Quantitative Stability Analysis in Distributionally Robust Optimization

Huifu Xu

School of Mathematical Sciences, University of Southampton, UK.

Presented at BIRS Workshop, Banff, Canada, 5 March 2018
The one-stage DRO model

Consider the following distributionally robust optimization problem:

\[
\begin{align*}
\text{(DRO)} \quad & \min_x \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] \\
\text{s.t.} \quad & x \in X,
\end{align*}
\]

(1)

where \( \mathcal{P} \) is a set of distributions which contains/approximates the true probability distribution of random variable \( \xi \).

Part I: Quantifying change of the ambiguity set
Construction of ambiguity sets

- Moment conditions
- Mixture distribution – contamination distribution
- Bayesian inference
- Marginal distributions
- Samples & empirical data
- Maximal likelihood
- Kullback-Leibler divergence
- Wasserstein ball, $\zeta$-ball
- ...
I. Quantifying change of the ambiguity set

Construction of ambiguity sets

- Moment conditions
- Mixture distribution – contamination distribution
- Bayesian inference
- Marginal distributions
- Samples & empirical data
- Maximal likelihood
- Kullback-Leibler divergence
- Wasserstein ball, $\zeta$-ball
- ...
Ambiguity set defined through moment conditions

\[ \mathcal{P} := \left\{ P : \begin{array}{l}
\mathbb{E}_P[\psi_i(\xi)] = \mu_i, \quad \text{for } i = 1, \cdots, p \\
\mathbb{E}_P[\psi_i(\xi)] \leq \mu_i, \quad \text{for } i = p + 1, \cdots, q
\end{array} \right\}, \tag{2} \]

where \( \psi_i : \Xi \to \mathbb{R}, i = 1, \cdots, q \), are measurable functions.
I. Quantifying change of the ambiguity set

Ambiguity set defined through moment conditions

\[ \mathcal{P} := \left\{ P : \begin{align*} \mathbb{E}_P[\psi_i(\xi)] &= \mu_i, \text{ for } i = 1, \cdots, p \\ \mathbb{E}_P[\psi_i(\xi)] &\leq \mu_i, \text{ for } i = p + 1, \cdots, q \end{align*} \right\}, \quad (2) \]

where \( \psi_i : \Xi \rightarrow \mathbb{R}, \ i = 1, \cdots, q, \) are measurable functions.

**Question:** What if the true \( \mu_i \) is not known?
I. Quantifying change of the ambiguity set

Moment conditions

Ambiguity set defined through moment conditions

\[ \mathcal{P} := \left\{ P : \begin{array}{l}
\mathbb{E}_P[\psi_i(\xi)] = \mu_i, \quad \text{for } i = 1, \ldots, p \\
\mathbb{E}_P[\psi_i(\xi)] \leq \mu_i, \quad \text{for } i = p + 1, \ldots, q
\end{array} \right\}, \quad (2) \]

where \( \psi_i : \Xi \to \mathbb{R}, \ i = 1, \ldots, q, \) are measurable functions.

**Question:** What if the true \( \mu_i \) is not known?

\[ \mathcal{P}_N := \left\{ P : \begin{array}{l}
\mathbb{E}_P[\psi_i(\xi)] = \mu_i^N, \quad \text{for } i = 1, \ldots, p \\
\mathbb{E}_P[\psi_i(\xi)] \leq \mu_i^N, \quad \text{for } i = p + 1, \ldots, q
\end{array} \right\} \quad (3) \]

where \( \mu_i^N \) is often constructed through samples.
I. Quantifying change of the ambiguity set

Difference between $\mathcal{P}$ and $\mathcal{P}_N$?

Let

$$\langle P, \psi \rangle := \int_{\Xi} \psi(\xi) P(d\xi).$$

We can write $\mathcal{P}$ and $\mathcal{P}_N$ as

$$\mathcal{P} = \{ P \in \mathcal{P}(\Xi) : \langle P, \psi_E(\xi) \rangle = \mu_E, \langle P, \psi_1(\xi) \rangle \leq \mu_1 \}$$

and

$$\mathcal{P}_N = \{ P \in \mathcal{P}(\Xi) : \langle P, \psi_E(\xi) \rangle = \mu^N_E, \langle P, \psi_1(\omega) \rangle \leq \mu^N_1 \}.$$
I. Quantifying change of the ambiguity set

Difference between $\mathcal{P}$ and $\mathcal{P}_N$?

Let

$$\langle P, \psi \rangle := \int_{\Xi} \psi(\xi) P(d\xi).$$

We can write $\mathcal{P}$ and $\mathcal{P}_N$ as

$$\mathcal{P} = \{ P \in \mathcal{P}(\Xi) : \langle P, \psi_E(\xi) \rangle = \mu_E, \langle P, \psi_I(\xi) \rangle \leq \mu_I \}$$

and

$$\mathcal{P}_N = \{ P \in \mathcal{P}(\Xi) : \langle P, \psi_E(\xi) \rangle = \mu_E^N, \langle P, \psi_I(\omega) \rangle \leq \mu_I^N \}.$$  

Q: Does $\mathcal{P}_N$ approximate $\mathcal{P}$? How?
Measuring the distance between probability measures: Metrics of $\zeta$-structure

Definition ($\zeta$-metric)

Let $\mathcal{P}(\Xi)$ denote the set of all probability distributions/measures over space $(\Xi, \mathcal{B})$. Let $P, Q \in \mathcal{P}(\Xi)$ and $\mathcal{G}$ be a family of real-valued measurable functions on $\Xi$. Define

$$d_{\mathcal{G}}(P, Q) := \sup_{g \in \mathcal{G}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|.$$ 

The (semi-) distance defined as such is called a metric with $\zeta$–structure.
Measuring the distance between probability measures: Metrics of $\zeta$-structure

Definition ($\zeta$-metric)

Let $\mathcal{P}(\Xi)$ denote the set of all probability distributions/measures over space $(\Xi, \mathcal{B})$. Let $P, Q \in \mathcal{P}(\Xi)$ and $\mathcal{G}$ be a family of real-valued measurable functions on $\Xi$. Define

$$d_{\mathcal{G}}(P, Q) := \sup_{g \in \mathcal{G}} \left| \mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)] \right|.$$ 

The (semi-) distance defined as such is called a metric with $\zeta$–structure.

Question: How to choose $\mathcal{G}$?
Metrics of $\zeta$-structure

\[ d_{\mathcal{G}}(P, Q) := \sup_{g \in \mathcal{G}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|. \]

- **Total variation metric** (denoted by $d_{TV}$):
  \[ \mathcal{G} := \{ g : \sup_{\xi \in \Xi} |g(\xi)| \leq 1 \}. \]
I. Quantifying change of the ambiguity set

Metrics of ζ-structure

\[ d_{\mathcal{G}}(P, Q) := \sup_{g \in \mathcal{G}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|. \]

- **Total variation metric** (denoted by \( d_{TV} \)):
  \[ \mathcal{G} := \{ g : \sup_{\xi \in \Xi} |g(\xi)| \leq 1 \}. \]

- **Kantorovich/Wasserstein metric** (denoted by \( d_K \)):
  \[ \mathcal{G} = \{ g : g \text{ is Lipschitz continuous with } L_1(g) \leq 1 \}. \]
Metrics of $\zeta$-structure

\[ d_G(P, Q) := \sup_{g \in G} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|. \]

- **Total variation metric** (denoted by $d_{TV}$): 
  \[ G := \{ g : \sup_{\xi \in \Xi} |g(\xi)| \leq 1 \} . \]

- **Kantorovich/Wasserstein metric** (denoted by $d_K$): 
  \[ G = \{ g : g \text{ is Lipschitz continuous with } L_1(g) \leq 1 \} . \]

- **Bounded Lipschitz metric** (denoted by $d_{BL}$): 
  \[ G := \{ g : \sup_{\xi \in \Xi} |g(\xi)| \leq 1, \ g \text{ is Lipschitz continuous with } L_1(g) \leq 1 \} \]
  where $L_1(g)$ denotes the Lipschitz modulus.
I. Quantifying change of the ambiguity set

Measuring the distance between probability measures

Hoffman’s lemma for moment problem

Hoffman’s lemma (Sun and Xu (2015))

There exists a positive constant $C$ depending on $\psi$ such that

$$d_{TV}(Q, \mathcal{P}) \leq C[(\|E_Q[\psi_I(\xi)] - \mu_I\|_+ + \|E_Q[\psi_E(\xi)] - \mu_E\|_+],$$

for any $Q \in \mathcal{P}(\Xi)$, where $\| \cdot \|$ denotes the Euclidean norm, $(a)_+ = \max(0, a)$ and the maximum is taken componentwise.
Quantifying the difference between $\mathcal{P}$ and $\mathcal{P}_N$

**Proposition 2.1**

*There exists a positive constant $C$ depending on $\psi$ such that*

\[
H_{TV}(\mathcal{P}_N, \mathcal{P}) \leq C \left[ \max(\|\mu^I - \mu_I\|, \|\mu_I - \mu^N_I\|) + \|\mu_E^N - \mu_E\| \right],
\]

*where $C$ is defined as in the Hoffman’s lemma and $H_{TV}$ denotes the Hausdorff distance under the total variation metric.*
Proof. Let $Q \in \mathcal{P}_N$. By the Hoffman lemma, there is a constant $C$ such that

$$d_{TV}(Q, \mathcal{P}) \leq C \left[ \left\| \mathbb{E}_Q[\psi_I(\xi(\omega))] - \mu_I \right\| + \left\| \mathbb{E}_Q[\psi_E(\xi(\omega))] - \mu_E \right\| \right]$$
Proof. Let $Q \in \mathcal{P}_N$. By the Hoffman lemma, there is a constant $C$ such that

$$d_{TV}(Q, \mathcal{P}) \leq C [\| (\mathbb{E}_Q[\psi_I(\xi(\omega))] - \mu_I) + \| + \| \mathbb{E}_Q[\psi_E(\xi(\omega))] - \mu_E \|]$$

$$\leq C \left[ \| (\mathbb{E}_Q[\psi_I(\xi(\omega))] - \mu_I) + \| + \| \mathbb{E}_Q[\psi_E(\xi(\omega))] - \mu_N \| + \| (\mu_I^N - \mu_I) + \| + \| \mu_E^N - \mu_E \| \right]$$
Proof. Let $Q \in \mathcal{P}_N$. By the Hoffman lemma, there is a constant $C$ such that

\[
d_{TV}(Q, \mathcal{P}) \leq C \left[ \| (E_Q[\psi_I(\xi(\omega))] - \mu_I) \| + \| E_Q[\psi_E(\xi(\omega))] - \mu_E \| \right]
\]

\[
\leq C \left[ \| (E_Q[\psi_I(\xi(\omega))] - \mu_I^N) \| + \| E_Q[\psi_E(\xi(\omega))] - \mu_E^N \| + \| (\mu_I^N - \mu_I) \| + \| \mu_E^N - \mu_E \| \right]
\]

\[
= C \left[ \| (\mu_I^N - \mu_I) \| + \| \mu_E^N - \mu_E \| \right],
\]

because $(a + b)_+ \leq (a)_+ + (b)_+$. 

Proof. Let $Q \in \mathcal{P}_N$. By the Hoffman lemma, there is a constant $C$ such that

$$d_{TV}(Q, \mathcal{P}) \leq C \left[ \| (\mathbb{E}_Q[\psi_I(\xi(\omega))] - \mu_I)_+ \| + \| \mathbb{E}_Q[\psi_E(\xi(\omega))] - \mu_E \| \right]$$

$$\leq C \left[ \| (\mathbb{E}_Q[\psi_I(\xi(\omega))] - \mu^N_I)_+ \| + \| \mathbb{E}_Q[\psi_E(\xi(\omega))] - \mu^N_E \| \right]$$

$$+ \| (\mu^N_I - \mu_I)_+ \| + \| \mu^N_E - \mu_E \|]$$

$$= C [\| (\mu^N_I - \mu_I)_+ \| + \| \mu^N_E - \mu_E \|],$$

because $(a + b)_+ \leq (a)_+ + (b)_+$. This gives

$$\mathbb{D}_{TV}(\mathcal{P}_N, P) = \sup_{Q \in \mathcal{P}_N} d_{TV}(Q, \mathcal{P}) \leq C (\| (\mu^N_I - \mu_I)_+ \| + \| \mu^N_E - \mu_E \|).$$
Likewise,

\[ D_{TV}(\mathcal{P}, \mathcal{P}_N) \leq C(\|\mu_I - \mu_I^N\|_1 + \|\mu_E - \mu_E^N\|) \]
Likewise,

$$\mathbb{D}_{TV}(\mathcal{P}, \mathcal{P}_N) \leq C\left(\| (\mu_I - \mu^N_I) \| + \| \mu_E - \mu^N_E \| \right).$$

Combining the inequalities, we have

$$\mathbb{H}_{TV}(\mathcal{P}_N, \mathcal{P}) \leq C \left[ \max(\| (\mu^N_I - \mu_I) \|, \| (\mu_I - \mu^N_I) \|) + \| \mu^N_E - \mu_E \| \right].$$
Generalizations

\[ \mathcal{P} := \left\{ P : \begin{array}{l} \mathbb{E}_P[\psi_i(\xi)] = \mu_i, \quad \text{for } i = 1, \ldots, p \\ \mathbb{E}_P[\psi_i(\xi)] \leq \mu_i, \quad \text{for } i = p + 1, \ldots, q \end{array} \right\} \]
I. Quantifying change of the ambiguity set

Generalizations

\[ \mathcal{P} := \left\{ P : \begin{align*} \mathbb{E}_P[\psi_i(\xi)] &= \mu_i, \quad \text{for } i = 1, \ldots, p \\ \mathbb{E}_P[\psi_i(\xi)] &\leq \mu_i, \quad \text{for } i = p + 1, \ldots, q \end{align*} \right\} \]

\[ \downarrow \]

- \( \mathcal{P} \) depends on the decision variable \( x \) and other parameter \( u \)

\[ \mathcal{P}(x, u) := \left\{ P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\Psi(x, u, \xi)] \in \mathcal{K} \right\}, \]

where \( \Psi \) is a mapping consisting of matrices and \( \mathcal{K} \) is a cone in the respective matrix spaces.
Generalizations

\[ \mathcal{P} := \left\{ P : \begin{array}{l} \mathbb{E}_P[\psi_i(\xi)] = \mu_i, \quad \text{for } i = 1, \cdots, p \\ \mathbb{E}_P[\psi_i(\xi)] \leq \mu_i, \quad \text{for } i = p + 1, \cdots, q \end{array} \right\} \]

\[ \Downarrow \]

- \( \mathcal{P} \) depends on the decision variable \( x \) and other parameter \( u \)

\[ \mathcal{P}(x, u) := \{ P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\Psi(x, u, \xi)] \in \mathcal{K} \}, \]

where \( \Psi \) is a mapping consisting of matrices and \( \mathcal{K} \) is a cone in the respective matrix spaces.

- Measuring the distance under \( \zeta \)-metric.
There exist $P_0 \in \mathcal{P}(\Xi)$ and a constant $\alpha > 0$ such that
\[ \langle P_0, \psi(x_0, u_0, \xi) \rangle + \alpha \mathbb{B} \subset K, \]
where $\mathbb{B}$ is the unit ball in the space that $K$ is defined.
Hoffman’s lemma under $\zeta$-metric

Proposition 2.2 (Liu, Pichler and Xu (2017))

Under the Slater condition (4)

\[
d_{\mathcal{G}}(Q, \mathcal{P}(x, u)) \leq \frac{\Delta}{\alpha} \inf_{w \in \mathcal{K}} \|w - \langle Q, \psi(x, u, \xi)\rangle\|
\]  

(5)

for any $Q \in \mathcal{P}(\Xi)$ and $(x, u)$ close to $(x_0, u_0)$, where $\alpha$ is the positive constant defined in the Slater condition and

\[
\Delta := \max_{P \in \mathcal{P}(\Xi)} d_{\mathcal{G}}(P, P_0).
\]  

(6)
Hoffman’s lemma under $\zeta$-metric

Proposition 2.2 (Liu, Pichler and Xu (2017))

Under the Slater condition (4)

$$d_G(Q, \mathcal{P}(x, u)) \leq \Delta \inf_{w \in K} \|w - \langle Q, \Psi(x, u, \xi) \rangle\|$$

(5)

for any $Q \in \mathcal{P}(\Xi)$ and $(x, u)$ close to $(x_0, u_0)$, where $\alpha$ is the positive constant defined in the Slater condition and

$$\Delta := \max_{P \in \mathcal{P}(\Xi)} d_G(P, P_0).$$

(6)

Question:

How to estimate $\Delta$?
I. Quantifying change of the ambiguity set

\[ \Delta := \max_{P \in \mathcal{P}(\Xi)} d_{\mathcal{G}}(P, P_0). \]

- \( \Delta \leq 2 \) under the total variation metric \( d_{TV} \) and the Bounded Lipschitz metric \( d_{BL} \);
\[ \Delta := \max_{P \in \mathcal{P}(\Xi)} d_\mathcal{G}(P, P_0). \]

- \( \Delta \leq 2 \) under the total variation metric \( d_{TV} \) and the Bounded Lipschitz metric \( d_{BL} \);
- \( \Delta \leq \text{diam}(\Xi) \) under the Kantorovich/Wasserstein metric when the support set \( \Xi \) is bounded.
Recall

\[ \mathcal{P}(x, u) := \{ P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\Psi(x, u, \xi)] \in \mathcal{K} \} \]

vs

\[ \mathcal{P}(x', u') := \{ P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\Psi(x', u', \xi)] \in \mathcal{K} \} . \]
Theorem 2.1

Assume:

(a) there exist positive constants $\gamma \in \mathbb{R}_+$ and $\nu_1, \nu_2 \in (0, 1]$ such that

$$
\| \Psi(x, u, \xi) - \Psi(x', u', \xi) \| \leq \gamma(\| x - x' \|^{\nu_1} + \| u - u' \|^{\nu_2})
$$

for all $\xi \in \Xi$ and $(x, u), (x', u') \in X \times U$ close to $(x_0, u_0)$. 

Theorem 2.1

Assume:

(a) there exist positive constants $\gamma \in \mathbb{R}_+$ and $\nu_1, \nu_2 \in (0, 1]$ such that

$$
\|\Psi(x, u, \xi) - \Psi(x', u', \xi)\| \leq \gamma (\|x - x'\|^{\nu_1} + \|u - u'\|^{\nu_2})
$$

for all $\xi \in \Xi$ and $(x, u), (x', u') \in X \times U$ close to $(x_0, u_0)$.

(b) $\Xi$ is a compact set.
Theorem 2.1

Assume:

(a) there exist positive constants \( \gamma \in \mathbb{R}_+ \) and \( \nu_1, \nu_2 \in (0, 1] \) such that

\[
\| \Psi(x, u, \xi) - \Psi(x', u', \xi) \| \leq \gamma (\| x - x' \|^{\nu_1} + \| u - u' \|^{\nu_2})
\]

for all \( \xi \in \Xi \) and \( (x, u), (x', u') \in X \times U \) close to \( (x_0, u_0) \).

(b) \( \Xi \) is a compact set.

(c) The Slater condition (4) is fulfilled.
Theorem 2.1

Assume:

(a) there exist positive constants $\gamma \in \mathbb{R}_+$ and $\nu_1, \nu_2 \in (0, 1]$ such that

$$\|\Psi(x, u, \xi) - \Psi(x', u', \xi)\| \leq \gamma(\|x - x'\|^\nu_1 + \|u - u'\|^\nu_2)$$

for all $\xi \in \Xi$ and $(x, u), (x', u') \in X \times U$ close to $(x_0, u_0)$.

(b) $\Xi$ is a compact set.

(c) The Slater condition (4) is fulfilled.

Then there exists a positive constant $C$ such that

$$H_G(\mathcal{P}(x, u), \mathcal{P}(x', u')) \leq C(\|x - x'\|^\nu_1 + \|u - u'\|^\nu_2)$$

(7)

for any $(x, u), (x', u') \in X \times U$ close to $(x_0, u_0)$. 
Let $P \in \mathcal{P}(\Xi)$ and $r$ be a positive number. We call the following set of probability distributions as $\zeta$-ball:

$$
\mathcal{B}(P, r) := \{ P' \in \mathcal{P}(\Xi) : d_\mathcal{G}(P', P) \leq r \},
$$

where

$$
d_\mathcal{G}(P, Q) := \sup_{g \in \mathcal{G}} | \mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)] |.
$$

and $P$ is a nominal distribution which may be an empirical probability distribution.
Quantifying change of the $\zeta$-ball

**Theorem 2.2 (Liu, Pichler and Xu (2017))**

Let $\mathcal{B}(P, r)$ be the $\zeta$-ball defined as in (8). For every $P, Q \in \mathcal{P}(\Xi)$ and $r_1, r_2 \in \mathbb{R}_+$, it holds that

$$
H_G(\mathcal{B}(P, r_1), \mathcal{B}(Q, r_2)) \leq d_G(P, Q) + |r_1 - r_2|,
$$

where $H_G$ denotes the Hausdorff distance in $\mathcal{P}(\Xi)$ associated with $\zeta$-metric.
Part II: Quantitative stability analysis of the DRO
Consider the following distributionally robust problem:

\[
\begin{align*}
\text{(DRO)} & \quad \min \sup_{x, P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] \\
\text{s.t.} & \quad x \in X,
\end{align*}
\]

(10)

and its perturbation

\[
\begin{align*}
\min \sup_{x, P \in \hat{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] \\
\text{s.t.} & \quad x \in X.
\end{align*}
\]

(11)
Distributionally robust formulation

Consider the following distributionally robust problem:

\[ \min_{x} \ \sup_{P \in \mathcal{P}} \mathbb{E}_{P}[f(x, \xi)] \]
\[ \text{s.t. } x \in X, \]  

and its perturbation

\[ \min_{x} \ \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_{P}[f(x, \xi)] \]
\[ \text{s.t. } x \in X. \]  

Question

How does perturbation of the ambiguity set \( \mathcal{P} \) affect the optimal value and the optimal solution?
Stability of optimal value

Theorem 3.1 (Liu, Pichler and Xu (2017))

Let \( \vartheta(\tilde{P}) \) and \( \vartheta(P) \) denote the optimal value of the DRO and its perturbation. Then the following assertions hold:

(i) \[
|\vartheta(\tilde{P}) - \vartheta(P)| \leq H_G(\tilde{P}, P)
\]

where \( H_G \) is the Hausdorff distance under \( \zeta \)-metric with

\[
G := \{f(x, \cdot) : x \in X\}.
\]
In particular, if \( \mathcal{P} = \mathcal{B}(P, r) \), \( \tilde{\mathcal{P}} = \mathcal{B}(\tilde{P}, \tilde{r}) \), then

\[
|\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| \leq d_G(P, \tilde{P}) + |r - \tilde{r}|.
\] (12)
In particular, if $\mathcal{P} = \mathcal{B}(\mathcal{P}, r)$, $\tilde{\mathcal{P}} = \mathcal{B}(\tilde{\mathcal{P}}, \tilde{r})$, then

$$|\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| \leq d_{\mathcal{G}}(\mathcal{P}, \tilde{\mathcal{P}}) + |r - \tilde{r}|. \quad (12)$$

- If the functions in the set $\mathcal{G}$ are Lipschitz continuous with modulus $\kappa$, then

$$|\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| \leq \kappa d_{K}(\mathcal{P}, \tilde{\mathcal{P}}) + |r - \tilde{r}|. \quad (13)$$

where $d_{K}$ denotes the Kantorovich/Wasserstein metric.
In particular, if $\mathcal{P} = \mathcal{B}(P, r)$, $\tilde{\mathcal{P}} = \mathcal{B}(\tilde{P}, \tilde{r})$, then

$$|\vartheta(\tilde{P}) - \vartheta(P)| \leq d_{\mathcal{G}}(P, \tilde{P}) + |r - \tilde{r}|.$$  \hfill (12)

- If the functions in the set $\mathcal{G}$ are Lipschitz continuous with modulus $\kappa$, then
  $$|\vartheta(\tilde{P}) - \vartheta(P)| \leq \kappa d_{K}(P, \tilde{P}) + |r - \tilde{r}|.$$  \hfill (13)

  where $d_{K}$ denotes the Kantorovich/Wasserstein metric.

- If the functions in $\mathcal{G}$ are bounded by a positive constant $C$, then
  $$|\vartheta(\tilde{P}) - \vartheta(P)| \leq C d_{TV}(P, \tilde{P}) + |r - \tilde{r}|.$$  \hfill (14)
II. Quantitative stability analysis of the DRO

Proof. Let

\[ v(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)], \quad \tilde{v}(x) := \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)]. \]
Proof. Let

\[ v(x) := \sup_{P \in \mathcal{P}} E_P[f(x, \xi)], \quad \tilde{v}(x) := \sup_{P \in \tilde{\mathcal{P}}} E_P[f(x, \xi)]. \]

\[ |\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| = \left| \sup_{x \in X} v(x) - \sup_{x \in X} \tilde{v}(x) \right| \leq \sup_{x \in X} |v(x) - \tilde{v}(x)|. \]
II. Quantitative stability analysis of the DRO

Proof. Let

\[ v(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)], \quad \tilde{v}(x) := \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)]. \]

\[ |\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| = \left| \sup_{x \in \mathcal{X}} v(x) - \sup_{x \in \mathcal{X}} \tilde{v}(x) \right| \leq \sup_{x \in \mathcal{X}} \left| v(x) - \tilde{v}(x) \right|. \]

\[ v(x) - \tilde{v}(x) = \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] - \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] \]
Proof. Let

\[ v(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)], \quad \tilde{v}(x) := \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)]. \]

\[ |\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| = \left| \sup_{x \in X} v(x) - \sup_{x \in X} \tilde{v}(x) \right| \leq \sup_{x \in X} |v(x) - \tilde{v}(x)|. \]

\[ v(x) - \tilde{v}(x) = \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] - \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] \]
\[ = \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)]. \]
Proof. Let

\[ \nu(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)], \quad \tilde{\nu}(x) := \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)]. \]

Then

\[ |\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| = \left| \sup_{x \in X} \nu(x) - \sup_{x \in X} \tilde{\nu}(x) \right| \leq \sup_{x \in X} |\nu(x) - \tilde{\nu}(x)|. \]

Also,

\[ \nu(x) - \tilde{\nu}(x) = \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] - \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] \]

\[ = \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)] \]

\[ \leq \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \sup_{x \in X} |\mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)]|. \]
Proof. Let

\[ \nu(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)], \quad \tilde{\nu}(x) := \sup_{\tilde{P} \in \tilde{\mathcal{P}}} \mathbb{E}_{\tilde{P}}[f(x, \xi)]. \]

\[ |\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| = \left| \sup_{x \in X} \nu(x) - \sup_{x \in X} \tilde{\nu}(x) \right| \leq \sup_{x \in X} \left| \nu(x) - \tilde{\nu}(x) \right|. \]

\[ \nu(x) - \tilde{\nu}(x) = \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] - \sup_{\tilde{P} \in \tilde{\mathcal{P}}} \mathbb{E}_{\tilde{P}}[f(x, \xi)] \]
\[ = \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)] \]
\[ \leq \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \sup_{x \in X} \mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)] \]
\[ = \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} d_{\mathcal{G}}(P, \tilde{P}) \]
II. Quantitative stability analysis of the DRO

**Proof.** Let

\[ v(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)], \quad \tilde{v}(x) := \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)]. \]

\[ |\vartheta(\tilde{\mathcal{P}}) - \vartheta(\mathcal{P})| = \left| \sup_{x \in X} v(x) - \sup_{x \in X} \tilde{v}(x) \right| \leq \sup_{x \in X} |v(x) - \tilde{v}(x)|. \]

\[ v(x) - \tilde{v}(x) = \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] - \sup_{P \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] \]
\[ = \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)] \]
\[ \leq \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} \sup_{x \in X} |\mathbb{E}_P[f(x, \xi)] - \mathbb{E}_{\tilde{P}}[f(x, \xi)]| \]
\[ = \sup_{P \in \mathcal{P}} \inf_{\tilde{P} \in \tilde{\mathcal{P}}} d_{\mathbb{Q}}(P, \tilde{P}) \]
\[ = D_{\mathbb{Q}}(\mathcal{P}, \tilde{\mathcal{P}}), \]
Stability of the optimal solutions

Theorem 3.1

(ii) If, in addition, \( \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] \) satisfies the **second order growth condition** at \( X^*(\mathcal{P}) \), that is, there exist positive constants \( C \) and \( \sigma \) such that

\[
\sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] \geq \vartheta(\mathcal{P}) + \sigma d(x, X^*(\mathcal{P}))^2 \quad \forall \ x \in X,
\]

then

\[
\mathbb{D}(X^*(\tilde{\mathcal{P}}), X^*(\mathcal{P})) \leq \sqrt{\frac{3}{\sigma} H_{\mathcal{G}}(\tilde{\mathcal{P}}, \mathcal{P})}.
\]
II. Quantitative stability analysis of the DRO

Stability of the optimal solutions

Theorem 3.1

(ii) If, in addition, \( \sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] \) satisfies the second order growth condition at \( X^*(\mathcal{P}) \), that is, there exist positive constants \( C \) and \( \sigma \) such that

\[
\sup_{P \in \mathcal{P}} \mathbb{E}_P[f(x, \xi)] \geq \vartheta(\mathcal{P}) + \sigma d(x, X^*(\mathcal{P}))^2 \quad \forall \ x \in \mathcal{X},
\]

then

\[
\mathbb{D}(X^*(\tilde{\mathcal{P}}), X^*(\mathcal{P})) \leq \sqrt{\frac{3}{\sigma}} \mathbb{H}_G(\tilde{\mathcal{P}}, \mathcal{P}).
\] (15)

A sufficient condition is that there exists a positive function \( \alpha(\xi) \) with \( \inf_{P \in \mathcal{P}} \mathbb{E}_P[\alpha(\xi)] > 0 \) such that

\[
f(x', \xi) \geq f(x, \xi) + \alpha(\xi)\|x' - x\|^2 \quad \forall \ x' \in \mathcal{X}, \xi \in \Xi.
\] (16)
Part III: Mathematical program with distributionally robust chance constraint (MPDRCC)

Guo, Xu and Zhang (2017)
We consider mathematical program with distributionally robust chance constraint:

\[
\begin{align*}
\text{(MPDRCC)} & : \quad \min_{x \in X} f(x) \\
\text{s.t.} & \quad \inf_{P \in \mathcal{P}} P(g(x, \xi) \leq 0) \geq 1 - \beta,
\end{align*}
\]

where \( \mathcal{P} \) is a set of distributions which contains/approximates the true probability distribution of random variable \( \xi \).
We consider mathematical program with distributionally robust chance constraint:

\[
\begin{align*}
\text{min} & \quad f(x) \\
\text{s.t.} & \quad \inf_{P \in \mathcal{P}} P(g(x, \xi) \leq 0) \geq 1 - \beta,
\end{align*}
\]  

(17)

where \( \mathcal{P} \) is a set of distributions which contains/approximates the true probability distribution of random variable \( \xi \).

Approximation of MPDRCC

\[
\text{(MPDRCC}_N) \quad \min_{x \in X} \quad f(x) \\
\text{s.t.} \quad \inf_{P \in \mathcal{P}_N} P(g(x, \xi) \leq 0) \geq 1 - \beta.
\]
Approximation of MPDRCC

\[ \begin{align*}
    \text{(MPDRCC}_N) & \quad \min_{x \in X} f(x) \\
    \text{s.t.} & \quad \inf_{P \in \mathcal{P}_N} P(g(x, \xi) \leq 0) \geq 1 - \beta.
\end{align*} \]  \quad (18)

Question:
What is the impact on the optimal value and the optimal solutions of MPDRCC as \( N \) increases?
Reformulation of the chance constraint

For each fixed $x \in X$, let

$$H(x) := \{ z \in \Xi : g(x, z) \leq 0 \}.$$

Then

$$P(g(x, \xi) \leq 0) \geq 1 - \beta \iff P(H(x)) \geq 1 - \beta.$$
Reformulation of the chance constraint

For each fixed $x \in X$, let

$$H(x) := \{ z \in \Xi : g(x, z) \leq 0 \}. $$

Then

$$ P(g(x, \xi) \leq 0) \geq 1 - \beta \iff P(H(x)) \geq 1 - \beta. $$

Let

$$ \mathbb{1}_{H(x)}(z) := \begin{cases} 
1 & \text{for } z \in H(x), \\
0 & \text{for } z \notin H(x), 
\end{cases} $$

denote the indicator function of $H(x)$. Then

$$ P(H(x)) = \mathbb{E}_P[\mathbb{1}_{H(x)}(\xi)]. $$
III. MPDRCC

\[(\text{MPDRCC}) \quad \min_{x \in X} \; f(x) \quad \text{s.t.} \quad v(x) := \inf_{P \in \mathcal{P}} \mathbb{E}_{P}[\mathbbm{1}_{H(x)}(\xi)] \geq 1 - \beta,\]

VS

\[(\text{MPDRCC}_N) \quad \min_{x \in X} \; f(x) \quad \text{s.t.} \quad v_N(x) := \inf_{P \in \mathcal{P}_N} \mathbb{E}_{P}[\mathbbm{1}_{H(x)}(\xi)] \geq 1 - \beta,\]

We call \(v(x)\) and \(v_N(x)\) robust probability function.
We call \( v(x) \) and \( v_N(x) \) robust probability function.
(MPDRCC) \[
\begin{align*}
\min_{x \in X} & \quad f(x) \\
\text{s.t.} & \quad v(x) := \inf_{P \in \mathcal{P}} \mathbb{E}_P[1_{H(x)}(\xi)] \geq 1 - \beta,
\end{align*}
\]

vs

(MPDRECC$_N$) \[
\begin{align*}
\min_{x \in X} & \quad f(x) \\
\text{s.t.} & \quad v_N(x) := \inf_{P \in \mathcal{P}_N} \mathbb{E}_P[1_{H(x)}(\xi)] \geq 1 - \beta,
\end{align*}
\]

We call $v(x)$ and $v_N(x)$ robust probability function.

**Question:** Are $v(x)$, $v_N(x)$ continuous? Does $v_N(x)$ converge to $v(x)$? How?
Consider the following set of random indicator functions
\[ \mathcal{G} := \{ \mathbb{1}_{H(x)}(\xi(\cdot)) : x \in X \}. \]

For \( P, Q \in \mathcal{P}(\Xi) \), let
\[ \mathcal{D}(P, Q) := \sup_{g \in \mathcal{G}} \left| \mathbb{E}_P[g] - \mathbb{E}_Q[g] \right| = \sup_{x \in X} \left| P(H(x)) - Q(H(x)) \right|. \]

We call \( \mathcal{D}(P, Q) \) pseudo-metric.
Assumption 4.1 (Convergence of the ambiguity set under the pseudo-metric)

The ambiguity sets $\mathcal{P}$ and $\mathcal{P}_N$ satisfy the following conditions:

(a) $\lim_{N \to \infty} \mathcal{D}(\mathcal{P}_N, \mathcal{P}) = 0$, (Outer semi-convergence)

(b) $\lim_{N \to \infty} \mathcal{D}(\mathcal{P}, \mathcal{P}_N) = 0$. (Inner semi-convergence)

Comment: A combination of (a) and (b) implies $\lim_{N \to \infty} \mathcal{H}(\mathcal{P}, \mathcal{P}_N) = 0$. 

III. MPDRCC
Abstract conditions for convergence and continuity

Assumptions
Theorem 4.1 (Uniform convergence)

*Under Assumption 4.1, i.e., \( \lim_{N \to \infty} \mathcal{H}(\mathcal{P}, \mathcal{P}_N) = 0 \), \( v_N(x) \) converges to \( v(x) \) uniformly over \( X \) as \( N \) tends to \( \infty \), i.e.,*

\[
\lim_{N \to \infty} \sup_{x \in X} |v_N(x) - v(x)| = 0.
\]
Proof.

\[ v_N(x) - v(x) = \inf_{P_N \in \mathcal{P}_N} P_N(H(x)) - \inf_{P \in \mathcal{P}} P(H(x)) \]

\[ = \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} P_N(H(x)) - P(H(x)) \]

\[ \leq \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} \sup_{x \in \mathcal{X}} \left| P(H(x)) - P_N(H(x)) \right| \]

\[ \leq \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} \sup_{x \in \mathcal{X}} \left| P(H(x)) - P_N(H(x)) \right| \]

\[ = \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} \sup_{g \in \mathcal{G}} \left| \mathbb{E}_P[g] - \mathbb{E}_{P_N}[g] \right| \]

\[ = \mathcal{D}(\mathcal{P}, \mathcal{P}_N). \]
Proof.

\[ v_N(x) - v(x) = \inf_{P_N \in \mathcal{P}_N} P_N(H(x)) - \inf_{P \in \mathcal{P}} P(H(x)) \]

\[ = \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} P_N(H(x)) - P(H(x)) \]

\[ \leq \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} \left| P(H(x)) - P_N(H(x)) \right| \]

\[ \leq \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} \sup_{x \in X} \left| E_P[g] - E_{P_N}[g] \right| \]

\[ = \sup_{P \in \mathcal{P}} \inf_{P_N \in \mathcal{P}_N} \sup_{g \in \mathcal{G}} \left| E_P[g] - E_{P_N}[g] \right| \]

This shows

\[ \sup_{x \in X} [v_N(x) - v(x)] \leq \mathcal{D}(\mathcal{P}, \mathcal{P}_N). \]
Stability/convergence of MPDRCC

- $\mathcal{F}$ and $\mathcal{F}_N$ denote the feasible set,
- $\vartheta := \inf \{ f(x) : x \in \mathcal{F} \}$ the optimal value,
- $\vartheta_N := \inf \{ f(x) : x \in \mathcal{F}_N \}$ the optimal value,
- $S := \{ x \in \mathcal{F} : \vartheta = f(x) \}$ the set of optimal solutions,
- $S_N := \{ x \in \mathcal{F}_N : \vartheta_N = f(x) \}$ the set of optimal solutions,
- $\mathcal{F}^s := \{ x \in X : v(x) > 1 - \beta \}$ strict feasible solution.
Assumption 4.2 (Continuity of robust probability function)

\[ \nu(\cdot) = \inf_{P \in \mathcal{P}} P(H(\cdot)) \text{ is continuous over } X. \]
Convergence of the optimal value and optimal solutions

Theorem 4.2

Suppose: (a) $\lim_{N \to \infty} \mathcal{H}(\mathcal{P}, \mathcal{P}_N) = 0$; (b) $\nu(x)$ is continuous; (c) $\text{cl} \mathcal{F}^s \cap S \neq \emptyset$. Then

(i) $\limsup_{N \to \infty} \mathcal{F}_N \subset \mathcal{F}$;
(ii) $\lim_{N \to \infty} \vartheta_N = \vartheta$;
(iii) $\limsup_{N \to \infty} S_N \subset S$. 
Sufficient for continuity of the robust probability function

Question

- Under what conditions is the robust probability function
  \[ \nu(x) := \inf_{P \in \mathcal{P}} P(H(x)) \]
  continuous? Recall that \( H(x) := \{z \in \Xi : g(x, z) \leq 0\} \).
Question

- Under what conditions is the robust probability function

\[ v(x) := \inf_{P \in \mathcal{P}} P(H(x)) \]

continuous? Recall that \( H(x) := \{ z \in \Xi : g(x, z) \leq 0 \} \).

- When is \( P(H(x)) \) continuous w.r.t \( x \)?
Conditions for continuity of $P(H(x))$

Condition 4.1

For $H(x) := \{ z \in \Xi : g(x,z) \leq 0 \}$, $K(x) := \{ z \in \Xi : g(x,z) = 0 \}$ and $P \in \mathcal{P}(\Xi)$,

(C1) $P(K(x)) = 0$ for any $x \in X$;

(C2) $H(\cdot)$ is continuous and convex-valued over $X$ and for any $x \in X$,

$$P(\text{bd } H(x)) = 0. \quad (19)$$

Note that $K(x) = \text{bd } H(x)$ when $g(x,\cdot)$ is strictly convex and $\Xi = \mathbb{R}^k$. 
Conditions for continuity of $P(H(x))$

Condition 4.1

For $H(x) := \{z \in \Xi : g(x, z) \leq 0\}$, $K(x) := \{z \in \Xi : g(x, z) = 0\}$ and $P \in \mathcal{P}(\Xi)$,

(C1) $P(K(x)) = 0$ for any $x \in X$;

(C2) $H(\cdot)$ is continuous and convex-valued over $X$ and for any $x \in X$,

$$P(\text{bd } H(x)) = 0.$$  \hspace{1cm} (19)

Note that $K(x) = \text{bd } H(x)$ when $g(x, \cdot)$ is strictly convex and $\Xi = \mathbb{R}^k$.

Theorem 4.3

Let $P \in \mathcal{P}(\Xi)$. Then $P(H(\cdot))$ is continuous over $X$ when either condition (C1) or (C2) is fulfilled.
Why do we need C2?

Example 4.1

Let $\xi : \Omega \rightarrow \mathbb{R}$ be a random variable with support set $\Xi = \mathbb{R}$. Let

$$g(x, z) := \begin{cases} 
  z + x & \text{for } z \geq -x, \\
  0 & \text{for } z \in [-x - 1, -x], \\
  z + x + 1 & \text{for } z \leq -x - 1.
\end{cases}$$
III. MPDRCC

Sufficient for continuity of the robust probability function

\[ H(x) = \{ \xi : g(x, \xi) \leq 0 \} \]

\[ \text{bd } H(x) = R - x - x - 1 \]

- \( H(x) = \{ z \in \mathbb{R} : g(x, z) \leq 0 \} = (-\infty, -x] \).
- \( H(\cdot) \) is convex set-valued, continuous and \( P(\text{bd } H(x)) = 0 \).
- So (C2) is satisfied!
III. MPDRCC

Sufficient for continuity of the robust probability function

$$K(x) = \begin{cases} g(x, z) = 0 & \text{if } z \in {-x, -x-1} \\ \Xi = \mathbb{R} & \text{otherwise} \end{cases}$$

- $K(x) := \{z \in \mathbb{R} : g(x, z) = 0\} = [-x - 1, -x]$.
- $P(K(x)) \neq 0$.
- (C1) is failed!
Pointwise continuity of the robust probability function

Theorem 4.4 (Continuity of $v(x) := \inf_{P \in \mathcal{P}} P(H(x))$)

Suppose $\mathcal{P}$ is weakly compact and one of the following condition holds:
Pointwise continuity of the robust probability function

Theorem 4.4 (Continuity of $v(x) := \inf_{P \in \mathcal{P}} P(H(x)))$

Suppose $\mathcal{P}$ is weakly compact and one of the following condition holds:

(a) (C1) holds for each $P \in \mathcal{P}$ and for each $x \in X$, $g(\cdot, \xi)$ is continuous at $x$ uniformly w.r.t. $\xi \in \Xi$;
III. MPDRCC

Pointwise continuity of the robust probability function

Theorem 4.4 (Continuity of $v(x) := \inf_{P \in \mathcal{P}} P(H(x))$)

Suppose $\mathcal{P}$ is weakly compact and one of the following condition holds:

(a) (C1) holds for each $P \in \mathcal{P}$ and for each $x \in X$, $g(\cdot, \xi)$ is continuous at $x$ uniformly w.r.t. $\xi \in \Xi$;

(b) (C2) holds for each $P \in \mathcal{P}$.

Then $v(\cdot)$ is continuous on $X$. 

The set $A$ is said to be weakly compact if every sequence $\{P_N\} \subset A$ contains a subsequence $\{P_{N'}\}$ and moreover there exists $P \in A$ such that $P_{N'}$ converges to $P$ weakly.
III. MPDRCC

Sufficient for continuity of the robust probability function

Pointwise continuity of the robust probability function

Theorem 4.4 (Continuity of $v(x) := \inf_{P \in \mathcal{P}} P(H(x)))$

Suppose $\mathcal{P}$ is weakly compact and one of the following condition holds:

(a) $(C1)$ holds for each $P \in \mathcal{P}$ and for each $x \in X$, $g(\cdot, \xi)$ is continuous at $x$ uniformly w.r.t. $\xi \in \Xi$;

(b) $(C2)$ holds for each $P \in \mathcal{P}$.

Then $v(\cdot)$ is continuous on $X$.

The set $\mathcal{A}$ is said to be weakly compact if every sequence $\{P_N\} \subset \mathcal{A}$ contains a subsequence $\{P_{N'}\}$ and moreover there exists $P \in \mathcal{A}$ such that $P_{N'}$ converges to $P$ weakly.
Lemma 4.1

Let \( \{P_N\} \subset \mathcal{P} \) be a sequence of probability measures and \( P \in \mathcal{P} \). Suppose \( P_N \) converges to \( P \) weakly. Then

\[
\lim_{N \to \infty} \mathcal{D}(P_N, P) = 0
\]

under one of the following conditions:

(a) \( g(\cdot, \xi) \) is continuous on \( X \) uniformly w.r.t. \( \xi \in \Xi \) + condition (C1) for \( P \).
Lemma 4.1

Let \( \{ P_N \} \subset \mathcal{P} \) be a sequence of probability measures and \( P \in \mathcal{P} \). Suppose \( P_N \) converges to \( P \) weakly. Then

\[
\lim_{N \to \infty} \mathcal{D}(P_N, P) = 0
\]

under one of the following conditions:

(a) \( g(\cdot, \xi) \) is continuous on \( X \) uniformly w.r.t. \( \xi \in \Xi \) + condition (C1) for \( P \).

(b) Condition (C2) for the \( P \).
Sufficient conditions for convergence of $D(P_N, P)$

Proposition 4.1

- $P_N$ converges to $P$ weakly, i.e., for every sequence $\{P_N\} \subseteq P_N$, $\{P_N\}$ has a subsequence $\{P_{N_k}\}$ converging to $P$ with $P \in P$. 
Sufficient conditions for convergence of $\mathcal{D}(\mathcal{P}_N, \mathcal{P})$

Proposition 4.1

- $\mathcal{P}_N$ converges to $\mathcal{P}$ weakly, i.e., for every sequence $\{P_N\} \subseteq \mathcal{P}_N$, $\{P_N\}$ has a subsequence $\{P_{N_k}\}$ converging to $P$ with $P \in \mathcal{P}$.
- Condition (a) or (b) in Lemma 4.1 holds for any $P \in \mathcal{P}$.

Then

$$\lim_{N \to \infty} \mathcal{D}(\mathcal{P}_N, \mathcal{P}) = 0.$$
Sufficient conditions for convergence of $\mathcal{D}(\mathcal{P}, \mathcal{P}_N)$

Proposition 4.2

Assume:

- $\mathcal{P}$ is weakly compact, i.e., for any $P \in \mathcal{P}$, there exists a sequence $\{P_N\} \in \mathcal{P}_N$ such that $P_N$ converges to $P$ weakly.
Proposition 4.2

Assume:

- \( \mathcal{P} \) is weakly compact, i.e., for any \( P \in \mathcal{P} \), there exists a sequence \( \{P_N\} \in \mathcal{P}_N \) such that \( P_N \) converges to \( P \) weakly.
- Condition (a) or (b) in Lemma 4.1 holds for any \( P \in \mathcal{P} \).

Then

\[
\lim_{N \to \infty} D(\mathcal{P}, \mathcal{P}_N) = 0.
\]


Y. Liu, A. Pichler and H. Xu, Discrete Approximation and Quantification in Distributionally Robust Optimization, to appear in MOR.