\xi)].$$

Mean-QDEV

Theorem

A Mean-QDEV SP is equivalent to a standard (expectation minimization) SP.

$$QDEV_\alpha[Z] = \min_{y \in \mathbb{R}} \mathbb{E}[\alpha(Z - y)_+ + (1-\alpha)(y - Z)_+]$$

$$\min_{x \in X} \mathbb{E}[F(x,\xi)] + \lambda QDEV_\alpha[F(x,\xi)]$$

$$\Updownarrow$$

$$\min_{x \in X, y \in \mathbb{R}} \mathbb{E}[\underbrace{F(x,\xi)] + \lambda\alpha(F(x,\xi) - y)_+ + \lambda(1-\alpha)(y - F(x,\xi))_+}_{\phi(x,y,\xi)}]$$

$$\Updownarrow$$

$$\min_{x \in X, y \in \mathbb{R}} \mathbb{E}[\phi(x,y,\xi)].$$

Mean-QDEV

Theorem

A Mean-QDEV SP is equivalent to a standard (expectation minimization) SP.

$$QDEV_\alpha[Z] = \min_{y\in\mathbb{R}} \mathbb{E}[\alpha(Z-y)_+ + (1-\alpha)(y-Z)_+]$$

$$\min_{x\in X} \mathbb{E}[F(x,\xi)] + \lambda QDEV_\alpha[F(x,\xi)]$$

$$\Updownarrow$$

$$\min_{x\in X, y\in\mathbb{R}} \underbrace{\mathbb{E}[F(x,\xi)] + \lambda\alpha(F(x,\xi)-y)_+ + \lambda(1-\alpha)(y-F(x,\xi))_+]}_{\phi(x,y,\xi)}$$

$$\Updownarrow$$

$$\min_{x\in X, y\in\mathbb{R}} \mathbb{E}[\phi(x,y,\xi)].$$

Mean-QDEV

Theorem

A Mean-QDEV SP is equivalent to a standard (expectation minimization) SP.

$$\boxed{QDEV_\alpha[Z] = \min_{y \in \mathbb{R}} \mathbb{E}[\alpha(Z - y)_+ + (1-\alpha)(y - Z)_+]}$$

$$\min_{x \in X} \; \mathbb{E}[F(x,\xi)] + \lambda QDEV_\alpha[F(x,\xi)]$$

$$\Updownarrow$$

$$\min_{x \in X, y \in \mathbb{R}} \; \mathbb{E}[\underbrace{F(x,\xi)] + \lambda\alpha(F(x,\xi) - y)_+ + \lambda(1-\alpha)(y - F(x,\xi))_+}_{\phi(x,y,\xi)}]$$

$$\Updownarrow$$

$$\min_{x \in X, y \in \mathbb{R}} \; \mathbb{E}[\phi(x,y,\xi)].$$

Solving Mean-QDEV SP

- When X is compact the domain of y is compact.
- Existing SAA analysis and method applies directly (dimension goes up by one).
- When F is linear and X is polyhedral, the corresponding SAA_N problem is a linear program.
- When F is convex (e.g. two-stage stochastic linear programs) and X is polyhedral, a specialized parametric decomposition algorithm has been developed to construct the efficient frontier (Mean vs QDEV).

Solving Mean-QDEV SP

- When X is compact the domain of y is compact.
- Existing SAA analysis and method applies directly (dimension goes up by one).
- When F is linear and X is polyhedral, the corresponding SAA_N problem is a linear program.
- When F is convex (e.g. two-stage stochastic linear programs) and X is polyhedral, a specialized parametric decomposition algorithm has been developed to construct the efficient frontier (Mean vs QDEV).

Solving Mean-QDEV SP

- When X is compact the domain of y is compact.
- Existing SAA analysis and method applies directly (dimension goes up by one).
- When F is linear and X is polyhedral, the corresponding SAA_N problem is a linear program.
- When F is convex (e.g. two-stage stochastic linear programs) and X is polyhedral, a specialized parametric decomposition algorithm has been developed to construct the efficient frontier (Mean vs QDEV).

Solving Mean-QDEV SP

- When X is compact the domain of y is compact.
- Existing SAA analysis and method applies directly (dimension goes up by one).
- When F is linear and X is polyhedral, the corresponding SAA_N problem is a linear program.
- When F is convex (e.g. two-stage stochastic linear programs) and X is polyhedral, a specialized parametric decomposition algorithm has been developed to construct the efficient frontier (Mean vs QDEV).

Solving Mean-QDEV SP

- When X is compact the domain of y is compact.
- Existing SAA analysis and method applies directly (dimension goes up by one).
- When F is linear and X is polyhedral, the corresponding SAA_N problem is a linear program.
- When F is convex (e.g. two-stage stochastic linear programs) and X is polyhedral, a specialized parametric decomposition algorithm has been developed to construct the efficient frontier (Mean vs QDEV).

Mean-MASD SP

Theorem

A Mean-MASD SP is equivalent to a minimax (expectation minimization) SP.

$$\min_{x \in X} \{\mathbb{E}[F(x,\xi)] + \lambda MASD[F(x,\xi)]\}$$

$$\Updownarrow$$

$$\min_{x \in X} \max_{\alpha \in [0,1]} \{\mathbb{E}[F(x,\xi)] + \lambda QDEV_\alpha[F(x,\xi)]\}$$

Mean-MASD SP

Theorem

A Mean-MASD SP is equivalent to a minimax (expectation minimization) SP.

$$\min_{x \in X} \{\mathbb{E}[F(x,\xi)] + \lambda MASD[F(x,\xi)]\}$$

$$\Updownarrow$$

$$\min_{x \in X} \max_{\alpha \in [0,1]} \{\mathbb{E}[F(x,\xi)] + \lambda QDEV_\alpha[F(x,\xi)]\}$$

$$\Updownarrow$$

$$\min_{x \in X} \max_{\alpha \in [0,1]} \min_{y \in \mathbb{R}} \left\{ \mathbb{E}[\underbrace{F(x,\xi) + \lambda\alpha(F(x,\xi) - y)_+ + \lambda(1-\alpha)(y - F(x,\xi))_+}_{\psi(x,y,\alpha,\xi)}] \right\}$$

$$\Updownarrow$$

$$\min_{x \in X, y \in \mathbb{R}} \max_{\alpha \in [0,1]} \{\mathbb{E}[\psi(x,y,\alpha,\xi)]\}$$

$$\Updownarrow$$

$$\min_{x \in X} \max_{\alpha \in [0,1]} \min_{y \in \mathbb{R}} \left\{ \mathbb{E}[\underbrace{F(x,\xi) + \lambda\alpha(F(x,\xi) - y)_+ + \lambda(1-\alpha)(y - F(x,\xi))_+}_{\psi(x,y,\alpha,\xi)}] \right\}$$

$$\Updownarrow$$

$$\min_{x \in X, y \in \mathbb{R}} \max_{\alpha \in [0,1]} \{\mathbb{E}[\psi(x,y,\alpha,\xi)]\}$$

Stochastic Minimax Problems

- The mean-ASD problem is equivalent to a stochastic minimax problem.
- Existing SAA approaches do not apply directly.
- Decomposition is not obvious.
- Consider general stochastic minimax problems

$$\min_{x \in X} \max_{y \in Y} \{\Psi(x, y) := \mathbb{E}[F(x, y, \xi)]\}$$

where X and Y are compact convex sets, and F is convex in x and concave in y a.e.

Stochastic Minimax Problems

- The mean-ASD problem is equivalent to a stochastic minimax problem.
- Existing SAA approaches do not apply directly.
- Decomposition is not obvious.
- Consider general stochastic minimax problems

$$\min_{x \in X} \max_{y \in Y} \{\Psi(x,y) := \mathbb{E}[F(x,y,\xi)]\}$$

where X and Y are compact convex sets, and F is convex in x and concave in y a.e.

Stochastic Minimax Problems

- The mean-ASD problem is equivalent to a stochastic minimax problem.
- Existing SAA approaches do not apply directly.
- Decomposition is not obvious.
- Consider general stochastic minimax problems

$$\min_{x \in X} \max_{y \in Y} \{\Psi(x, y) := \mathbb{E}[F(x, y, \xi)]\}$$

where X and Y are compact convex sets, and F is convex in x and concave in y a.e.

Stochastic Minimax Problems

- The mean-ASD problem is equivalent to a stochastic minimax problem.
- Existing SAA approaches do not apply directly.
- Decomposition is not obvious.
- Consider general stochastic minimax problems

$$\min_{x \in X} \max_{y \in Y} \{\Psi(x, y) := \mathbb{E}[F(x, y, \xi)]\}$$

where X and Y are compact convex sets, and F is convex in x and concave in y a.e.

Stochastic Minimax Problems

- The mean-ASD problem is equivalent to a stochastic minimax problem.
- Existing SAA approaches do not apply directly.
- Decomposition is not obvious.
- Consider general stochastic minimax problems

$$\min_{x \in X} \max_{y \in Y} \{\Psi(x, y) := \mathbb{E}[F(x, y, \xi)]\}$$

where X and Y are compact convex sets, and F is convex in x and concave in y a.e.

SAA for Stochastic Minimax Problems

$$SAA_N : \quad v_N = \min_{x \in X} \max_{y \in Y} \{N^{-1} \sum_{i=1}^{N} [F(x, y, \xi^i)]\}$$

Theorem

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^* exponentially fast. So with

$$N = O\left(\frac{n_x + n_y}{\epsilon^2}\right)$$

an optimal solution to SAA_N is an ϵ-optimal solution to minmax SP with very high probability.

SAA for Stochastic Minimax Problems

$$SAA_N: \quad v_N = \min_{x \in X} \max_{y \in Y} \{ N^{-1} \sum_{i=1}^{N} [F(x, y, \xi^i)] \}$$

Theorem

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^* exponentially fast. So with

$$N = O\left(\frac{n_x + n_y}{\epsilon^2}\right)$$

an optimal solution to SAA_N is an ϵ-optimal solution to minmax SP with very high probability.

SAA for Stochastic Minimax Problems

$$SAA_N: \quad v_N = \min_{x \in X} \max_{y \in Y} \{N^{-1} \sum_{i=1}^{N} [F(x, y, \xi^i)]\}$$

Theorem

As $N \to \infty$, v_N and X_N converges to their true counterparts v^* and X^* exponentially fast. So with

$$N = O\left(\frac{n_x + n_y}{\epsilon^2}\right)$$

an optimal solution to SAA_N is an ϵ-optimal solution to minmax SP with very high probability.

Solving SAA_N for mean-MASD

- Sharper sample size estimates for Mean-MASD stochastic programs.
- We have developed statistical validation techniques to screen candidate solutions from SAA_N.
- How to decompose?

Solving SAA_N for mean-MASD

- Sharper sample size estimates for Mean-MASD stochastic programs.
- We have developed statistical validation techniques to screen candidate solutions from SAA_N.
- How to decompose?

Solving SAA_N for mean-MASD

- Sharper sample size estimates for Mean-MASD stochastic programs.
- We have developed statistical validation techniques to screen candidate solutions from SAA_N.
- How to decompose?

Solving SAA_N for mean-MASD

- Sharper sample size estimates for Mean-MASD stochastic programs.
- We have developed statistical validation techniques to screen candidate solutions from SAA_N.
- How to decompose?

Decomposition of mean-MASD SP

$$
\begin{aligned}
& \mathbb{E}[F(x,\xi)] + \lambda MASD[F(x,\xi)] \\
= \; & \mathbb{E}[F(x,\xi)] + \lambda \mathbb{E}[F(x,\xi) - \mathbb{E}[F(x,\xi)]]_+ \\
= \; & (1-\lambda)\mathbb{E}[F(x,\xi)] + \lambda \max\{F(x,\xi), \mathbb{E}[F(x,\xi)]\} \\
= \; & (1-\lambda)\mu(x) + \lambda\nu(x)
\end{aligned}
$$

Let $I(\xi) = 1$ if $F(x,\xi) > \mathbb{E}[F(x,\xi)]$ and 0 otherwise. Given $s(\xi) \in \partial F(x,\xi)$ and $\bar{s} = \mathbb{E}[s(\xi)]$, let

$$\hat{s} = \mathbb{E}[I(\xi)s(\xi) + (1 - I(\xi))\bar{s}]$$

then $s \in \partial\nu(x)$.

Thus, evaluation of μ and ν and its subgradients can be done in a decomposed manner.

Decomposition of mean-MASD SP

$$\begin{aligned} & \mathbb{E}[F(x,\xi)] + \lambda MASD[F(x,\xi)] \\ = \; & \mathbb{E}[F(x,\xi)] + \lambda \mathbb{E}[F(x,\xi) - \mathbb{E}[F(x,\xi)]]_+ \\ = \; & (1-\lambda)\mathbb{E}[F(x,\xi)] + \lambda \max\{F(x,\xi), \mathbb{E}[F(x,\xi)]\} \\ = \; & (1-\lambda)\mu(x) + \lambda\nu(x) \end{aligned}$$

Let $I(\xi) = 1$ if $F(x,\xi) > \mathbb{E}[F(x,\xi)]$ and 0 otherwise. Given $s(\xi) \in \partial F(x,\xi)$ and $\overline{s} = \mathbb{E}[s(\xi)]$, let

$$\hat{s} = \mathbb{E}[I(\xi)s(\xi) + (1 - I(\xi))\overline{s}]$$

then $s \in \partial\nu(x)$.

Thus, evaluation of μ and ν and its subgradients can be done in a decomposed manner.

Decomposition of mean-MASD SP

$$
\begin{aligned}
& \mathbb{E}[F(x,\xi)] + \lambda MASD[F(x,\xi)] \\
= \; & \mathbb{E}[F(x,\xi)] + \lambda\mathbb{E}[F(x,\xi) - \mathbb{E}[F(x,\xi)]]_+ \\
= \; & (1-\lambda)\mathbb{E}[F(x,\xi)] + \lambda \max\{F(x,\xi), \mathbb{E}[F(x,\xi)]\} \\
= \; & (1-\lambda)\mu(x) + \lambda\nu(x)
\end{aligned}
$$

Let $I(\xi) = 1$ if $F(x,\xi) > \mathbb{E}[F(x,\xi)]$ and 0 otherwise. Given $s(\xi) \in \partial F(x,\xi)$ and $\overline{s} = \mathbb{E}[s(\xi)]$, let

$$\hat{s} = \mathbb{E}[I(\xi)s(\xi) + (1 - I(\xi))\overline{s}]$$

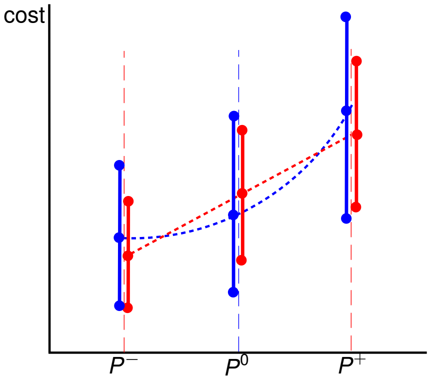
then $s \in \partial\nu(x)$.

Thus, evaluation of μ and ν and its subgradients can be done in a decomposed manner.

Illustration: Inventory Problems

- Considered Distribution robust newsvendor = Risk averse newsendor
- Analytical optimality equations.
- Analysis of sensitivity of order quantity to distributional inaccuracies = risk aversion.
- For multi-period problems optimal policy structure is identical to that of standard expectation minimization problems.

Effect of distribution robustness



Classical SP solution
Robust SP solution

Illustration: Supply Chain Network Design

- A chemical supply chain adapted from Tsiakis, Shah & Pantelides, 2001
- Two-stage problem (Mixed-integer first stage)
- First-stage: Capacity of 6 warehouses and 8 DCs
- Customer demand is uncertain (6 random variables)
- Second-stage: Ship to customers + Outsource penalty
- Minimize (annualized) capacity costs + shipping costs + outsourcing penalty
- Solved Expectation + 0.5*MASD model
- Sample size = 500, Replications = 20, Evaluation sample size = 10000

Results: Supply Chain Network Design

Model	Deterministic	Traditional SP	Mean-MASD
cost_1	17541	22095	26556
$\hat{\mathbb{E}}[\text{cost}_2]$	144834	130156	126408
$SD[\text{cost}_2]$	861	681	600
$\Pr\{\text{infeas}\}_{0.95}(\%)$	43.52	13.29	4.39
$\hat{\mathbb{E}}[\text{cost}_2 \mid \text{feas}]$	109076	110104	118032
total-time	7.57	1414.32	1694.10
W1 (1300)	0	0	620
W2 (1100)	0	0	0
W3 (1200)	0	0	0
W4 (1200)	0	0	0
W5 (1100)	1100	1100	1100
W6 (1300)	610	1300	1239
D1 (1000)	800	1000	1000
D2 (900)	0	0	0
D3 (950)	0	0	0
D4 (1050)	910	1050	1050
D5 (1100)	0	0	0
D6 (1000)	0	0	0
D7 (980)	0	350	909
D8 (1050)	0	0	0

Conclusions

- Classical SP assumes accurate distribution and is risk-neutral.
- SP with risk functions offer a unifying treatment of these deficiencies.
- Classical sampling and decomposition algorithms extended to mean-MASD and mean-QDEV models.
- Extended models not much harder than classical SP.
- Additional research: Risk averse extensions of dynamic (multi-stage) stochastic inventory problems.

Conclusions

- Classical SP assumes accurate distribution and is risk-neutral.
- SP with risk functions offer a unifying treatment of these deficiencies.
- Classical sampling and decomposition algorithms extended to mean-MASD and mean-QDEV models.
- Extended models not much harder than classical SP.
- Additional research: Risk averse extensions of dynamic (multi-stage) stochastic inventory problems.

Conclusions

- Classical SP assumes accurate distribution and is risk-neutral.
- SP with risk functions offer a unifying treatment of these deficiencies.
- Classical sampling and decomposition algorithms extended to mean-MASD and mean-QDEV models.
- Extended models not much harder than classical SP.
- Additional research: Risk averse extensions of dynamic (multi-stage) stochastic inventory problems.

Conclusions

- Classical SP assumes accurate distribution and is risk-neutral.
- SP with risk functions offer a unifying treatment of these deficiencies.
- Classical sampling and decomposition algorithms extended to mean-MASD and mean-QDEV models.
- Extended models not much harder than classical SP.
- Additional research: Risk averse extensions of dynamic (multi-stage) stochastic inventory problems.

Conclusions

- Classical SP assumes accurate distribution and is risk-neutral.
- SP with risk functions offer a unifying treatment of these deficiencies.
- Classical sampling and decomposition algorithms extended to mean-MASD and mean-QDEV models.
- Extended models not much harder than classical SP.
- Additional research: Risk averse extensions of dynamic (multi-stage) stochastic inventory problems.

Conclusions

- Classical SP assumes accurate distribution and is risk-neutral.
- SP with risk functions offer a unifying treatment of these deficiencies.
- Classical sampling and decomposition algorithms extended to mean-MASD and mean-QDEV models.
- Extended models not much harder than classical SP.
- Additional research: Risk averse extensions of dynamic (multi-stage) stochastic inventory problems.