Bootstrap Robust Analytics

Dimitris Bertsimas, **Bart P.G. Van Parys**

MIT Sloan School of Management

06 Feb. 2018

Stochastic Optimization

Decision Problem : $z^\star \in \arg\min_{z \in Z} \mathbf{E}_{\mathbb{M}^\star} [L(z, Y)]$ Cost Optimal

Portfolio Management

- z : Financial position
- Y : Market prices
- L : Risk / return

Autonomous Driving

- z : Travel route
- Y : Traffic
- L : Travel time

Stochastic Optimization

Decision Problem : $z^\star \in \arg\min_{z \in Z} \mathbf{E}_{\mathbb{M}^\star} [L(z, Y) | X = \bar{x}]$ Cost Optimal

Portfolio Management

- z : Financial position
- Y : Market prices
- L : Risk / return
- X : Oil, Twitter chatter,

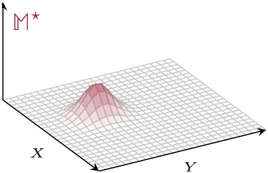
Autonomous Driving

- z : Travel route
- Y : Traffic
- L : Travel time
- X : Holiday, Weather,

Stochastic Optimization

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z,Y)|X=\bar{x}\right]$ Cost Optimal

Portfolio Management



- z : Financial position
- Y : Market prices
- L : Risk / return
- X : Oil, Twitter chatter,

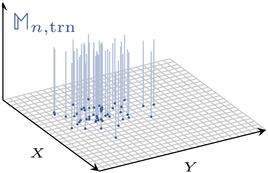
Autonomous Driving

- z : Travel route
- Y : Traffic
- L : Travel time
- X : Holiday, Weather,

Stochastic Optimization

Decision Problem : $z^\star \in \arg\min_{z \in Z} \mathbf{E}_{\mathbb{M}^\star}[L(z, Y) | X = \bar{x}]$ Cost Optimal

Portfolio Management



Autonomous Driving

- z : Financial position
- Y : Market prices
- L : Risk / return
- X : Oil, Twitter chatter,

- z : Travel route
- Y : Traffic
- L : Travel time
- X : Holiday, Weather,

Supervised Training Data :

- $X := (x_1, \dots, x_n) \in \mathrm{R}^{n \times p}$
- $Y := (y_1, \dots, y_n) \in \mathrm{R}^{n}$

Prescriptive Analytics : Formulate decisions based on **data**, not **distributions**.

Naive Formulations

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y) | X = \bar{x}\right]$ Cost Optimal

Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

Naive Formulations

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z,Y)|X=\bar{x}\right]$ Cost Optimal

Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Sample Average Formulation :**

$$z^{\mathrm{saa}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \sum_{(x_i,y_i)} L(z,y_i)\cdot \frac{1}{n}$$

Fails to learn from covariate information $X=\bar{x}$.

Naive Formulations

Decision Problem : $z^\star \in \arg\min_{z \in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y) | X = \bar{x}\right]$ Cost Optimal

Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Sample Average Formulation :**

$$z^{\mathrm{saa}}_{n,\mathrm{trn}} \in \arg\min_{z \in Z} \sum_{(x_i, y_i)} L(z, y_i) \cdot \frac{1}{n}$$

Fails to learn from covariate information $X = \bar{x}$.

- **Empirical Formulation :**

$$z^{\mathrm{emp}}_{n,\mathrm{trn}} \in \arg\min_{z \in Z} \sum_{(x_i = \bar{x}, y_i)} L(z, y_i) \cdot \frac{1}{n}$$

Fails to generalize to covariate information $X = \bar{x} \neq x_i$.

Naive Formulations

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z,Y)|X=\bar{x}\right]$ Cost Optimal

Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Sample Average Formulation :**

$$z^{\mathrm{saa}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \sum_{(x_i,y_i)} L(z,y_i)\cdot \frac{1}{n}$$

Fails to learn from covariate information $X=\bar{x}$.

- **Empirical Formulation :**

$$z^{\mathrm{emp}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \sum_{(x_i=\bar{x},y_i)} L(z,y_i)\cdot \frac{1}{n}$$

Fails to generalize to covariate information $X=\bar{x}\neq x_i$.

- **ML Formulations** : Learn and generalize

Bertsimas and Kallus. "From predictive to prescriptive analytics." (2014)

Nearest Neighbors Formulation

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y) | X = \bar{x}\right]$ Cost Optimal

Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Nearest Neighbors Formulation :**

$$z^{\mathrm{nn}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \quad \sum_{(x_i,y_i)} L(z, y_i) \cdot \mathbb{1}\{x_i \in N_k(\bar{x})\} \cdot s$$
$$\text{where} \quad 1 = \sum_{(x_i.y_i)} \mathbb{1}\{x_i \in N_k(\bar{x})\} \cdot s$$

where $N_k(\bar{x})$ the k nearest observations to $\bar{x}$.

Nearest Neighbors Formulation

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}[L(z,Y)|X=\bar{x}]$ Cost Optimal

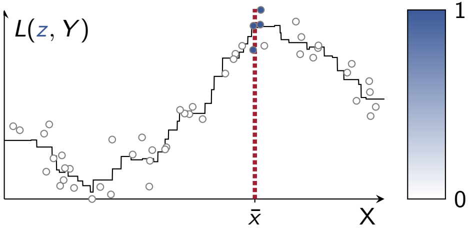
Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Nearest Neighbors Formulation :**

$$z^{\mathrm{nn}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \quad \sum_{(x_i,y_i)} L(z,y_i)\cdot \mathbb{1}\{x_i \in N_k(\bar{x})\}\cdot s$$
$$\text{where} \quad 1 = \sum_{(x_i.y_i)} \mathbb{1}\{x_i \in N_k(\bar{x})\}\cdot s$$

where $N_k(\bar{x})$ the k nearest observations to $\bar{x}$.

- Budgets cost of decisions in context $X=\bar{x}$ based on **k-NN**



Nadaraya-Watson Formulation

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y)|X = \bar{x}\right]$ Cost Optimal

Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Nadaraya-Watson Formulation :**

$$z^{\mathrm{nw}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \quad \sum_{(x_i, y_i)} L(z, y_i) \cdot S_n(x_i, \bar{x}) \cdot s$$
$$\text{where} \quad 1 = \sum_{(x_i . y_i)} S_n(x_i, \bar{x}) \cdot s$$

where $S(x, \bar{x})$ a smoothing function.

Nadaraya-Watson Formulation

Decision Problem : $z^\star \in \arg\min_{z \in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y) | X = \bar{x}\right]$ Cost Optimal

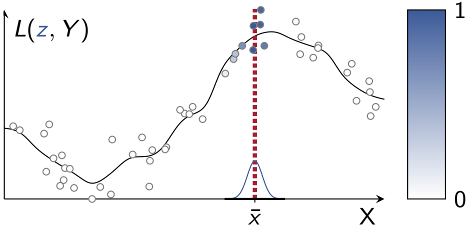
Formulate decision $z_{n,\mathrm{trn}}$ based on data, not distributions.

- **Nadaraya-Watson Formulation :**

$$
\begin{aligned}
z^{\mathrm{nw}}_{n,\mathrm{trn}} \in \arg\min_{z \in Z} \quad & \textstyle\sum_{(x_i, y_i)} L(z, y_i) \cdot S_n(x_i, \bar{x}) \cdot s \\
\text{where} \quad & 1 = \textstyle\sum_{(x_i . y_i)} S_n(x_i, \bar{x}) \cdot s
\end{aligned}
$$

where $S(x, \bar{x})$ a smoothing function.

- Budgets cost of decisions in context $X = \bar{x}$ based on **kernel smoothing**



Prescriptive Analytics

Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y)|X = \bar{x}\right]$ Cost Optimal

$\mathbb{M}_{n,\mathrm{trn}}$

$$\begin{cases} \text{Budget :} & c^{\mathrm{nn}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) \\ \text{Decision :} & z^{\mathrm{nn}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z}\ c^{\mathrm{nn}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) \end{cases}$$

$\mathbb{M}_{n,\mathrm{trn}}$

$$\begin{cases} \text{Budget :} & c^{\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) \\ \text{Decision :} & z^{\mathrm{nw}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z}\ c^{\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) \end{cases}$$

Bertsimas and Kallus. "From predictive to prescriptive analytics." (2014)

Decision Problem : $z^\star \in \arg\min_{z \in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z, Y) | X = \bar{x}\right]$ Cost Optimal

Prescriptive Analytics : $z_{n,\mathrm{trn}} \in \arg\min_{z \in Z} c(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$ Budget Optimal

$\mathbb{M}_{n,\mathrm{trn}}$

Budget : $c^{\mathrm{nn}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

Decision : $z^{\mathrm{nn}}_{n,\mathrm{trn}} \in \arg\min_{z \in Z} c^{\mathrm{nn}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

$\mathbb{M}_{n,\mathrm{trn}}$

Budget : $c^{\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

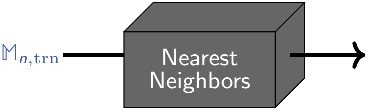
Decision : $z^{\mathrm{nw}}_{n,\mathrm{trn}} \in \arg\min_{z \in Z} c^{\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

Bertsimas and Kallus. "From predictive to prescriptive analytics." (2014)

Prescriptive Analytics

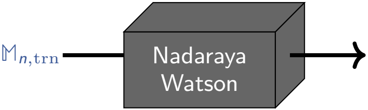
Decision Problem : $z^\star \in \arg\min_{z\in Z} \mathbf{E}_{\mathbb{M}^\star}\left[L(z,Y)|X=\bar{x}\right]$ Cost Optimal

Prescriptive Analytics : $z_{n,\mathrm{trn}} \in \arg\min_{z\in Z} c(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$ Budget Optimal



Budget : $c^{\mathrm{nn}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

Decision : $z^{\mathrm{nn}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} c^{\mathrm{nn}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$



Budget : $c^{\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

Decision : $z^{\mathrm{nw}}_{n,\mathrm{trn}} \in \arg\min_{z\in Z} c^{\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})$

Bertsimas and Kallus. “From predictive to prescriptive analytics.” (2014)

Comparison:

+ Decision $z_{n,\mathrm{trn}}$ based on data, not distributions.
– Budget optimal $\neq$ cost optimal

Robust Prescriptive Analytics

Decisions $z_{n,\mathrm{trn}}$ based on a **nominal budget** c often **disappoint** on **test data** :

$$c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{tst}}}^{\text{test}}, \bar{x}) > c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x}).$$

Better known as the "*optimizer's curse*".

We propose making prescriptions $z^r_{n,\mathrm{trn}}$ *based on a* **robust budget** c^r*.*

Decisions $z_{n,\mathrm{trn}}$ based on a **nominal budget** c often **disappoint** on **test data** :

$$c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{tst}}}^{\mathrm{test}}, \bar{x}) > c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\mathrm{training}}, \bar{x}).$$

Better known as the "*optimizer's curse*".

We propose making prescriptions $z^r_{n,\mathrm{trn}}$ *based on a* **robust budget** c^r*.*

- Statistical robustness, i.e.,

$$\mathbb{M}^{\star n}\left(c(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{tst}}}^{\mathrm{test}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\mathrm{training}}, \bar{x})\right) \leq b,$$

requires "stochastic data"; a contradiction in terms.

Decisions $z_{n,\mathrm{trn}}$ based on a **nominal budget** c often **disappoint** on **test data** :

$$c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{tst}}}^{\text{test}}, \bar{x}) > c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x}).$$

Better known as the "*optimizer's curse*".

We propose making prescriptions $z^r_{n,\mathrm{trn}}$ *based on a* **robust budget** c^r.

- Statistical robustness, i.e.,

$$\mathbb{M}^{\star n}\left(c(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{tst}}}^{\text{test}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

  requires "stochastic data"; a contradiction in terms.

- **Bootstrap robustness**, i.e.,

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

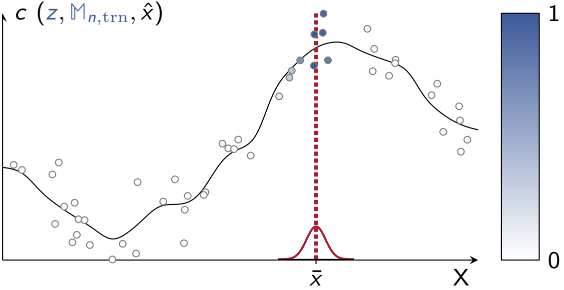
  is the next best thing. Bootstrap data is **synthetic** test data drawn with replacement from training data.

A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

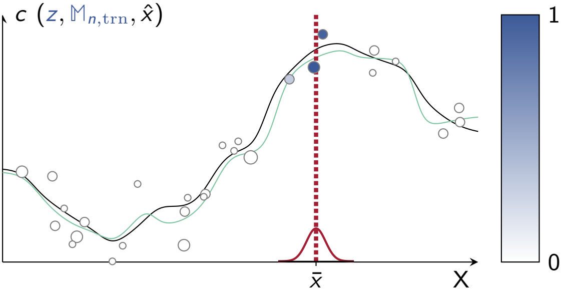


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

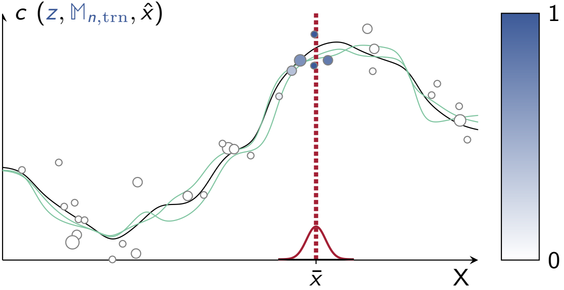


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

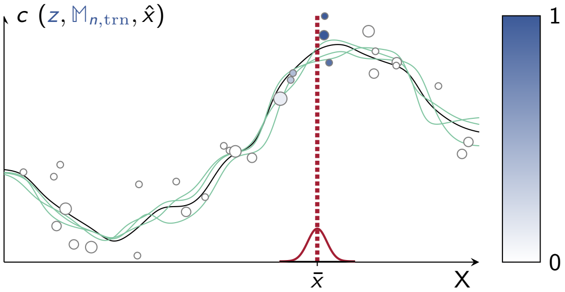


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\Big(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\Big) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

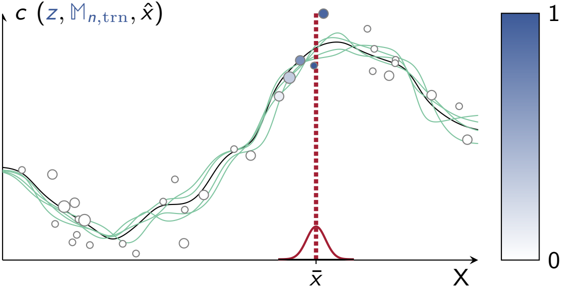


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

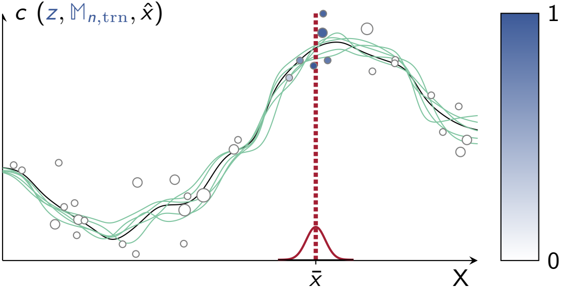


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\Big(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\Big) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

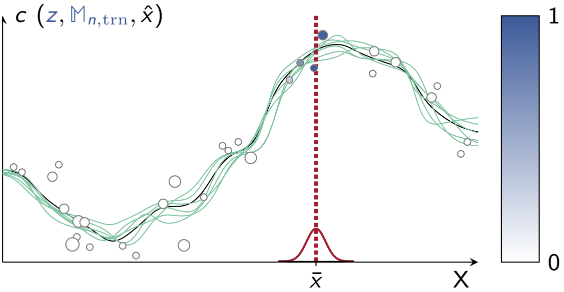


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x}) \right) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

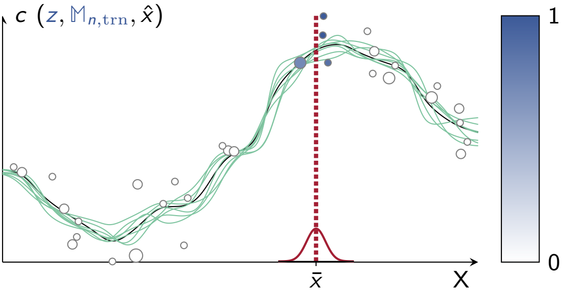


A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

Nominal vs Robust Budgets:

(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$



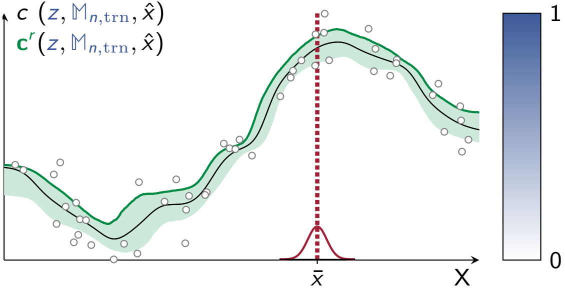
A budget c^r is **bootstrap robust** relative to a nominal budget c if

$$\mathbb{M}^n_{n,\mathrm{trn}}\left(c(z_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{training}}, \bar{x})\right) \leq b,$$

Nominal vs Robust Budgets:

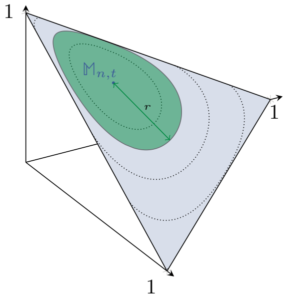
(N) Cost (over)-calibrated to given training data $\mathbb{M}_{n,\mathrm{trn}}$

(R) Cost calibrated to bootstrap data $\mathbb{M}_{n,\mathrm{btsp}}$ as well.



Robust Predictive Analytics :

$$z^r_{n,\mathrm{trn}} \in \arg\min_{z\in Z}\; c^r(z,\mathbb{M}_{n,\mathrm{trn}},\bar{x}) := \max_{\mathbb{M}}\; c(z,\mathbb{M},\bar{x}) \quad \text{s.t. } B(\mathbb{M},\mathbb{M}_{n,\mathrm{trn}}) \leq r.$$



DRO counterpart with respect to **convex set**

$$\{\mathbb{M} \;:\; B(\mathbb{M},\mathbb{M}_{n,\mathrm{trn}}) \leq r\}\,.$$

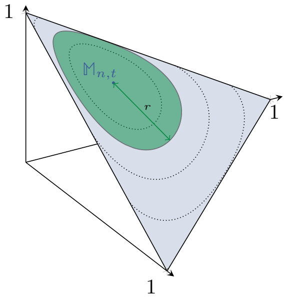
Theorem 1 :

Robust Predictive Analytics :

$$z^r_{n,\mathrm{trn}} \in \arg\min_{z\in Z} c^r(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) := \max_{\mathbb{M}} c(z, \mathbb{M}, \bar{x}) \quad \text{s.t. } B(\mathbb{M}, \mathbb{M}_{n,\mathrm{trn}}) \leq r.$$

DRO counterpart with respect to **convex set**

$$\{\mathbb{M} \ : \ B(\mathbb{M}, \mathbb{M}_{n,\mathrm{trn}}) \leq r\}.$$



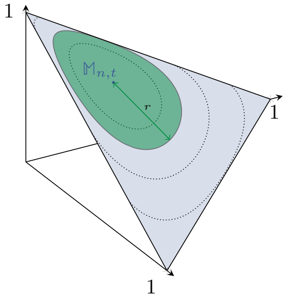
Theorem 1 :

- Both robust NN and NW are **as tractable their nominal counterparts**. E.g.

$$\begin{aligned} z^{r,\mathrm{nw}}_{n,t} \in \arg\min_{z\in Z} c^{r,\mathrm{nw}}(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) = \max_{\mathbb{P},\, s>0} \ & \textstyle\sum_{(x_i,y_i)} S_n(x_i,\bar{x}) \cdot L(z, y_i) \cdot \mathbb{P}(x_i, y_i) \\ \text{s.t. } & \textstyle\sum_{(x_i,y_i)} \mathbb{P}(x_i,y_i) = s, \\ & \textstyle\sum_{(x_i,y_i)} S_n(x_i,\bar{x}) \cdot \mathbb{P}(x_i,y_i) = 1, \\ & s \cdot B(\mathbb{P}/s, \mathbb{M}_{n,t}) \leq s \cdot r. \end{aligned}$$

Robust Predictive Analytics :

$$z^r_{n,\mathrm{trn}} \in \arg\min_{z\in Z} \; c^r(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) := \max_{\mathbb{M}} \; c(z, \mathbb{M}, \bar{x}) \quad \text{s.t.}\; B(\mathbb{M}, \mathbb{M}_{n,\mathrm{trn}}) \le r.$$



Entropy ball :

$$B(\mathbb{M}, \mathbb{M}_{n,t}) := \sum_{(x_i, y_i)} \mathbb{M}(x_i, y_i) \cdot \log\left(\frac{\mathbb{M}(x_i,y_i)}{\mathbb{M}_{n,t}(x_i,y_i)}\right).$$

Theorem 2 :

- Robust predictive analytics with entropy ball is **bootstrap robust**:

$$\mathbb{M}^n_{n,\mathrm{trn}}\Big(c(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{train}}, \bar{x})\Big) \le \exp(-n\cdot r), \quad \forall n$$

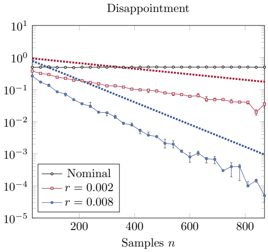
Distributionally Robust Optimization – Disappointment

Robust Predictive Analytics :

$$z^r_{n,\mathrm{trn}} \in \arg\min_{z\in Z}\ c^r(z, \mathbb{M}_{n,\mathrm{trn}}, \bar{x}) := \max_{\mathbb{M}}\ c(z, \mathbb{M}, \bar{x}) \quad \text{s.t. } B(\mathbb{M}, \mathbb{M}_{n,\mathrm{trn}}) \le r.$$

Entropy ball :

$$B(\mathbb{M}, \mathbb{M}_{n,t}) := \sum_{(x_i,y_i)} \mathbb{M}(x_i, y_i) \cdot \log\left(\frac{\mathbb{M}(x_i,y_i)}{\mathbb{M}_{n,t}(x_i,y_i)}\right).$$



Theorem 2 :

- Robust predictive analytics with entropy ball is **bootstrap robust**:

$$\mathbb{M}^n_{n,\mathrm{trn}}\Big(c(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \overbrace{\mathbb{M}_{n,\mathrm{trn}}}^{\text{train}}, \bar{x})\Big) \le \exp(-n\cdot r), \quad \forall n$$

$$\frac{1}{n}\log \mathbb{M}^n_{n,\mathrm{trn}}\left(c(z^r_{n,\mathrm{trn}}, \mathbb{M}_{n,\mathrm{btsp}}, \bar{x}) > c^r(z^r_{n,\mathrm{trn}}, \mathbb{M}_{n,\mathrm{trn}}, \bar{x})\right) \to -r, \quad n \to \infty$$

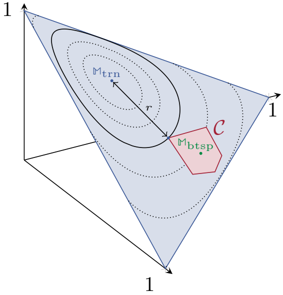
Large deviation theory:

Csiszàr (1984): Let $\mathcal{C}$ be a convex set of distributions. Then,

$$\mathbb{M}^n_{\mathrm{trn}}(\mathbb{M}_{\mathrm{btsp}} \in \mathcal{C}) \leq \exp(-n \cdot \inf_{\mathbb{M}\in\mathcal{C}} B(\mathbb{M}, \mathbb{M}_{\mathrm{trn}})), \quad \forall n.$$

At the same time

$$\frac{1}{n} \log \mathbb{M}^n_{\mathrm{trn}}(\mathbb{M}_{\mathrm{btsp}} \in \mathcal{C}) \to -\inf_{\mathbb{M}\in\mathcal{C}} B(\mathbb{M}, \mathbb{M}_{\mathrm{trn}}), \quad n \to \infty.$$



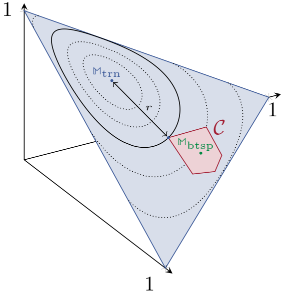
Large deviation theory:

Csiszàr (1984): Let $\mathcal{C}$ be a convex set of distributions. Then,

$$\mathbb{M}^n_{\text{trn}}(\mathbb{M}_{\text{btsp}} \in \mathcal{C}) \leq \exp(-n \cdot \inf_{\mathbb{M}\in\mathcal{C}} B(\mathbb{M}, \mathbb{M}_{\text{trn}})), \quad \forall n.$$

At the same time

$$\frac{1}{n}\log \mathbb{M}^n_{\text{trn}}(\mathbb{M}_{\text{btsp}} \in \mathcal{C}) \to -\inf_{\mathbb{M}\in\mathcal{C}} B(\mathbb{M}, \mathbb{M}_{\text{trn}}), \quad n \to \infty.$$



Proof sketch : It only remains to be shown that

$$\mathcal{C} := \left\{\mathbb{M}_{\text{btsp}} : c(z^r_{\text{trn}}, \overbrace{\mathbb{M}_{\text{btsp}}}^{\text{bootstrap}}, \bar{x}) > c^r(z^r_{\text{trn}}, \overbrace{\mathbb{M}_{\text{trn}}}^{\text{train}}, \bar{x})\right\}$$

is convex for the NW and NN formulations. Indeed, $\inf_{\mathbb{M}\in\mathcal{C}} B(\mathbb{M}, \mathbb{M}_{\text{trn}}) = r$, by construction of the robust budget c^r from its nominal counterpart c.

News Vendor Problem

$z^\star \in \arg\min_{z\geq 0} \mathbf{E}_{\mathbb{M}^\star}\left[c_{\mathrm{buy}} \cdot z - c_{\mathrm{sell}} \cdot \min(z, Y) | X = \bar{x}\right]$

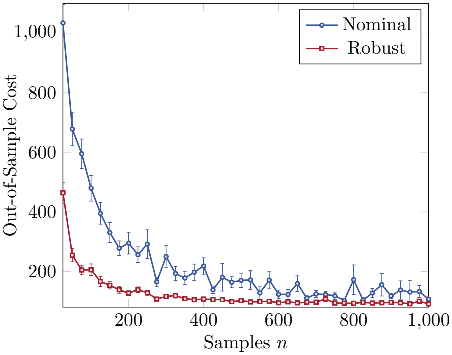
- Y : Uncertain demand
- X : Weekday, Extra Edition, Weather, ...

News Vendor Problem

$$z^\star \in \arg\min_{z\geq 0} \mathbf{E}_{\mathbb{M}^\star}\left[c_{\text{buy}} \cdot z - c_{\text{sell}} \cdot \min(z, Y) | X = \bar{x}\right]$$

- Y : Uncertain demand
- X : Weekday, Extra Edition, Weather, ...

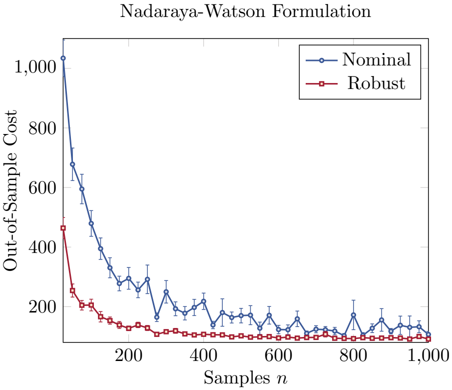
Nadaraya-Watson Formulation

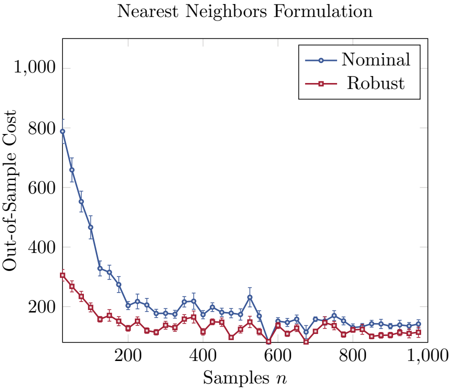


News Vendor Problem

$$z^\star \in \arg\min_{z\geq 0} \mathbf{E}_{\mathbb{M}^\star}\left[c_{\text{buy}} \cdot z - c_{\text{sell}} \cdot \min(z, Y) | X = \bar{x}\right]$$

- Y : Uncertain demand
- X : Weekday, Extra Edition, Weather, ...





Conclusions

1. The statistical bootstrap provides a data-driven robustness notion
2. Distributional robust analytics is as tractable as its nominal counterpart
3. Distributional robust optimization safeguards against over-calibrated decisions

Conclusions

1. The statistical bootstrap provides a data-driven robustness notion
2. Distributional robust analytics is as tractable as its nominal counterpart
3. Distributional robust optimization safeguards against over-calibrated decisions

Bertimas and Van Parys. "Bootstrap Robust Prescriptive Analytics", 2017.
https://arxiv.org/abs/1711.09974

github

https://github.com/vanparys/bootstrap-robust-analytics-julia