Identifying Effective Scenarios in Distributionally Robust Optimization

Güzin Bayraksan

Department of Integrated Systems Engineering
The Ohio State University

8 March 2018

Joint work with Hamed Rahimian (The Ohio State University) and
Tito Homem-de-Mello (Universidad Adolfo Ibañez)

Shameless Plug

Optimization Under Uncertainty Sessions @INFORMS 2018

INFORMS Annual Meeting

November 4-7, 2018

`https://tinyurl.com/IOSOptUncertainty2018`

Outline

Outline

Convex Stochastic Optimization Problem

$$\min_{x \in \mathbb{X}} \mathbb{E}_{\mathbf{q}} \left[h_\omega(x) \right],$$

where

- $\mathbb{X} \subseteq \mathbb{R}^m$ is a deterministic and non-empty convex compact set,
- Ω is a finite sample space, $\Omega = \{1, 2, \ldots, n\}$,
- $\mathbf{q}$ denotes a known probability distribution,
- cost function $h_\omega(\cdot) : \mathbb{X} \mapsto \mathbb{R}$ is convex and real-valued for all $\omega \in \Omega$.

Convex Stochastic Optimization Problem

$$\min_{x \in \mathbb{X}} \sum_{\omega \in \mathbf{\Omega}} \mathbf{q}_\omega h_\omega(x),$$

where

- $\mathbb{X} \subseteq \mathbb{R}^m$ is a deterministic and non-empty convex compact set,
- **Ω is a finite sample space, $\Omega = \{1, 2, \dots, n\}$,**
- $\mathbf{q}$ denotes a known probability distribution,
- cost function $h_\omega(\cdot) : \mathbb{X} \mapsto \mathbb{R}$ is convex and real-valued for all $\omega \in \Omega$.

Convex Stochastic Optimization Problem

$$\min_{x \in \mathbb{X}} \sum_{\omega \in \Omega} q_\omega \mathbf{h}_\omega(\mathbf{x}),$$

where

- $\mathbb{X} \subseteq \mathbb{R}^m$ is a deterministic and **non-empty convex compact** set,
- Ω is a finite sample space, $\Omega = \{1, 2, \ldots, n\}$,
- $\mathbf{q}$ denotes a known probability distribution,
- cost function $h_\omega(\cdot) : \mathbb{X} \mapsto \mathbb{R}$ is **convex and real-valued** for all $\omega \in \Omega$.

Example: Influenza Vaccine Production

Determining the number of influenza vaccines to produce before the influenza season

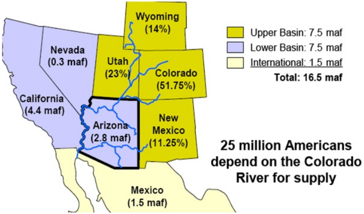
The decision maker

1. might be risk-averse,
2. might have partial knowledge about the underlying probability distribution, e.g., from historical data and/or expert opinions.

Example: Water Allocation Problem

Water managers are risk averse: they are concerned about low-probability but high-impact events

Ambiguous uncertainties: climate, population, water-use trends, hydrological changes, etc.



How to best allocate water among different users (agriculture, hydroelectric, municipal, etc.) while meeting uncertain water demand and not exceeding uncertain water supply over?

Distributionally Robust Convex Optimization (DRO)

$$\min_{x \in \mathbb{X}} \left\{ f(x) := \max_{\mathbf{p} \in \mathcal{P}} \mathbb{E}_{\mathbf{p}} \left[h_\omega(x) \right] \right\},$$

where

- $\mathcal{P}$ is the ambiguity set of distributions
- $\mathcal{P}$ is a subset of all probability distributions on Ω
- Hedge against the worst expectation with a probability distribution in the ambiguity set

Distributionally Robust Convex Optimization (DRO)

$$\min_{x \in \mathbb{X}} \left\{ f(x) := \max_{\mathbf{p} \in \mathcal{P}} \sum_{\omega \in \Omega} p_\omega h_\omega(x) \right\},$$

where

- $\mathcal{P}$ is the ambiguity set of distributions
- $\mathcal{P}$ is a subset of all probability distributions on Ω
- Hedge against the worst expectation with a probability distribution in the ambiguity set

Approaches to Construct the Ambiguity Set

Different ways to form the ambiguity set of distributions $\mathcal{P}$:

- *Moment* based sets:
  e.g., (Scarf, 1958; Delage and Ye, 2010; Wiesemann, Kuhn and Sim, 2014; Hasunanto et. al, 2014)
- *Distance* based sets:
  - *Prokhorov metric* (Erdoğan and Iyengar, 2006)
  - *Kantorovich/Wassertein metric* (Pflug and Wozabal, 2007; Wozabal, 2012; Mehrotra and Zhang, 2014; Esfahani and Kuhn, 2015; Gao and Kleywegt, 2016)
  - ζ*-Structure metrics* (Zhao and Guan, 2015)
  - ϕ*-divergence* based sets:
    - χ^2*-distance* (Klabjan et al., 2013)
    - *Kullback-Leibler divergence* (Calafiore, 2007),(Hu and Hong, 2015)
    - *Variation distance:* (Jiang and Guan, 2015), (R., B., H.-d.-M., 2017)
    - *General* ϕ*-divergences* (Ben Tal et al., 2013), (Yanikoglu and den Hertog, 2013), (Jiang and Guan, 2015), (Bayraksan and Love, 2015)

Approaches to Construct the Ambiguity Set

Different ways to form the ambiguity set of distributions $\mathcal{P}$:

- *Moment* based sets:
  e.g., (Scarf, 1958; Delage and Ye, 2010; Wiesemann, Kuhn and Sim, 2014; Hasunanto et. al, 2014)
- *Distance* based sets:
  - *Prokhorov metric* (Erdoğan and Iyengar, 2006)
  - *Kantorovich/Wassertein metric* (Pflug and Wozabal, 2007; Wozabal, 2012; Mehrotra and Zhang, 2014; Esfahani and Kuhn, 2015; Gao and Kleywegt, 2016)
  - ζ*-Structure metrics* (Zhao and Guan, 2015)
  - ϕ*-divergence* based sets:
    - χ^2*-distance* (Klabjan et al., 2013)
    - *Kullback-Leibler divergence* (Calafiore, 2007),(Hu and Hong, 2015)
    - *Variation distance:* (Jiang and Guan, 2015), (R., B., H.-d.-M., 2017)
    - *General* ϕ*-divergences* (Ben Tal et al., 2013), (Yanikoglu and den Hertog, 2013), (Jiang and Guan, 2015), (Bayraksan and Love, 2015)

Approaches to Construct the Ambiguity Set

Different ways to form the ambiguity set of distributions $\mathcal{P}$:

- *Moment* based sets:
  e.g., (Scarf, 1958; Delage and Ye, 2010; Wiesemann, Kuhn and Sim, 2014; Hasunanto et. al, 2014)
- *Distance* based sets:
  - *Prokhorov metric* (Erdoğan and Iyengar, 2006)
  - *Kantorovich/Wassertein metric* (Pflug and Wozabal, 2007; Wozabal, 2012; Mehrotra and Zhang, 2014; Esfahani and Kuhn, 2015; Gao and Kleywegt, 2016)
  - ζ*-Structure metrics* (Zhao and Guan, 2015)
  - ϕ*-divergence* based sets:
    - χ^2*-distance* (Klabjan et al., 2013)
    - *Kullback-Leibler divergence* (Calafiore, 2007),(Hu and Hong, 2015)
    - *Variation distance:* (Jiang and Guan, 2015), (R., B., H.-d.-M., 2017)
    - *General* ϕ*-divergences* (Ben Tal et al., 2013), (Yanikoglu and den Hertog, 2013), (Jiang and Guan, 2015), (Bayraksan and Love, 2015)

Relation to Risk-Averse Optimization

Under appropriate conditions on $\mathcal{P}$, DRO can be written as

$$\min_{x \in \mathbb{X}} \left\{ \mathcal{R}(h(x)) \right\},$$

where $\mathcal{R}$ is a coherent risk measure.

Relation to Risk-Averse Optimization

Under appropriate conditions on $\mathcal{P}$, DRO can be written as

$$\min_{x \in \mathbb{X}} \left\{ \mathcal{R}(h(x)) \right\},$$

where $\mathcal{R}$ is a coherent risk measure.

Coherent risk measures can be interpreted as worst-case expectations from a set of probability measures:

$$\mathcal{R}(Y) = \max_{A \in \mathcal{A}} \mathbb{E}_A [Y],$$

where $\mathcal{A}$ is a closed convex set of probability measures
(Artzner et al., 1999), (Rockafellar and Uryasev, 2013) (Ruszczyński and Shapiro, 2006)

Coherent Risk Measures (Rockafellar, 2007)

$\mathcal{R} : L^2 \to (-\infty, \infty]$ is a coherent risk measure (in the basic sense), defined on random variables (e.g., $Y, Y' \in L^2$) such that

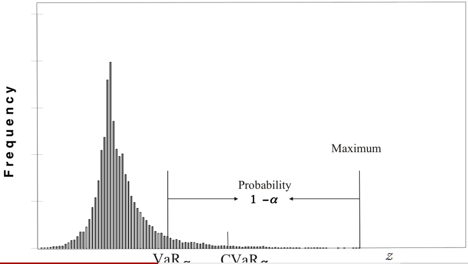
1. $\mathcal{R}(C) = C$ for all constants C,
2. $\mathcal{R}((1-\lambda)Y + \lambda Y') \leq (1-\lambda)\mathcal{R}(Y) + \lambda\mathcal{R}(Y')$ for all $\lambda \in [0, 1]$, i.e., $\mathcal{R}$ is convex,
3. $\mathcal{R}(Y) \leq \mathcal{R}(Y')$ when $Y \leq Y'$, i.e., $\mathcal{R}$ is monotonic,
4. $\mathcal{R}(Y) \leq 0$ when $||Y^k - Y||_2 \to 0$ with $\mathcal{R}(Y^k) \leq 0$, i.e., $\mathcal{R}$ is closed,
5. $\mathcal{R}(\lambda R) = \lambda\mathcal{R}(Y)$ for $\lambda > 0$, i.e., $\mathcal{R}$ is positively homogeneous.

Other related definitions (Artzner et al., 1999) (Ruszczyński and Shapiro, 2006)

Examples of Risk Measures

Let Z be an integrable random variable with distribution function $F_Z(\cdot)$

1. **Expectation** (risk-neutral): $\mathcal{R}(Z) = \mathbb{E}[Z]$
2. **Value-at-Risk (VaR)** at level $\alpha \in (0,1)$ (left-side α-quantile)
   $\mathcal{R}(Z) = \mathrm{VaR}_\alpha[Z] = \inf_\eta \{\eta : F_Z(\eta) \geq \alpha\}$
3. **Conditional Value-at-Risk (CVaR)** at $\alpha \in (0,1)$
   $\mathcal{R}(Z) = \mathrm{CVaR}_\alpha[Z] = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\tau[Z]\, d\tau$



Outline

DRO

$$\min_{x \in \mathbb{X}} \left\{ f(x) := \max_{\mathbf{p} \in \mathcal{P}} \mathbb{E}_{\mathbf{p}} \left[h_\omega(x) \right] \right\},$$

where

- $\mathcal{P}$ is the ambiguity set of distributions
- $\mathcal{P}$ is a subset of all probability distributions on Ω
- Hedge against the worst expectation with a probability distribution in the ambiguity set

Some Questions. . .

Consider an optimal solution $(x^*, \mathbf{p}^*) \in \mathbb{X} \times \mathcal{P}$ to DRO:

$$x^* \in \operatorname*{argmin}_{x \in \mathbb{X}} \mathbb{E}_{\mathbf{p}*} [h(x)]$$
$$\mathbf{p}^* \in \operatorname*{argmax}_{\mathbf{p} \in \mathcal{P}} \mathbb{E}_{\mathbf{p}} [h(x^*)]$$

Some Questions...

Consider an optimal solution $(x^*, \mathbf{p}^*) \in \mathbb{X} \times \mathcal{P}$ to DRO:

$$x^* \in \operatorname*{argmin}_{x \in \mathbb{X}} \mathbb{E}_{\mathbf{p}*}[h(x)]$$
$$\mathbf{p}^* \in \operatorname*{argmax}_{\mathbf{p} \in \mathcal{P}} \mathbb{E}_{\mathbf{p}}[h(x^*)]$$

With respect to the optimal 'worst-case' probability $\mathbf{p}^*$:

- What does the worst-case probability distribution look like?
- What information can be inferred from it?
- Which scenarios are more important?

Research Questions

Divide the scenarios into two groups:

$$E := \{\omega \in \Omega : p^*_\omega > 0\}$$

$$E^c := \{\omega \in \Omega : p^*_\omega = 0\} = \Omega \setminus E$$

Are the scenarios in set E important?

Research Questions

Divide the scenarios into two groups:

$$E := \{\omega \in \Omega : p^*_\omega > 0\}$$
$$E^c := \{\omega \in \Omega : p^*_\omega = 0\} = \Omega \setminus E$$

Are the scenarios in set E important?

AIM: Help the decision maker to better understand their problem and encourage them to collect more accurate information surrounding certain scenarios.

Motivating Example: A Simple Inventory Problem

Decide now how many vaccines, x, should be produced to minimize expected cost, z.

The distribution of demand is not known.

Suppose the ambiguity set contains <u>all</u> probability distributions on the support of demand.

- Unit cost $= 2$
- Unit revenue$= 3$
- Demand$=(2, 5, 1)^\top$
- $\mathbf{q} = (0.30, 0.70, 0)^\top$

- $(x^*, z^*) = (1, -1)$
- $\mathbf{p}^* = (0.5, 0.50, 0)^\top$

Motivating Example: Remove Scenarios One by One

- Unit cost $= 2$
- Unit revenue$= 3$
- Demand$=(2, 5, 1)^\top$
- $\mathbf{q} = (0.30, 0.70, 0)^\top$

- $(x^*, z^*) = (1, -1)$
- $\mathbf{p}^* = (0.5, 0.50, 0)^\top$

Motivating Example: Remove Scenarios One by One

- Unit cost $= 2$
- Unit revenue$= 3$
- Demand$=(2, 5, 1)^\top$
- $\mathbf{q} = (0.30, 0.70, 0)^\top$

- $(x^*, z^*) = (1, -1)$
- $\mathbf{p}^* = (0.5, 0.50, 0)^\top$

- If $p_1^* = 0$ and/or $p_2^* = 0 \Longrightarrow$ optimal value does not change
- If $p_3^* = 0$ the optimal value does change.

Motivating Example: Remove Scenarios One by One

- Unit cost $= 2$
- Unit revenue$= 3$
- Demand$=(2, 5, 1)^\top$
- $\mathbf{q} = (0.30, 0.70, 0)^\top$

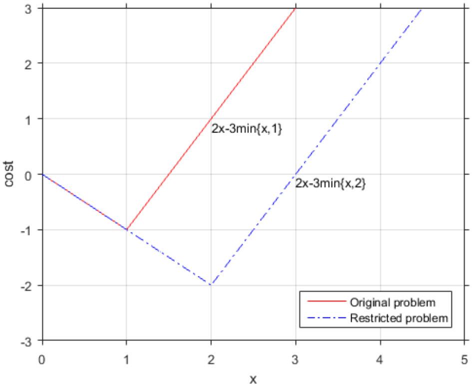
- $(x^*, z^*) = (1, -1)$
- $\mathbf{p}^* = (0.5, 0.50, 0)^\top$

- If $p_1^* = 0$ and/or $p_2^* = 0 \Longrightarrow$ optimal value does <u>not</u> change
- If $p_3^* = 0$ the optimal value does change.

→ The worst-case probability of a scenario is not necessarily an indication of its effect on the optimal value.

→ multiple optimal probabilities

Motivating Example: Third Scenario is Removed



Basic Definitions

Definition (**Effective Scenario)**

At an optimal solution x^*, a scenario is called effective if its removal causes the optimal objective function (and possibly the optimal x^*) to differ from that of the original problem.

Definition (**Ineffective Scenario)**

A scenario that is not effective is called ineffective.

How to Find Effective/Ineffective Scenarios?

Suppose that we have $(x^*, \mathbf{p}^*)$. How can we determine the effectiveness of a scenario?

- Resolve for any scenario $\omega \in \Omega$
  - Restrict the ambiguity set to $\mathbf{p}^*_\omega = 0$,
  - Resolve the resulting distributionally robust problem,
  - Compare the optimal values to determine the effectiveness of the scenario.

- Exploit the structure of the ambiguity set
  - Propose easy-to-check conditions (based on $\mathbf{p}^*$) to identify the effectiveness of a scenario,
  - Low computational cost
  - We might not be able to identify the effectiveness of all scenarios

How to Determine Effective Scenarios?

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathbb{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0, \ \omega \in \mathcal{F}\}$.

How to Determine Effective Scenarios?

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathbb{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

How to Determine Effective Scenarios?

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathbb{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

How to Determine Effective Scenarios?

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathbb{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0, \ \omega \in \mathcal{F}\}$.

How to Determine Effective Scenarios?

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathbb{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

How to Determine Effective Scenarios?

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathbb{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

If Inner Max of the Assessment Problem is Infeasible: $f^{\mathsf{A}}(x; \mathcal{F}) = -\infty$

Formal Definition

Let

- $S := \text{argmin}_{x \in \mathbb{X}} f(x)$ Set of opt. sol.s to DRO
- $S^{\text{A}}(\mathcal{F}) := \text{argmin}_{x \in \mathbb{X}} f^{\text{A}}(x; \mathcal{F})$ Set of opt. sol.s to Assess. Prob. in $\mathcal{F}$
- Suppose $x^* \in S$ and $\bar{x} \in S^{\text{A}}(\mathcal{F})$

Formal Definition

Let

- $S := \operatorname{argmin}_{x \in \mathbb{X}} f(x)$ Set of opt. sol.s to DRO
- $S^{\mathrm{A}}(\mathcal{F}) := \operatorname{argmin}_{x \in \mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F})$ Set of opt. sol.s to Assess. Prob. in $\mathcal{F}$
- Suppose $x^* \in S$ and $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$

Then

$$f^{\mathrm{A}}(\bar{x}; \mathcal{F}) \leq f^{\mathrm{A}}(x^*; \mathcal{F}) \leq f(x^*) \leq f(\bar{x}).$$

Formal Definition

Definition (**Effective Subset of Scenarios**)

A subset $\mathcal{F} \subset \Omega$ is called effective if $\min_{x \in \mathbb{X}} f^{\mathsf{A}}(x; \mathcal{F}) < \min_{x \in \mathbb{X}} f(x)$.

Formal Definition

Definition (**Effective Subset of Scenarios**)

A subset $\mathcal{F} \subset \Omega$ is called effective if $\min_{x \in \mathbb{X}} f^{\mathsf{A}}(x; \mathcal{F}) < \min_{x \in \mathbb{X}} f(x)$.

Formal Definition

Definition (**Effective Subset of Scenarios**)

A subset $\mathcal{F} \subset \Omega$ is called effective if $\min_{x \in \mathbb{X}} f^{\mathsf{A}}(x; \mathcal{F}) < \min_{x \in \mathbb{X}} f(x)$.

Essentially, a scenario is effective if its “removal” causes the optimal value of the problem to change.

Relating Effectiveness to Sets of Optimal Solutions

Proposition (**A sufficient condition for effectiveness)**

A subset $\mathcal{F} \subset \Omega$ *is effective if* $S \not\subset S^A(\mathcal{F})$.

Proposition (**Alternative characterization of effectiveness focusing on <u>all</u> opt. sol.s of the Assessment Problem)**

A subset $\mathcal{F} \subset \Omega$ *is effective if and only if* $f^A(\bar{x}; \mathcal{F}) < f(\bar{x})$ *for all* $\bar{x} \in S^A(\mathcal{F})$.

Effectiveness in Interaction With Other Subsets

Proposition (**Effectiveness in Interaction With Other Subsets**)

i. *The union of an effective subset with any other subset of Ω is effective.*

ii. *The intersection of an ineffective subset with any other subset of Ω is ineffective.*

Effectiveness in Interaction With Other Subsets

Proposition (**Effectiveness in Interaction With Other Subsets)**

i. *The union of an effective subset with any other subset of Ω is effective.*

ii. *The intersection of an ineffective subset with any other subset of Ω is ineffective.*

- Any subset of an ineffective subset is ineffective.
- The union of two or more ineffective subsets — unknown

How to Find Effective/Ineffective Scenarios?

Suppose that we have $(x^*, \mathbf{p}^*)$. How can we determine the effectiveness of a scenario?

- Resolve for any scenario $\omega \in \Omega$
  - Restrict the ambiguity set to $\mathbf{p}^*_\omega = 0$,
  - Resolve the resulting distributionally robust problem,
  - Compare the optimal values to determine the effectiveness of the scenario.

- Exploit the structure of the ambiguity set
  - Propose easy-to-check conditions (based on $\mathbf{p}^*$) to identify the effectiveness of a scenario,
  - Low computational cost
  - We might not be able to identify the effectiveness of all scenarios

Outline

Total Variation Distance

- $\mathbf{p} = \{p_\omega\}_{\omega=1}^n$ and $\mathbf{q} = \{q_\omega\}_{\omega=1}^n$ two probability vectors

$$\mathrm{V}(\mathbf{p}, \mathbf{q}) := \frac{1}{2} \sum_{\omega \in \Omega} |p_\omega - q_\omega|$$

  denotes the total variation distance between **p** and **q**.

- Metric
- $0 \leq \mathrm{V}(\mathbf{p}, \mathbf{q}) \leq 1$

DRO with Total Variation Distance

Instead of considering one ("nominal") distribution $\mathbf{q}$,

DRO with Total Variation Distance

Instead of considering one ("nominal") distribution $\mathbf{q}$,

Consider all distributions $\mathbf{p}$ in

DRO with Total Variation Distance

Instead of considering one ("nominal") distribution $\mathbf{q}$,

Consider all distributions $\mathbf{p}$ in

$$\mathcal{P} = \left\{ \mathsf{V}(\mathbf{p},\mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega} |p_\omega - q_\omega| \le \gamma, \right.$$

$$\sum_{\omega=1}^{n} p_\omega = 1,$$

$$\left. p_\omega \ge 0,\ \forall \omega \right\}.$$

DRO with Total Variation Distance

Instead of considering one ("nominal") distribution $\mathbf{q}$,

Consider all distributions $\mathbf{p}$ in

$$\mathcal{P} = \Big\{ \mathsf{V}(\mathbf{p}, \mathbf{q}) := \frac{1}{2} \sum_{\omega \in \Omega} |p_\omega - q_\omega| \leq \gamma, \\ \sum_{\omega=1}^{n} p_\omega = 1, \\ p_\omega \geq 0, \ \forall \omega \Big\}.$$

- all distributions sufficiently close to the nominal distribution

DRO with Total Variation Distance

Instead of considering one ("nominal") distribution **q**,

Consider all distributions **p** in

$$\mathcal{P} = \Big\{ \mathsf{V}(\mathbf{p}, \mathbf{q}) := \frac{1}{2} \sum_{\omega \in \Omega} |p_\omega - q_\omega| \leq \gamma, \\ \sum_{\omega=1}^{n} p_\omega = 1, \\ p_\omega \geq 0, \ \forall \omega \Big\} .$$

- ensure it is a probability measure

DRO with Total Variation Distance

$$\min_{x \in X} \max_{p \in \mathcal{P}} \sum_{\omega=1}^{n} p_\omega h_\omega(x)$$

where

$$\mathcal{P} = \left\{ \mathsf{V}(\mathbf{p}, \mathbf{q}) \leq \gamma, \sum_{\omega=1}^{n} p_\omega = 1, p_\omega \geq 0 \right\},$$

Right-Sided and Left-Sided Variation Distances

Right-Sided

$$\mathrm{V}^{\mathrm{R}}(\mathbf{p},\mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega}(p_\omega - q_\omega)_+$$

Left-Sided

$$\mathrm{V}^{\mathrm{L}}(\mathbf{p},\mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega}(q_\omega - p_\omega)_+$$

where $(\cdot)_+ := \max\{0,\cdot\}$.

- $0 \leq \mathrm{V}^{\mathrm{R}}(\mathbf{p},\mathbf{q}), \mathrm{V}^{\mathrm{L}}(\mathbf{p},\mathbf{q}) \leq \frac{1}{2}$
- Not Metrics

Risk-Averse Interpretation

Proposition (**Risk-Averse Interpretation of DRO with Total Variation**)

$$f_\gamma(x) = \begin{cases} \mathbb{E}_{\mathbf{q}}\left[h(x,\xi)\right], & \textit{if } \gamma = 0, \\ \gamma \sup_{\omega \in \Omega} h_\omega(x) + (1-\gamma)\,\mathrm{CVaR}_\gamma\left[\mathbf{h}(x)\right], & \textit{if } 0 < \gamma < 1, \\ \sup\limits_{\omega \in \Omega} h_\omega(x), & \textit{if } \gamma \geq 1, \end{cases}$$

By (Jiang and Guan, 2015).

Risk-Averse Interpretation

Proposition (**Risk-Averse Interpretation of DRO with Total Variation)**

$$f_\gamma(x) = \begin{cases} \mathbb{E}_{\mathbf{q}}\left[h(x,\xi)\right], & \text{if } \gamma = 0, \\ \gamma \sup_{\omega\in\Omega} h_\omega(x) + (1-\gamma)\,\mathrm{CVaR}_\gamma\left[\mathbf{h}(x)\right], & \text{if } 0 < \gamma < 1, \\ \sup_{\omega\in\Omega} h_\omega(x), & \text{if } \gamma \geq 1, \end{cases}$$

By (Jiang and Guan, 2015).

Proposition (**Equivalence of Left-Sided and Right-Sided Variation Distances)**

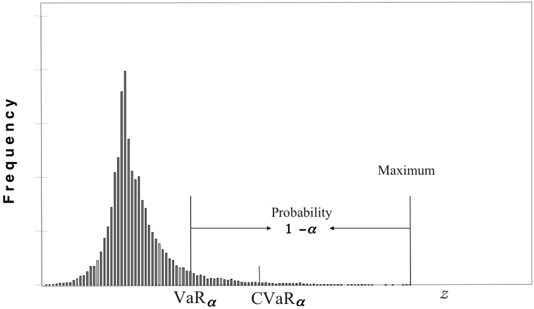
$$f^R_{\frac{\gamma}{2}}(x) = f^L_{\frac{\gamma}{2}}(x) = f_\gamma(x)$$

Recall VaR and CVaR

Let Z be an integrable random variable with distribution function $F_Z(\cdot)$

1. **Value-at-Risk (VaR)** at level $\alpha \in (0,1)$ (left-side α-quantile)
   $\mathcal{R}(Z) = \mathrm{VaR}_\alpha[Z] = \inf_\eta\{\eta : F_Z(\eta) \geq \alpha\}$

2. **Conditional Value-at-Risk (CVaR)** at $\alpha \in (0,1)$
   $\mathcal{R}(Z) = \mathrm{CVaR}_\alpha[Z] = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\tau[Z]\,d\tau$



Characterization of the Worst-Case Probabilities

Proposition (**Worst-Case Probabilities at $x \in \mathbb{X}$**)

$$\begin{cases} p_\omega = 0, & \omega \in \Omega_1(x) := \left[\mathbf{h}(x) < \mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right]\right], \\ p_\omega \leq q_\omega, & \omega \in \Omega_2(x) := \left[\mathbf{h}(x) = \mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right]\right], \\ p_\omega = q_\omega, & \omega \in \Omega_3(x) := \left[\mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right] < \mathbf{h}(x) < \sup_{\omega \in \Omega} h_\omega(x)\right], \\ p_\omega \geq q_\omega, & \omega \in \Omega_4(x) := \left[\mathbf{h}(x) = \sup_{\omega \in \Omega} h_\omega(x)\right], \end{cases}$$

Characterization of the Worst-Case Probabilities

Proposition (**Worst-Case Probabilities at $x \in \mathbb{X}$**)

$$
\begin{cases}
p_\omega = 0, & \omega \in \Omega_1(x) := [\mathbf{h}(x) < \mathrm{VaR}_\gamma [\mathbf{h}(x)]], \\
p_\omega \leq q_\omega, & \omega \in \Omega_2(x) := [\mathbf{h}(x) = \mathrm{VaR}_\gamma [\mathbf{h}(x)]], \\
p_\omega = q_\omega, & \omega \in \Omega_3(x) := [\mathrm{VaR}_\gamma [\mathbf{h}(x)] < \mathbf{h}(x) < \sup_{\omega \in \Omega} h_\omega(x)], \\
p_\omega \geq q_\omega, & \omega \in \Omega_4(x) := [\mathbf{h}(x) = \sup_{\omega \in \Omega} h_\omega(x)],
\end{cases}
$$

Characterization of the Worst-Case Probabilities

Proposition (Worst-Case Probabilities at $x \in \mathbb{X}$)

$$\begin{cases} p_\omega = 0, & \omega \in \Omega_1(x) := \left[\mathbf{h}(x) < \mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right]\right], \\ p_\omega \le q_\omega, & \omega \in \Omega_2(x) := \left[\mathbf{h}(x) = \mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right]\right], \\ p_\omega = q_\omega, & \omega \in \Omega_3(x) := \left[\mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right] < \mathbf{h}(x) < \sup_{\omega \in \Omega} h_\omega(x)\right], \\ p_\omega \ge q_\omega, & \omega \in \Omega_4(x) := \left[\mathbf{h}(x) = \sup_{\omega \in \Omega} h_\omega(x)\right], \end{cases}$$

Characterization of the Worst-Case Probabilities

Proposition (**Worst-Case Probabilities at $x \in \mathbb{X}$**)

$$\begin{cases} p_\omega = 0, & \omega \in \Omega_1(x) := [\mathbf{h}(x) < \mathrm{VaR}_\gamma [\mathbf{h}(x)]] , \\ p_\omega \leq q_\omega, & \omega \in \Omega_2(x) := [\mathbf{h}(x) = \mathrm{VaR}_\gamma [\mathbf{h}(x)]] , \\ p_\omega = q_\omega, & \omega \in \Omega_3(x) := [\mathrm{VaR}_\gamma [\mathbf{h}(x)] < \mathbf{h}(x) < \sup_{\omega \in \Omega} h_\omega(x)] , \\ p_\omega \geq q_\omega, & \omega \in \Omega_4(x) := [\mathbf{h}(x) = \sup_{\omega \in \Omega} h_\omega(x)] , \end{cases}$$

Characterization of the Worst-Case Probabilities

Proposition (Worst-Case Probabilities at $x \in \mathbb{X}$)

$$\begin{cases} p_\omega = 0, & \omega \in \Omega_1(x) := [\mathbf{h}(x) < \mathrm{VaR}_\gamma[\mathbf{h}(x)]], \\ p_\omega \le q_\omega, & \omega \in \Omega_2(x) := [\mathbf{h}(x) = \mathrm{VaR}_\gamma[\mathbf{h}(x)]], \\ p_\omega = q_\omega, & \omega \in \Omega_3(x) := [\mathrm{VaR}_\gamma[\mathbf{h}(x)] < \mathbf{h}(x) < \sup_{\omega \in \Omega} h_\omega(x)], \\ p_\omega \ge q_\omega, & \omega \in \Omega_4(x) := [\mathbf{h}(x) = \sup_{\omega \in \Omega} h_\omega(x)], \end{cases}$$

Characterization of the Worst-Case Probabilities

Proposition (Worst-Case Probabilities at $x \in \mathbb{X}$)

$$\begin{cases} p_\omega = 0, & \omega \in \Omega_1(x) := [\mathbf{h}(x) < \mathrm{VaR}_\gamma[\mathbf{h}(x)]], \\ p_\omega \le q_\omega, & \omega \in \Omega_2(x) := [\mathbf{h}(x) = \mathrm{VaR}_\gamma[\mathbf{h}(x)]], \\ p_\omega = q_\omega, & \omega \in \Omega_3(x) := [\mathrm{VaR}_\gamma[\mathbf{h}(x)] < \mathbf{h}(x) < \sup_{\omega \in \Omega} h_\omega(x)], \\ p_\omega \ge q_\omega, & \omega \in \Omega_4(x) := [\mathbf{h}(x) = \sup_{\omega \in \Omega} h_\omega(x)], \end{cases}$$

coupled with additional constraints

$$\sum_{\omega \in \Omega_2(x)} p_\omega = \sum_{\omega \in \Omega_1(x) \cup \Omega_2(x)} q_\omega - \gamma,$$

$$\sum_{\omega \in \Omega_4(x)} p_\omega = \gamma + \sum_{\omega \in \Omega_4(x)} q_\omega,$$

in addition to $\sum_{\omega \in \Omega} p_\omega = 1$ *and* $p_\omega \ge 0$, *for all* $\omega \in \Omega$.

Outline

Ineffective Scenarios of DRO-V

Theorem (**Easy-to-Check Conditions for Ineffective Scenarios)**

Suppose $(x^*, \mathbf{p}^*)$ *solves DRO-V. Then, a scenario* ω' *with* $q_{\omega'} \leq \gamma$*, is ineffective if any of the following conditions holds:*

- $\omega' \in \Omega_1(x^*)$,
- $\omega' \in \Omega_2(x^*)$ *and* $q_{\omega'} = 0$,
- $\omega' \in \Omega_2(x^*)$ *and* $\sum_{\omega \in \Omega_2(x^*)} p^*_\omega = 0$,
- $\omega' \in \Omega_3(x^*)$ *and* $q_{\omega'} = 0$.

Effective Scenarios of DRO-V

Theorem (**Easy-to-Check Conditions for Effective Scenarios)**

Suppose $(x^*, \mathbf{p}^*)$ solves DRO-V. Then, a scenario ω' is effective if any of the following conditions holds:

- $q_{\omega'} > \gamma$,
- $\Omega_2(x^*) = \{\omega'\}$ *and* $p^*_{\omega'} > 0$,
- $\omega' \in \Omega_3(x^*)$ *and* $q_{\omega'} > 0$,
- $\omega' \in \Omega_4(x^*)$ *and* $q_{\omega'} > 0$,
- $\Omega_4(x^*) = \{\omega'\}$.

Effectiveness of (Non-Singleton) Subsets of Scenarios

Proposition (Recall... the General Results)

- *The union of an effective subset with any other subset of Ω is effective.*
- *The intersection of an ineffective subset with any other subset of Ω is ineffective.*

Effectiveness of (Non-Singleton) Subsets of Scenarios

Proposition (Recall... the General Results)

- *The union of an effective subset with any other subset of Ω is effective.*
- *The intersection of an ineffective subset with any other subset of Ω is ineffective.*

Theorem (**Union of "Some" Ineffective Scenarios of DRO-V**)

The union of any of the ineffective scenarios identified in the first theorem is ineffective.

Effectiveness of (Non-Singleton) Subsets of Scenarios

Proposition (Recall... the General Results)

- *The union of an effective subset with any other subset of Ω is effective.*
- *The intersection of an ineffective subset with any other subset of Ω is ineffective.*

Theorem (**Union of "Some" Ineffective Scenarios of DRO-V)**

The union of any of the ineffective scenarios *identified in the first theorem is ineffective.*

Union of arbitrary ineffective scenarios is not necessarily ineffective.

Example: Lot-Sizing Problem

Consider $\gamma = 0.15$:

- Unit ordering cost $= 1$
- Unit backorder $= 4$
- Unit holding cost $= 8$
- Demand$=(1, 2, 3, 4)^\top$
- $\mathbf{q} = (0, 0.5, 0.5, 0)^\top$

Example: Lot-Sizing Problem

Consider $\gamma = 0.15$:

- Unit ordering cost $= 1$
- Unit backorder $= 4$
- Unit holding cost $= 8$
- Demand$=(1, 2, 3, 4)^\top$
- $\mathbf{q} = (0, 0.5, 0.5, 0)^\top$

- $(x^*, z^*) = (2, 5.2)$
- $\Omega_4(x^*) := \{1, 4\}$.

Example: Lot-Sizing Problem

Consider $\gamma = 0.15$:

- Unit ordering cost $= 1$
- Unit backorder $= 4$
- Unit holding cost $= 8$
- Demand$=(1, 2, 3, 4)^\top$
- $\mathbf{q} = (0, 0.5, 0.5, 0)^\top$

- $(x^*, z^*) = (2, 5.2)$
- $\Omega_4(x^*) := \{1, 4\}$.

- We do not have easy-to-check conditions to identify the effectiveness of scenarios $d = 1$ and 4.
- Solving the assessment problem concludes both scenarios $d = 1$ and 4 are individually ineffective.

Example: Lot-Sizing Problem

Consider $\gamma = 0.15$:

- Unit ordering cost $= 1$
- Unit backorder $= 4$
- Unit holding cost $= 8$
- Demand$=(1, 2, 3, 4)^\top$
- $\mathbf{q} = (0, 0.5, 0.5, 0)^\top$

- $(x^*, z^*) = (2, 5.2)$
- $\Omega_4(x^*) := \{1, 4\}$.

- We do not have easy-to-check conditions to identify the effectiveness of scenarios $d = 1$ and 4.
- Solving the assessment problem concludes both scenarios $d = 1$ and 4 are individually ineffective.
- If $p_1^* = 0$ and $p_4^* = 0 \Longrightarrow$ optimal value does change.
- The union of two ineffective scenarios $d = 1$ and $d = 4$ is effective.

Effective/Ineffective Scenarios of DRO-V

- Together, the above theorems identify only a subset of the ineffective and effective scenarios.
- They use only cost and probability (worst-case p^*_ω or nominal q_ω) information
- By examining the Assessment Problem in more detail, we can do more. . .

Assessment Problem of DRO-V

Assessment Problem of scenarios in $\mathcal{F} \subset \Omega$:

$$\min_{x \in \mathbb{X}} \left\{ f_\gamma^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}_\gamma^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}_\gamma^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P}_\gamma : p_\omega = 0, \ \omega \in \mathcal{F}\}$.

Assessment Problem of DRO-V

Assessment Problem of scenarios in $\mathcal{F} \subset \Omega$:

$$\min_{x \in \mathbb{X}} \left\{ f_\gamma^{\mathsf{A}}(x;\mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}_\gamma^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\},$$

where $\mathcal{P}_\gamma^{\mathsf{A}} := \{\mathbf{p} \in \mathcal{P}_\gamma : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

Let

- Let $\mathbb{Q}(\mathcal{F}) := \sum_{\omega \in \mathcal{F}} q_\omega$ be the Nominal Probability of $\mathcal{F}$
- If $\mathbb{Q}(\mathcal{F}) > \gamma$, then $\mathcal{F}$ is effective

Risk-Averse Interpretation of the DRO-V Assessment Problem

For $\mathbb{Q}(\mathcal{F}) \leq \gamma \leq 1$:

$$f_{\gamma}^{\mathsf{A}}(x;\mathcal{F}) = \begin{cases} \gamma \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_{\omega}(x) + (1-\gamma) \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} q_{\omega|\mathcal{F}^{\mathsf{c}}} h_{\omega}(x), & \gamma = \mathbb{Q}(\mathcal{F}), \\ \gamma \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_{\omega}(x) + (1-\gamma)\, \mathrm{CVaR}_{\gamma_{\mathcal{F}}} \left[\mathbf{h}(x) | \mathcal{F}^{\mathsf{c}}\right], & \mathbb{Q}(\mathcal{F}) < \gamma < 1, \\ \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_{\omega}(x), & \gamma = 1, \end{cases}$$

Risk-Averse Interpretation of the DRO-V Assessment Problem

For $\mathbb{Q}(\mathcal{F}) \leq \gamma \leq 1$:

$$f_\gamma^{\mathsf{A}}(x;\mathcal{F}) = \begin{cases} \gamma \sup_{\omega \in \mathcal{F}^c} h_\omega(x) + (1-\gamma) \sum_{\omega \in \mathcal{F}^c} q_{\omega|\mathcal{F}^c} h_\omega(x), & \gamma = \mathbb{Q}(\mathcal{F}), \\ \gamma \sup_{\omega \in \mathcal{F}^c} h_\omega(x) + (1-\gamma)\, \mathrm{CVaR}_{\gamma_{\mathcal{F}}} \left[\mathbf{h}(x)|\mathcal{F}^c\right], & \mathbb{Q}(\mathcal{F}) < \gamma < 1, \\ \sup_{\omega \in \mathcal{F}^c} h_\omega(x), & \gamma = 1, \end{cases}$$

where

- sup taken over $\mathcal{F}^c$
- $\mathrm{CVaR}_{\gamma_{\mathcal{F}}} \left[\mathbf{h}(x)|\mathcal{F}^c\right]$ is the conditional CVaR of $\mathbf{h}(x)$ at level $\gamma_{\mathcal{F}}$ with respect to the conditional distribution $\mathbf{q}_{\cdot|\mathcal{F}^c}$.

Risk-Averse Interpretation of the DRO-V Assessment Problem

The CVaR level has to be adjusted to

$$\gamma_{\mathcal{F}} := \frac{\gamma - \mathbb{Q}(\mathcal{F})}{1 - \mathbb{Q}(\mathcal{F})}$$

- Probability mass function of scenario ω *conditioned on* $\mathcal{F}^{\mathsf{c}}$:

$$q_{\omega|\mathcal{F}^{\mathsf{c}}} := \frac{q_\omega}{1 - \mathbb{Q}(\mathcal{F})}$$

Additional Notation

Let $\mathcal{F} \subset \Omega$, $x \in \mathbb{X}$, and $\eta \in \mathbb{R}$.

- Cumulative distribution function *conditioned on* $\mathcal{F}^{\mathsf{c}}$:

$$\Psi_{|\mathcal{F}^{\mathsf{c}}}(x, \eta) := \sum_{\omega \in \mathcal{F}^{\mathsf{c}} \cap [\mathbf{h}(x) \leq \eta]} q_{\omega|\mathcal{F}^{\mathsf{c}}}$$

- VaR of $\mathbf{h}(x)$ at level $0 \leq \gamma_{\mathcal{F}} \leq 1$ *conditioned on* $\mathcal{F}^{\mathsf{c}}$:

$$\mathrm{VaR}_{\gamma_{\mathcal{F}}}[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}] := \inf\{\eta : \Psi_{|\mathcal{F}^{\mathsf{c}}}(x, \eta) \geq \gamma_{\mathcal{F}}\}$$

Beyond Previous Theorems: Identify Undetermined Scenarios

Theorem (**Easy-to-Check Conditions to Identify Undetermined Scenarios)**

Suppose $(x^*, \mathbf{p}^*)$ *solves DRO-V. For a scenario* $\omega' \in \Omega_2(x^*)$ *with* $q_{\omega'} > 0$*, suppose that the effectiveness of scenario* ω' *is* *not* *identified by the previous theorems. Let* $\mathcal{F} = \{\omega'\}$*. If*

1. $\mathrm{VaR}_{\gamma_{\mathcal{F}}}[\mathbf{h}(x^*)|\mathcal{F}^c] < \mathrm{VaR}_{\gamma}[\mathbf{h}(x^*)]$*, and*
2. *either there exists a scenario* $\omega \in \left[\mathrm{VaR}_{\gamma_{\mathcal{F}}}[\mathbf{h}(x^*)|\mathcal{F}^c] < \mathbf{h}(x^*) < \mathrm{VaR}_{\gamma}[\mathbf{h}(x^*)]\right]$ *with* $q_{\omega} > 0$ *or* $\Psi_{|\mathcal{F}^c}\left(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}[\mathbf{h}(x^*)|\mathcal{F}^c]\right) > \gamma_{\mathcal{F}}$,

then scenario ω' *is effective.*

The Relationship Between Effective/Ineffective Scenarios and Multiple Worst-Case Probability Distributions

Theorem (**Necessary Conditions for Zero-Prob. Effective or Positive-Prob. Ineffective Scenarios**)

Suppose $(x^*, \mathbf{p}^*)$ *solves DRO-V. If there is either a zero-probability effective scenario* ω' *(*$p^*_{\omega'} = 0$*) or a positive-probability ineffective scenario* ω' *(*$p^*_{\omega'} > 0$*), then,*

- *there is a scenario* ω'' *with the same cost as that of* ω'*, i.e.,* $h_{\omega'}(x^*) = h_{\omega''}(x^*)$*, and*
- *the (primal) worst-case expected problem at* x^* *has multiple optimal solutions.*

Outline

Obtaining Optimal x^* and Worst-Case $\mathbf{p}^*$

- $\mathcal{P}$ is a polytope
- $\{p^k\}_{k \in K}$: the set of extreme points of $\mathcal{P}$

$$\min_{x \in \mathbb{X}} \max_{\mathbf{p} \in \mathcal{P}} \sum_{\omega \in \Omega} p_\omega h_\omega(x)$$

Obtaining Optimal x^* and Worst-Case $\mathbf{p}^*$

- $\mathcal{P}$ is a polytope
- $\{p^k\}_{k \in K}$: the set of extreme points of $\mathcal{P}$

$$\min_{x \in \mathbb{X}} \max_{k \in K} \sum_{\omega \in \Omega} p_\omega^k h_\omega(x)$$

Obtaining Optimal x^* and Worst-Case $\mathbf{p}^*$

- $\mathcal{P}$ is a polytope
- $\{p^k\}_{k \in K}$: the set of extreme points of $\mathcal{P}$

$$\text{(MP)} \quad \min_{x \in \mathbb{X}} \alpha$$
$$\text{s.t. } \alpha \geq \sum_{\omega \in \Omega} p_\omega^k h_\omega(x), \; k \in K$$

Obtaining Optimal x^* and Worst-Case $\mathbf{p}^*$

- $\mathcal{P}$ is a polytope
- $\{p^k\}_{k \in K}$: the set of extreme points of $\mathcal{P}$

$$
\text{(MP)} \min_{x \in \mathbb{X}} \alpha \\
\text{s.t. } \alpha \geq \sum_{\omega \in \Omega} p^k_\omega h_\omega(x), \; k \in K
$$

- Solve (MP) with a cut generation approach

Obtaining Optimal x^* and Worst-Case $\mathbf{p}^*$

- $\mathcal{P}$ is a polytope
- $\{p^k\}_{k\in K}$: the set of extreme points of $\mathcal{P}$

$$\text{(MP)} \min_{x\in\mathbb{X}} \alpha$$
$$\text{s.t. } \alpha \geq \sum_{\omega\in\Omega} p^k_\omega h_\omega(x),\ k \in K$$

- Solve (MP) with a cut generation approach
- Finite convergence

Obtaining Optimal x^* and Worst-Case $\mathbf{p}^*$

- $\mathcal{P}$ is a polytope
- $\{p^k\}_{k \in K}$: the set of extreme points of $\mathcal{P}$

$$
\text{(MP)} \min_{x \in \mathbb{X}} \alpha \quad \text{s.t. } \alpha \geq \sum_{\omega \in \Omega} p_\omega^k h_\omega(x), \; k \in K
$$

- Solve (MP) with a cut generation approach
- Finite convergence

This idea can be applied to any polyhedral ambiguity set, with finite convergence guaranteed

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$

$$\text{s.t. } \alpha \geq \sum_{\omega} p_{\omega}^{j} h_{\omega}(x), \; j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_{\omega} h_{\omega}(\hat{x}),$$

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$

$$\text{s.t. } \alpha \geq \sum_{\omega} p^j_\omega h_\omega(x), \; j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_\omega h_\omega(\hat{x}),$$

1. Solve (RMP), obtain $(\hat{x}, \hat{\alpha})$ and LB$=c\hat{x} + \hat{\alpha}$

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$

$$\text{s.t. } \alpha \geq \sum_{\omega} p^j_\omega h_\omega(x), \; j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_\omega h_\omega(\hat{x}),$$

1. Solve (RMP), obtain $(\hat{x}, \hat{\alpha})$ and LB$=c\hat{x} + \hat{\alpha}$
2. Obtain $h_\omega(\hat{x})$ for all ω

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$
$$\text{s.t. } \alpha \geq \sum_{\omega} p_{\omega}^{j} h_{\omega}(x),\ j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_{\omega} h_{\omega}(\hat{x}),$$

1. Solve (RMP), obtain $(\hat{x}, \hat{\alpha})$ and LB$=c\hat{x} + \hat{\alpha}$
2. Obtain $h_{\omega}(\hat{x})$ for all ω
3. Solve row-generation subproblem at $\hat{x}$, obtain $z(\hat{x})$ and p^{j}

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$
$$\text{s.t. } \alpha \geq \sum_{\omega} p_{\omega}^{j} h_{\omega}(x),\ j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_{\omega} h_{\omega}(\hat{x}),$$

1. Solve (RMP), obtain $(\hat{x}, \hat{\alpha})$ and LB$=c\hat{x} + \hat{\alpha}$
2. Obtain $h_{\omega}(\hat{x})$ for all ω
3. Solve row-generation subproblem at $\hat{x}$, obtain $z(\hat{x})$ and p^{j}
4. Update UB and (x^{*}, p^{*}), if needed
   - UB$= cx^{*} + z(x^{*})$

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$
$$\text{s.t. } \alpha \geq \sum_{\omega} p_{\omega}^{j} h_{\omega}(x),\ j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_{\omega} h_{\omega}(\hat{x}),$$

1. Solve (RMP), obtain $(\hat{x}, \hat{\alpha})$ and LB$=c\hat{x} + \hat{\alpha}$
2. Obtain $h_{\omega}(\hat{x})$ for all ω
3. Solve row-generation subproblem at $\hat{x}$, obtain $z(\hat{x})$ and p^{j}
4. Update UB and (x^{*}, p^{*}), if needed
   - UB$= cx^{*} + z(x^{*})$
5. Add the cut, corresponding to p^{j}, if $\hat{\alpha} < z(\hat{x})$. Stop, otherwise.

Primal Decomposition

Restricted Master Problem (RMP)

$$\min_{x \in \mathbb{X}} cx + \alpha$$
$$\text{s.t. } \alpha \geq \sum_{\omega} p^j_\omega h_\omega(x),\ j \in J \subseteq K$$

Row-generation Subproblem

$$z(\hat{x}) = \max_{p \in \mathcal{P}} \sum_{\omega} p_\omega h_\omega(\hat{x}),$$

1. Solve (RMP), obtain $(\hat{x}, \hat{\alpha})$ and LB$=c\hat{x} + \hat{\alpha}$
2. Obtain $h_\omega(\hat{x})$ for all ω
3. Solve row-generation subproblem at $\hat{x}$, obtain $z(\hat{x})$ and p^j
4. Update UB and (x^*, p^*), if needed
   - UB$= cx^* + z(x^*)$
5. Add the cut, corresponding to p^j, if $\hat{\alpha} < z(\hat{x})$. Stop, otherwise.

Can apply a decomposition-based method to obtain an outer approximation to $h_\omega(x)$

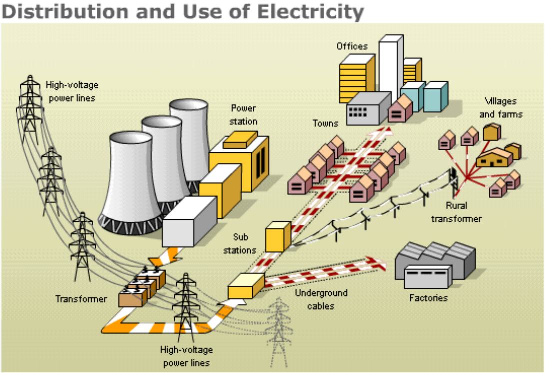
Outline

Test Problems

Table: Test problems used.

Problem name	# of 1st stage variables	# of 2nd stage variables	# of stochastic parameters	# of scenarios
APL1P	2	9	5	1280
Water Allocation	636	1060	7	200

APL1P (Infanger, 1992)



Power network with uncertain demand:

- First-stage decisions: What capacities to install at the generators?
- Second-stage decisions: Purchase additional capacities to fulfill unmet demands

Table: APL1P, $n = 1280$.

γ	# of scenarios $\Omega_1(x^*)$	$\Omega_2(x^*)$	$\Omega_3(x^*)$	$\Omega_4(x^*)$	# of scenarios ineffective	effective	undetermined
0.00	0	3	1276	1	0	1280	0
0.05	74	2	1203	1	74	1205	1
0.10	136	1	1142	1	136	1144	0
0.15	189	1	1089	1	189	1091	0
0.20	226	1	1052	1	226	1054	0
0.25	267	1	1011	1	267	1013	0
0.30	312	4	963	1	312	966	2
0.35	353	4	922	1	353	924	3
0.40	384	3	892	1	384	893	3
0.45	431	6	842	1	431	843	6
0.50	471	6	802	1	471	803	6
0.55	510	7	762	1	510	763	7
0.60	561	7	711	1	561	712	7
0.65	600	6	673	1	600	674	6
0.70	671	3	605	1	671	609	0
0.75	728	11	540	1	728	541	11
0.80	804	10	465	1	804	466	10
0.85	899	9	371	1	899	379	2
0.90	988	12	279	1	988	280	12
0.95	1076	12	191	1	1076	192	12
1.00	1279	-	-	1	1279	1	0

Water Allocation Problem

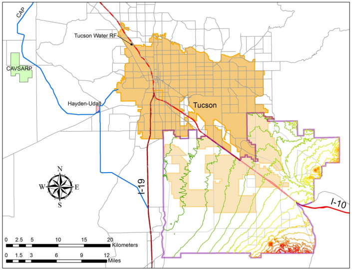


Figure: The southeastern region of Tucson, AZ.

How to best allocate Colorado River water among different users while meeting uncertain water demand and not exceeding uncertain water supply over the next 16 years? (Zhang, Rahimian, B., 2016)

Table: Water problem with equally likely scenarios, $n = 200$.

γ	# of scenarios				# of scenarios		
	$\Omega_1(x^*)$	$\Omega_2(x^*)$	$\Omega_3(x^*)$	$\Omega_4(x^*)$	ineffective	effective	undetermined
0.00	0	1	198	1	0	200	0
0.05	9	1	189	1	10	190	0
0.10	19	1	179	1	20	180	0
0.15	29	1	169	1	30	170	0
0.20	39	1	159	1	40	160	0
0.25	49	1	149	1	50	150	0
0.30	59	1	139	1	60	140	0
0.35	69	1	129	1	70	130	0
0.40	79	1	119	1	80	120	0
0.45	79	1	109	1	90	110	0
0.50	99	1	99	1	100	100	0
0.55	109	1	89	1	110	90	0
0.60	119	1	79	1	120	80	0
0.65	129	1	69	1	130	70	0
0.70	139	1	59	1	140	60	0
0.75	149	1	49	1	150	50	0
0.80	159	1	39	1	160	40	0
0.85	169	1	29	1	170	30	0
0.90	179	1	19	1	180	20	0
0.95	189	1	9	1	190	10	0
1.00	199	1	-	-	199	1	0

Are the Effective/Ineffective Subsets Monotone?

Conjecture 1 (C1)

The number of effective/ineffective scenarios is **monotone**

Conjecture 2 (C2)

The sets of effective/ineffective scenarios are **nested**

Conjecture 3 (C3)

The effectiveness of scenarios is **monotone**: Once a scenario becomes ineffective (or effective) it remains the same at higher values of γ

(C1) and (C3) are implied by (C2). But, neither of them imply (C2).

- For APL1P, PGP2, and the Water Allocation problem, (C1)–(C3) does hold.

Counterexample 1

Breakage of (C1), and hence Breakage of (C2), while (C3) Does Hold

Lot-sizing problem

- Unit ordering cost $= 4$
- Unit backorder $= 5$
- Unit holding cost $= 5$
- Demand$=(1, 2, 3, 4, 5, 6)^\top$
- $\mathbf{q} = (0, 0.2, 0.25, 0.2, 0.35, 0)^\top$

Table: Lot-sizing problem, $n = 6$.

γ	# of scen.s Ineffective	# of scen.s Effective	scen. # 1	2	3	4	5	6
0.00	2	4	I	E	E	E	E	I
0.05	1	5	I	E	E	E	E	E
0.10	1	5	I	E	E	E	E	E
0.15	1	5	I	E	E	E	E	E
0.20	1	5	I	E	E	E	E	E
0.25	1	5	I	E	E	E	E	E
0.30	1	5	I	E	E	E	E	E
0.35	2	4	I	E	I	E	E	E
0.40	2	4	I	E	I	E	E	E
0.45	2	4	I	E	I	E	E	E
0.50	2	4	I	E	I	E	E	E
0.55	3	3	I	E	I	I	E	E
0.60	3	3	I	E	I	I	E	E
0.65	4	2	I	I	I	I	E	E
0.70	4	2	I	I	I	I	E	E
0.75	4	2	I	I	I	I	E	E
0.80	4	2	I	I	I	I	E	E
0.85	4	2	I	I	I	I	E	E
0.90	5	1	I	I	I	I	I	E
0.95	**4**	**2**	**E**	**I**	**I**	**I**	**I**	**E**
1.00	4	2	E	I	I	I	I	E

Why This Happened?

- This problem contains a high-cost scenario with zero nominal probability ($d_1 = 1$)
- For high-enough γ, this scenario is
  - "popped" (that is, given a positive probability p_1^*), and
  - eventually becomes an effective scenario.

Counterexample 2

Breakage of (C3), and hence Breakage of (C2), while (C1) Does Hold.

- A variant of APL1P, but with 16 scenarios
- Set the availability of the second generator, first and third demands to their expected values
- Stochastic parameters: the availability of the first generator and second demand.
- Consider the scenario with 900 units of demand and 90% generator availability

Table: Variant of APL1P, $n = 16$.

	# of scen.s		
γ	Ineffective	Effective	Effectiveness
0.00	0	16	E
0.05	1	15	E
0.10	**2**	**14**	**I**
0.15	2	14	E
0.20	3	13	I
0.25	3	13	I
0.30	4	12	I
0.35	5	11	I
0.40	5	11	I
0.45	6	10	I
0.50	8	8	I
0.55	8	8	I
0.60	9	7	I
0.65	9	7	I
0.70	9	7	I
0.75	10	6	I
0.80	10	6	I
0.85	11	5	I
0.90	12	4	I
0.95	13	3	I
1.00	15	1	I

Why This Happened?

- This problem contains several scenarios of similar cost (e.g., 1000 units of demand with 100% generator availability and 900 units of demand with 90% generator availability).
- When x^* changes at different γ values, these similar scenarios exchange primal categories (e.g., from $\Omega_2(x^*)$ to $\Omega_1(x^*)$ and back) because the order of their costs change at different x^*.
- This exchange of primal categories causes an effective scenario to become ineffective and return back to being effective.

Observe that for a lower resolution of 0.1-steps in the γ values, we do not observe breakage of monotonicity.

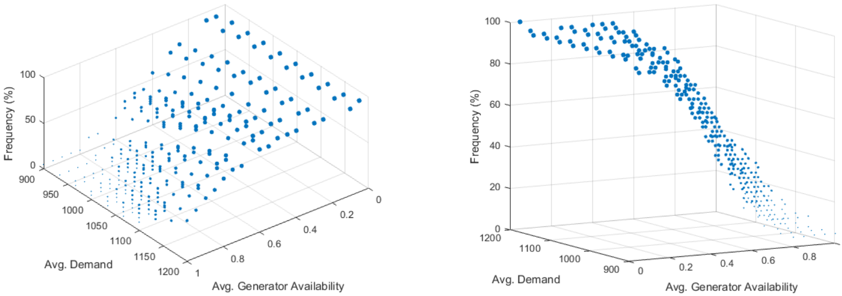
Summary

In our test problems:

- The easy-to-check conditions work quite efficiently,
- "undetermined" scenarios form a small portion of scenarios.

While obtaining general conditions under which one can guarantee monotonicity may be difficult, (Rahimian, B., Homem-de-Mello, 2017) show monotonicity for newsvendor problems under appropriate assumptions.

Managerial Insights from Effective Scenarios for APL1P



Percentage of time (out of 20 γ values) a scenario is identified effective

Outline

How to Use Effective Scenarios?

- **Managerial insight about the underlying uncertainty**
- **Guide on choosing the level of robustness:** (Rahimian, B., Homem-de-Mello, 2017) propose levels of robustness based on the *price of optimism/pessimism* and *regrets* for a class of inventory problems
- **Scenario reduction/generation:** (Homem-de-Mello, Arṕon, Pagnoncelli, 2017) propose a probability metrics approach for risk-based models while iteratively reducing to/updating the set of effective scenarios based on the current solution

How to Use Effective Scenarios and Other Future Directions

- **Cut Management in Decomposition Algorithms?**

  Can we have better performance by focusing on "effective scenarios" to speed computation time?

- **Influential (potentially Outlier and Leverage) Points in Regression?**

  Relation to Regularized Statistical Learning??

- **Degree of Effectiveness:**

  Here, we had a binary decision (effective, ineffective). One could generalize results to consider "how effective" as well

Other Future Directions

- **Generalizations and Extensions:**
  - other ϕ-divergences and other ambiguity sets
  - Multi-stage setting
  - More general probability spaces

Rahimian, B., Hommem-de-Mello, Identifying effective scenarios in distributionally robust stochastic programs with total variation distance, **Mathematical Programming**, published online, 2018.

`https://doi.org/10.1007/s10107-017-1224-6`

Acknowledgments

Grateful to Hamed Rahimian and Tito Homem-de-Mello.

This research is funded in part by the National Science Foundation through CAREER Grant CMMI-1345626 and CMMI-1563504.

Thank you

(bayraksan.1@osu.edu)

Removal of a Scenario

$$
\begin{aligned}
\min_{x \in \mathbb{X}} \; \max_{p} \; & \sum_{\omega \neq \omega'} p_\omega h_\omega(x) \\
\text{s.t.} \quad & \sum_{\omega \neq \omega'} |p_\omega - q_\omega| \leq \gamma - q_{\omega'} \\
& \sum_{\omega \neq \omega'} p_\omega = 1 \\
& p_\omega \geq 0, \; \omega \neq \omega'
\end{aligned}
$$