Data-Driven Auction Design III

Learnability of Product Distributions and Strong Revenue Monotonicity

Zhiyi Huang

University of Hong Kong

Recap: Single-Item Auctions

- Sell 1 item to n bidders, to maximize revenue
- Bidder i's value v_i is drawn independently from D_i

Recap: Single-Item Auctions

- Sell 1 item to n bidders, to maximize revenue
- Bidder i's value v_i is drawn independently from D_i
- Direct revelation auction
  1. Bidders bid $b_1, b_2, \ldots, b_n$
  2. Seller picks allocations $x_1, x_2, \ldots, x_n$ and payments $p_1, p_2, \ldots, p_n$
  3. Bidder i wins the item w.p. x_i, pays p_i, gets utility $v_i x_i - p_i$

Recap: Single-Item Auctions

- Sell 1 item to n bidders, to maximize revenue
- Bidder i's value v_i is drawn independently from D_i
- Direct revelation auction
  1. Bidders bid $b_1, b_2, \ldots, b_n$
  2. Seller picks allocations $x_1, x_2, \ldots, x_n$ and payments $p_1, p_2, \ldots, p_n$
  3. Bidder i wins the item w.p. x_i, pays p_i, gets utility $v_i x_i - p_i$
- Dominant-Strategy Incentive Compatible (DSIC)

$$\forall i, v_i, b_i, b_{-i} : \quad v_i x_i(v_i, b_{-i}) - p_i(v_i, b_{-i}) \geq v_i x_i(b_i, b_{-i}) - p_i(b_i, b_{-i})$$

Recap: Single-Item Auctions

- Sell 1 item to n bidders, to maximize revenue
- Bidder i's value v_i is drawn independently from D_i
- Direct revelation auction
  1. Bidders bid $b_1, b_2, \dots, b_n$
  2. Seller picks allocations $x_1, x_2, \dots, x_n$ and payments $p_1, p_2, \dots, p_n$
  3. Bidder i wins the item w.p. x_i, pays p_i, gets utility $v_i x_i - p_i$
- Dominant-Strategy Incentive Compatible (DSIC)

$$\forall i, v_i, b_i, b_{-i} : \quad v_i x_i(v_i, b_{-i}) - p_i(v_i, b_{-i}) \geq v_i x_i(b_i, b_{-i}) - p_i(b_i, b_{-i})$$

- Individually Rational (IR)

$$\forall i, v_i, b_{-i} : \quad v_i x_i(v_i, b_{-i}) - p_i(v_i, b_{-i}) \geq 0$$

Recap: Myerson's Theory

- DSIC and IR are equivalent to
  1. $x_i(v_i, b_{-i})$ is monotone (e.g., step function)
  2. $p_i(v_i, b_{-i})$ is the area on the left of $x_i(v_i, b_{-i})$ as a function of v_i (e.g., threshold price above which $x_i = 1$, if x_i is a step function)

Recap: Myerson's Theory

- DSIC and IR are equivalent to
  1. $x_i(v_i, b_{-i})$ is monotone (e.g., step function)
  2. $p_i(v_i, b_{-i})$ is the area on the left of $x_i(v_i, b_{-i})$ as a function of v_i (e.g., threshold price above which $x_i = 1$, if x_i is a step function)
- Expected revenue is equivalent to expected virtual welfare

$$\mathbf{E} \sum_{i=1}^{n} \varphi_i(v_i) x_i$$

where the virtual value φ_i is

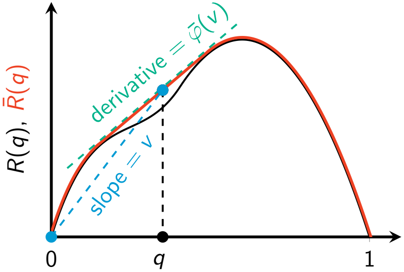
$$\varphi_i(v_i) = v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}$$

Recap: Myerson's Optimal (Single-Item) Auction

- $\bar{R}(q)$ is concave closure of revenue curve
  - Max expected revenue given sale prob. q
- Ironed virtual value $\bar{\varphi}_i(v_i)$ is $\bar{R}(q)$'s derivative
  - Quantile q's marginal revenue contribution

- Highest non-negative ironed virtual value wins
- Winner pays threshold winning bid
  i.e., lowest bid above which he/she wins

- Expected revenue is at most $\mathbf{E} \sum_{i=1}^{n} \bar{\varphi}_i(v_i) x_i$
  with equality if values in an ironed interval are treated as the same



Recap: Data-Driven Optimal (Single-Item) Auction

- Sample Complexity/Statistical Learning Model
  - Take m i.i.d. samples from $D = D_1 \times D_2 \times \cdots \times D_m$ as input
  - Output a DSIC and IR auction A
- How many samples are needed to pick a near optimal A "up to an ε margin"?
  - ε additive approximation

    $[0, 1]$-bounded distributions **(illustrative example)**
  - $1 - \varepsilon$ (multiplicative) approximation

    Regular distributions (i.e., concave revenue curve)

    MHR distributions (i.e., "strongly concave" revenue curve)

    $[1, H]$-bounded distributions
- The sample complexity is smallest number of samples needed

Recap: Summary of Upper and Lower Bounds So Far

Distributions	Upper Bound	Lower Bound
$[0, 1]$-Bounded	$\frac{n}{\varepsilon^3}$	$\frac{n}{\varepsilon^2}$
Regular distributions	$\frac{n}{\varepsilon^4}$	$\frac{n}{\varepsilon^3}$
MHR distributions	$\frac{n}{\varepsilon^3}$	$\frac{n}{\varepsilon^2}$
$[1, H]$-bounded distributions	$\frac{Hn}{\varepsilon^3}$	$\frac{Hn}{\varepsilon^2}$

- **Upper Bound:**
  Concentration inequality + covering of auction space + union bound
- **Lower Bound:**
  Assouad's method

Recap: Concentration Inequalities

Theorem (Chernoff-Hoeffding, User-Friendly Version)

$X_1, X_2, \ldots, X_m$ are i.i.d. RV over $[0,1]$. Let $\mu = \mathbf{E}\, X_i$. With probability $1-\delta$ we have

$$\left|\frac{1}{m}\sum_{i=1}^{m} X_i - \mu\right| \lesssim \sqrt{\frac{\log\frac{1}{\delta}}{m}}$$

Theorem (Bernstein Inequality, User-Friendly Version)

$X_1, X_2, \ldots, X_m$ are i.i.d. RV over $[0,1]$. Let $\mu = \mathbf{E}\, X_i$. With probability $1-\delta$ we have

$$\left|\frac{1}{m}\sum_{i=1}^{m} X_i - \mu\right| \lesssim \max\left\{\sqrt{\frac{\mu(1-\mu)\log\frac{1}{\delta}}{m}}, \frac{\log\frac{1}{\delta}}{m}\right\}$$

Learning Prices' Revenue vs. Learning Value Distribution

- Recall our approach for data-driven optimal pricing
  - It suffices to learn the revenue of every price up to ε
  - For each price, estimating its revenue reduces to estimating its quantile

Learning Prices' Revenue vs. Learning Value Distribution

- Recall our approach for data-driven optimal pricing
  - It suffices to learn the revenue of every price up to ε
  - For each price, estimating its revenue reduces to estimating its quantile

- Next consider an alternative approach

Learning Prices' Revenue vs. Learning Value Distribution

- Recall our approach for data-driven optimal pricing
  - It suffices to learn the revenue of every price up to ε
  - For each price, estimating its revenue reduces to estimating its quantile
- Next consider an alternative approach
  - It suffices to learn the distribution up to ε w.r.t. its CDF/quantile, i.e.

$$\sup_{v \in [0,1]} \left| \underbrace{F_D(v)}_{\text{true CDF}} - \underbrace{F_E(v)}_{\text{estimated CDF}} \right| \leq \varepsilon \qquad \text{(Kolmogorov distance)}$$

Learning Prices' Revenue vs. Learning Value Distribution

- Recall our approach for data-driven optimal pricing
  - It suffices to learn the revenue of every price up to ε
  - For each price, estimating its revenue reduces to estimating its quantile
- Next consider an alternative approach
  - It suffices to learn the distribution up to ε w.r.t. its CDF/quantile, i.e.

$$\sup_{v \in [0,1]} \left| \underbrace{F_D(v)}_{\text{true CDF}} - \underbrace{F_E(v)}_{\text{estimated CDF}} \right| \leq \varepsilon \qquad \text{(Kolmogorov distance)}$$

  - Then, for each price, we know its revenue up to ε

Learning Prices' Revenue vs. Learning Value Distribution

- Recall our approach for data-driven optimal pricing
  - It suffices to learn the revenue of every price up to ε
  - For each price, estimating its revenue reduces to estimating its quantile
- Next consider an alternative approach
  - It suffices to learn the distribution up to ε w.r.t. its CDF/quantile, i.e.

$$\sup_{v\in[0,1]} \Big|\underbrace{F_D(v)}_{\text{true CDF}} - \underbrace{F_E(v)}_{\text{estimated CDF}}\Big| \leq \varepsilon \qquad \text{(Kolmogorov distance)}$$

  - Then, for each price, we know its revenue up to ε
- Two approaches coincide for pricing...

Learning Prices' Revenue vs. Learning Value Distribution

- Recall our approach for data-driven optimal pricing
  - It suffices to learn the revenue of every price up to ε
  - For each price, estimating its revenue reduces to estimating its quantile
- Next consider an alternative approach
  - It suffices to learn the distribution up to ε w.r.t. its CDF/quantile, i.e.

$$\sup_{v\in[0,1]} \left| \underbrace{F_D(v)}_{\text{true CDF}} - \underbrace{F_E(v)}_{\text{estimated CDF}} \right| \le \varepsilon \qquad \text{(Kolmogorov distance)}$$

  - Then, for each price, we know its revenue up to ε
- Two approaches coincide for pricing... but not for auctions

Lower Confidence (Revenue) Bounds vs. Underestimating Distribution

- Recall the take-home question regarding optimal pricing
  - Different value distributions require different regularization in Lecture I
  - Can we get all upper bounds using the same algorithm?

Lower Confidence (Revenue) Bounds vs. Underestimating Distribution

- Recall the take-home question regarding optimal pricing
  - Different value distributions require different regularization in Lecture I
  - Can we get all upper bounds using the same algorithm?

- Lower Confidence Bound (LCB), e.g., choose p to maximize

$$p \cdot \left(\frac{\left|\text{number of samples} \geq p\right|}{m} - \sqrt{\frac{\log \frac{1}{\delta}}{m}} \right)$$

Lower Confidence (Revenue) Bounds vs. Underestimating Distribution

- Recall the take-home question regarding optimal pricing
  - Different value distributions require different regularization in Lecture I
  - Can we get all upper bounds using the same algorithm?

- Lower Confidence Bound (LCB), e.g., choose p to maximize

$$p \cdot \left(\frac{\left|\text{number of samples} \geq p\right|}{m} - \sqrt{\frac{\log \frac{1}{\delta}}{m}} \right)$$

- Alternatively, consider an underestimation of the value distribution, e.g.

$$F_E(v) = \frac{\left|\text{number of samples} \leq v\right|}{m} + \sqrt{\frac{\log \frac{1}{\delta}}{m}}$$

Lower Confidence (Revenue) Bounds vs. Underestimating Distribution

- Recall the take-home question regarding optimal pricing
  - Different value distributions require different regularization in Lecture I
  - Can we get all upper bounds using the same algorithm?
- Lower Confidence Bound (LCB), e.g., choose p to maximize

$$p \cdot \left(\frac{\left|\text{number of samples} \geq p\right|}{m} - \sqrt{\frac{\log \frac{1}{\delta}}{m}} \right)$$

- Alternatively, consider an underestimation of the value distribution, e.g.

$$F_E(v) = \frac{\left|\text{number of samples} \leq v\right|}{m} + \sqrt{\frac{\log \frac{1}{\delta}}{m}}$$

- Two approaches coincide for pricing...

Lower Confidence (Revenue) Bounds vs. Underestimating Distribution

- Recall the take-home question regarding optimal pricing
  - Different value distributions require different regularization in Lecture I
  - Can we get all upper bounds using the same algorithm?
- Lower Confidence Bound (LCB), e.g., choose p to maximize

$$p \cdot \left(\frac{\left|\text{number of samples} \geq p\right|}{m} - \sqrt{\frac{\log \frac{1}{\delta}}{m}} \right)$$

- Alternatively, consider an underestimation of the value distribution, e.g.

$$F_E(v) \ = \ \frac{\left|\text{number of samples} \leq v\right|}{m} + \sqrt{\frac{\log \frac{1}{\delta}}{m}}$$

- Two approaches coincide for pricing... but not for auctions

Data-Driven (Single-Item) Auction via Learning Value Distribution

- Product empirical distribution $E = E_1 \times E_2 \times \cdots \times E_n$
  - E_i is uniform distribution over bidder i's value samples

Data-Driven (Single-Item) Auction via Learning Value Distribution

- Product empirical distribution $E = E_1 \times E_2 \times \cdots \times E_n$
  - E_i is uniform distribution over bidder i's value samples
- Dominated product empirical distribution $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$F_{\bar{E}_i}(v) = F_{E_i}(v) + \sqrt{\frac{F_{E_i}(v)\big(1 - F_{E_i}(v)\big)}{m}}$$

(simplified incorrect form for illustration)

Data-Driven (Single-Item) Auction via Learning Value Distribution

- Product empirical distribution $E = E_1 \times E_2 \times \cdots \times E_n$
  - E_i is uniform distribution over bidder i's value samples
- Dominated product empirical distribution $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$F_{\bar{E}_i}(v) \;=\; F_{E_i}(v) \;+\; \sqrt{\frac{F_{E_i}(v)\big(1 - F_{E_i}(v)\big)}{m}}$$

(simplified incorrect form for illustration)

- Return Myerson's optimal auction w.r.t. E or $\bar{E}$

Hellinger Distance

$$\mathrm{H}(P, Q) = \frac{1}{\sqrt{2}} \left\| \sqrt{P} - \sqrt{Q} \right\|_2 = \sqrt{\frac{1}{2} \sum_{v} \left(\sqrt{P(v)} - \sqrt{Q(v)} \right)^2}$$

Hellinger Distance

$$\mathrm{H}(P, Q) = \frac{1}{\sqrt{2}} \left\| \sqrt{P} - \sqrt{Q} \right\|_2 = \sqrt{\frac{1}{2} \sum_{v} \left(\sqrt{P(v)} - \sqrt{Q(v)} \right)^2}$$

- Direct product

$$1 - \mathrm{H}(P_1 \times \cdots \times P_n, Q_1 \times \cdots \times Q_n)^2 \; = \; \prod_{i=1}^{n} \left(1 - \mathrm{H}(P_i, Q_i)^2 \right)$$

Hellinger Distance

$$\mathrm{H}(P, Q) = \frac{1}{\sqrt{2}} \left\| \sqrt{P} - \sqrt{Q} \right\|_2 = \sqrt{\frac{1}{2} \sum_v \left(\sqrt{P(v)} - \sqrt{Q(v)} \right)^2}$$

- Direct product

$$1 - \mathrm{H}(P_1 \times \cdots \times P_n, Q_1 \times \cdots \times Q_n)^2 \;=\; \prod_{i=1}^{n} \left(1 - \mathrm{H}(P_i, Q_i)^2\right)$$

- This implies sub-additivity

$$\mathrm{H}(P_1 \times \cdots \times P_n, Q_1 \times \cdots \times Q_n)^2 \;\leq\; \sum_{i=1}^{n} \mathrm{H}(P_i, Q_i)^2$$

Hellinger, Kullback–Leibler, and Total Variation

- Relation to TV

$$\mathrm{H}(P, Q)^2 \leq \mathrm{TV}(P, Q) \leq \sqrt{2} \cdot \mathrm{H}(P, Q)$$

Hellinger, Kullback–Leibler, and Total Variation

- Relation to TV

$$\mathrm{H}(P, Q)^2 \leq \mathrm{TV}(P, Q) \leq \sqrt{2} \cdot \mathrm{H}(P, Q)$$

- Relation to KL

$$\mathrm{H}(P, Q)^2 \leq \mathrm{KL}(P \| Q)$$

Hellinger, Kullback–Leibler, and Total Variation

- Relation to TV
$$\mathrm{H}(P, Q)^2 \le \mathrm{TV}(P, Q) \le \sqrt{2} \cdot \mathrm{H}(P, Q)$$
- Relation to KL
$$\mathrm{H}(P, Q)^2 \le \mathrm{KL}(P \| Q)$$
- Why TV is called total variation distance?

Hellinger, Kullback–Leibler, and Total Variation

- Relation to TV

$$\mathrm{H}(P, Q)^2 \leq \mathrm{TV}(P, Q) \leq \sqrt{2} \cdot \mathrm{H}(P, Q)$$

- Relation to KL

$$\mathrm{H}(P, Q)^2 \leq \mathrm{KL}(P \| Q)$$

- Why TV is called total variation distance?
  - P and Q are distributions over $\mathcal{T}$
  - $h : \mathcal{T} \to [0, 1]$ is a function

Hellinger, Kullback–Leibler, and Total Variation

- Relation to TV

$$\mathrm{H}(P, Q)^2 \leq \mathrm{TV}(P, Q) \leq \sqrt{2} \cdot \mathrm{H}(P, Q)$$

- Relation to KL

$$\mathrm{H}(P, Q)^2 \leq \mathrm{KL}(P \| Q)$$

- Why TV is called total variation distance?
  - P and Q are distributions over $\mathcal{T}$
  - $h : \mathcal{T} \to [0, 1]$ is a function
  - We have

$$\left| \mathbf{E}_{v \sim P} h(v) - \mathbf{E}_{v \sim Q} h(v) \right| \leq \mathrm{TV}(P, Q)$$

Learnability of Distribution

Theorem

If D has support size k, E is empirical distribution over $m \precsim \frac{k+\log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

Learnability of Distribution

Theorem

If D has support size k, E is empirical distribution over $m \eqsim \frac{k+\log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Here we prove a weaker result $\mathbf{E}\,\mathrm{H}(D, E) \lesssim \sqrt{\frac{k}{m}}$

$$\left(\mathbf{E}\,\mathrm{H}(D, E)\right)^2 \leq \mathbf{E}\,\mathrm{H}(D, E)^2$$

Learnability of Distribution

Theorem

If D has support size k, E is empirical distribution over $m \approx \frac{k+\log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Here we prove a weaker result $\mathbf{E}\,\mathrm{H}(D, E) \lesssim \sqrt{\frac{k}{m}}$

$$\left(\mathbf{E}\,\mathrm{H}(D, E)\right)^2 \leq \mathbf{E}\,\mathrm{H}(D, E)^2 = \sum_v \mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$$

Learnability of Distribution

Theorem

If D has support size k, E is empirical distribution over $m \approx \frac{k+\log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D,E) \le \varepsilon$$

- Here we prove a weaker result $\mathbf{E}\,\mathrm{H}(D,E) \lesssim \sqrt{\frac{k}{m}}$

$$\left(\mathbf{E}\,\mathrm{H}(D,E)\right)^2 \le \mathbf{E}\,\mathrm{H}(D,E)^2 = \sum_v \mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$$

- It suffices to bound $\mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$ for any v

$$\mathbf{E}\,\frac{\left(D(v) - E(v)\right)^2}{\left(\sqrt{D(v)} + \sqrt{E(v)}\right)^2}$$

Learnability of Distribution

Theorem

If D has support size k, E is empirical distribution over $m \approx \frac{k+\log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D,E) \le \varepsilon$$

- Here we prove a weaker result $\mathbf{E}\,\mathrm{H}(D,E) \lesssim \sqrt{\frac{k}{m}}$

$$\Big(\mathbf{E}\,\mathrm{H}(D,E)\Big)^2 \le \mathbf{E}\,\mathrm{H}(D,E)^2 = \sum_v \mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$$

- It suffices to bound $\mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$ for any v

$$\mathbf{E}\,\frac{\big(D(v)-E(v)\big)^2}{\left(\sqrt{D(v)}+\sqrt{E(v)}\right)^2} \le \underbrace{\mathbf{E}\,\frac{\big(D(v)-E(v)\big)^2}{D(v)}}_{\chi^2 \text{ distance}}$$

Learnability of Distribution

Theorem

If D has support size k, E is empirical distribution over $m \approx \frac{k+\log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D,E) \le \varepsilon$$

- Here we prove a weaker result $\mathbf{E}\,\mathrm{H}(D,E) \lesssim \sqrt{\frac{k}{m}}$

$$\left(\mathbf{E}\,\mathrm{H}(D,E)\right)^2 \le \mathbf{E}\,\mathrm{H}(D,E)^2 = \sum_v \mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$$

- It suffices to bound $\mathbf{E}\left(\sqrt{D(v)} - \sqrt{E(v)}\right)^2$ for any v

$$\mathbf{E}\,\frac{\left(D(v) - E(v)\right)^2}{\left(\sqrt{D(v)} + \sqrt{E(v)}\right)^2} \le \underbrace{\mathbf{E}\,\frac{\left(D(v) - E(v)\right)^2}{D(v)}}_{\chi^2 \text{ distance}} = \frac{1}{m}\left(1 - D(v)\right) \le \frac{1}{m}$$

Learnability of Product Distribution

Theorem

If $D = D_1 \times D_2 \times \cdots \times D_n$ and each D_i has support size k, $E = E_1 \times E_2 \times \cdots \times E_n$ is the product empirical distribution over $m \eqsim \frac{kn + \log \frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

Learnability of Product Distribution

Theorem

If $D = D_1 \times D_2 \times \cdots \times D_n$ and each D_i has support size k, $E = E_1 \times E_2 \times \cdots \times E_n$ is the product empirical distribution over $m \approx \frac{kn + \log \frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Again, we prove a weaker result $\mathbf{E}\,\mathrm{H}(D, E) \lesssim \sqrt{\frac{kn}{m}}$

Learnability of Product Distribution

Theorem

If $D = D_1 \times D_2 \times \cdots \times D_n$ and each D_i has support size k, $E = E_1 \times E_2 \times \cdots \times E_n$ is the product empirical distribution over $m \approx \frac{kn + \log \frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Again, we prove a weaker result $\mathbf{E}\, \mathrm{H}(D, E) \lesssim \sqrt{\frac{kn}{m}}$

$$\left(\mathbf{E}\, \mathrm{H}(D, E)\right)^2 \leq \mathbf{E}\, \mathrm{H}(D, E)^2$$

Learnability of Product Distribution

Theorem

If $D = D_1 \times D_2 \times \cdots \times D_n$ and each D_i has support size k, $E = E_1 \times E_2 \times \cdots \times E_n$ is the product empirical distribution over $m \approx \frac{kn + \log \frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Again, we prove a weaker result $\mathbf{E}\,\mathrm{H}(D, E) \lesssim \sqrt{\frac{kn}{m}}$

$$\Big(\mathbf{E}\,\mathrm{H}(D, E)\Big)^2 \;\leq\; \mathbf{E}\,\mathrm{H}(D, E)^2 \leq \sum_{i=1}^{n} \mathbf{E}\,\mathrm{H}(D_i, E_i)^2$$

Learnability of Product Distribution

Theorem

If $D = D_1 \times D_2 \times \cdots \times D_n$ and each D_i has support size k, $E = E_1 \times E_2 \times \cdots \times E_n$ is the product empirical distribution over $m \approx \frac{kn + \log\frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Again, we prove a weaker result $\mathbf{E}\,\mathrm{H}(D, E) \lesssim \sqrt{\frac{kn}{m}}$

$$\Big(\mathbf{E}\,\mathrm{H}(D, E)\Big)^2 \leq \mathbf{E}\,\mathrm{H}(D, E)^2 \leq \sum_{i=1}^{n} \mathbf{E}\,\mathrm{H}(D_i, E_i)^2 \leq \sum_{i=1}^{n} \frac{k}{m}$$

Learnability of Product Distribution

Theorem

If $D = D_1 \times D_2 \times \cdots \times D_n$ and each D_i has support size k, $E = E_1 \times E_2 \times \cdots \times E_n$ is the product empirical distribution over $m \approx \frac{kn + \log \frac{1}{\delta}}{\varepsilon^2}$ i.i.d. samples, then

$$\mathrm{H}(D, E) \leq \varepsilon$$

- Again, we prove a weaker result $\mathbf{E}\,\mathrm{H}(D, E) \lesssim \sqrt{\frac{kn}{m}}$

$$\Big(\mathbf{E}\,\mathrm{H}(D, E)\Big)^2 \leq \mathbf{E}\,\mathrm{H}(D, E)^2 \leq \sum_{i=1}^{n} \mathbf{E}\,\mathrm{H}(D_i, E_i)^2 \leq \sum_{i=1}^{n} \frac{k}{m} = \frac{kn}{m}$$

Implications to Data-Driven Auction Design

- $D = D_1 \times D_2 \times \cdots \times D_n$ is an n-dimensional product value distribution

Implications to Data-Driven Auction Design

- $D = D_1 \times D_2 \times \cdots \times D_n$ is an n-dimensional product value distribution
  - We may think of each dimension's support size as $k = \frac{1}{\varepsilon}$ because we can round values v_i to $\lfloor v_i \rfloor_\varepsilon$ (closest multiple of ε)

Implications to Data-Driven Auction Design

- $D = D_1 \times D_2 \times \cdots \times D_n$ is an n-dimensional product value distribution
  - We may think of each dimension's support size as $k = \frac{1}{\varepsilon}$ because we can round values v_i to $\lfloor v_i \rfloor_\varepsilon$ (closest multiple of ε)
  - Formally, let $\lfloor D \rfloor_\varepsilon$ be the distribution of rounded value profile

$$OPT(\lfloor D \rfloor_\varepsilon) \ \geq \ OPT(D) - \varepsilon$$

Implications to Data-Driven Auction Design

- $D = D_1 \times D_2 \times \cdots \times D_n$ is an n-dimensional product value distribution
  - We may think of each dimension's support size as $k = \frac{1}{\varepsilon}$ because we can round values v_i to $\lfloor v_i \rfloor_\varepsilon$ (closest multiple of ε)
  - Formally, let $\lfloor D \rfloor_\varepsilon$ be the distribution of rounded value profile

$$OPT(\lfloor D \rfloor_\varepsilon) \;\geq\; OPT(D) - \varepsilon$$

- $E = E_1 \times E_2 \times \cdots \times E_n$ is product empirical distribution from m rounded samples
  - E_i is the uniform distribution over bidder i's rounded sample values

Implications to Data-Driven Auction Design

- $D = D_1 \times D_2 \times \cdots \times D_n$ is an n-dimensional product value distribution
  - We may think of each dimension's support size as $k = \frac{1}{\varepsilon}$ because we can round values v_i to $\lfloor v_i \rfloor_\varepsilon$ (closest multiple of ε)
  - Formally, let $\lfloor D \rfloor_\varepsilon$ be the distribution of rounded value profile

$$OPT(\lfloor D \rfloor_\varepsilon) \ \geq \ OPT(D) - \varepsilon$$

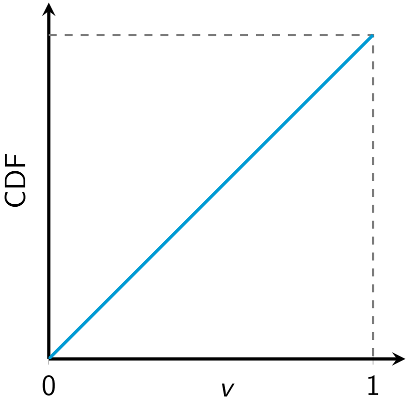
- $E = E_1 \times E_2 \times \cdots \times E_n$ is product empirical distribution from m rounded samples
  - E_i is the uniform distribution over bidder i's rounded sample values

Theorem

With $m \gtrsim \frac{n}{\varepsilon^3} + \frac{\log \frac{1}{\delta}}{\varepsilon^2}$ samples, Myerson's optimal auction M_E w.r.t. E is an ε additiive approximation w.p. $1 - \delta$.

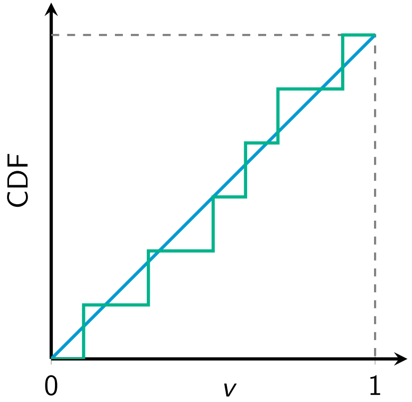
Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0, 1]$



Underestimating Value Distribution

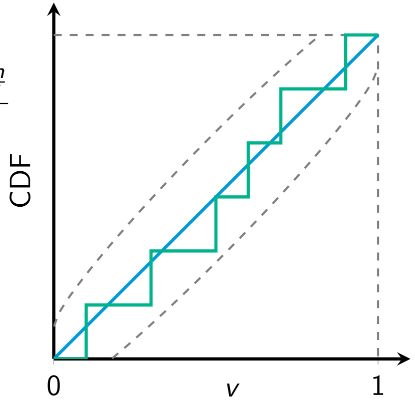
- Value distribution D, e.g., uniform on $[0, 1]$
- Empirical distribution E over m samples



Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0,1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1-F_D(v))\log\frac{m}{\delta}}{m}} + \frac{\log\frac{m}{\delta}}{m}$$



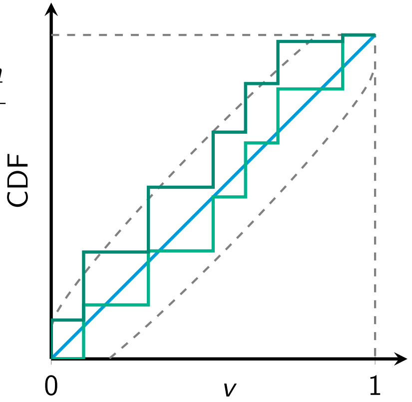
Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0, 1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

- Dominated empirical $\bar{E}$

$$F_{\bar{E}}(v) - F_E(v) \approx \sqrt{\frac{F_E(v)(1 - F_E(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$



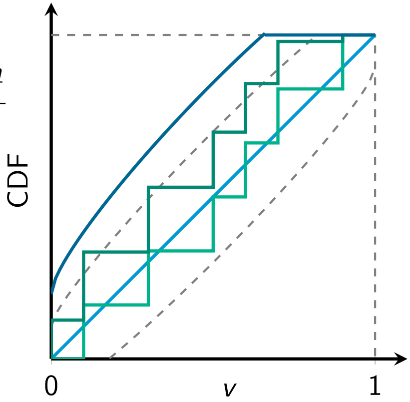
Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0, 1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

- Dominated empirical $\bar{E}$

$$F_{\bar{E}}(v) - F_E(v) \approx \sqrt{\frac{F_E(v)(1 - F_E(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$



Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0, 1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

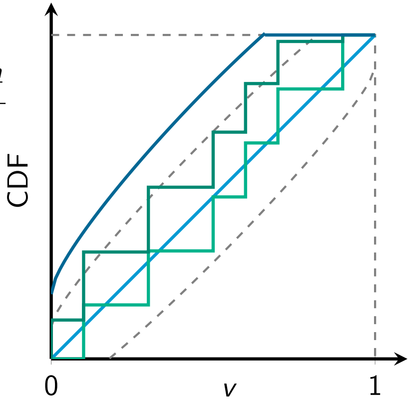
$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

- Dominated empirical $\bar{E}$

$$F_{\bar{E}}(v) - F_E(v) \approx \sqrt{\frac{F_E(v)(1 - F_E(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

- Auxiliary distribution $\bar{D}$ $\quad (\bar{D} \preceq \bar{E} \preceq D)$

$$F_{\bar{D}}(v) - F_D(v) \approx \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$



Dominated Empirical Myerson's Auction

- Compute dominated empirical distribution $\bar{E}_i$ for each bidder i
- Return Myerson's optimal auction $M_{\bar{E}}$ w.r.t. $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$M_{\bar{E}}(D) \quad \text{vs.} \quad OPT(D)$$

Dominated Empirical Myerson's Auction

- Compute dominated empirical distribution $\bar{E}_i$ for each bidder i
- Return Myerson's optimal auction $M_{\bar{E}}$ w.r.t. $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$M_{\bar{E}}(D) \quad \text{vs.} \quad OPT(D)$$

- What's the best conceivable lower bound for $M_{\bar{E}}(D)$ given $\bar{E} \preceq D$?

Dominated Empirical Myerson's Auction

- Compute dominated empirical distribution $\bar{E}_i$ for each bidder i
- Return Myerson's optimal auction $M_{\bar{E}}$ w.r.t. $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$M_{\bar{E}}(D) \quad \text{vs.} \quad OPT(D)$$

- What's the best conceivable lower bound for $M_{\bar{E}}(D)$ given $\bar{E} \preceq D$?

$$M_{\bar{E}}(D) \geq OPT(\bar{E}) \qquad \text{(strong monotonicity)}$$

Dominated Empirical Myerson's Auction

- Compute dominated empirical distribution $\bar{E}_i$ for each bidder i
- Return Myerson's optimal auction $M_{\bar{E}}$ w.r.t. $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$M_{\bar{E}}(D) \quad \text{vs.} \quad OPT(D)$$

- What's the best conceivable lower bound for $M_{\bar{E}}(D)$ given $\bar{E} \preceq D$?

$$M_{\bar{E}}(D) \geq OPT(\bar{E}) \qquad \text{(strong monotonicity)}$$

- What's the best conceivable lower bound for $OPT(\bar{E})$ given $\bar{E} \succeq \bar{D}$?

Dominated Empirical Myerson's Auction

- Compute dominated empirical distribution $\bar{E}_i$ for each bidder i
- Return Myerson's optimal auction $M_{\bar{E}}$ w.r.t. $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$M_{\bar{E}}(D) \quad \text{vs.} \quad OPT(D)$$

- What's the best conceivable lower bound for $M_{\bar{E}}(D)$ given $\bar{E} \preceq D$?

$$M_{\bar{E}}(D) \geq OPT(\bar{E}) \qquad \text{(strong monotonicity)}$$

- What's the best conceivable lower bound for $OPT(\bar{E})$ given $\bar{E} \succeq \bar{D}$?

$$OPT(\bar{E}) \geq OPT(\bar{D}) \qquad \text{(weak monotonicity)}$$

Dominated Empirical Myerson's Auction

- Compute dominated empirical distribution $\bar{E}_i$ for each bidder i
- Return Myerson's optimal auction $M_{\bar{E}}$ w.r.t. $\bar{E} = \bar{E}_1 \times \bar{E}_2 \times \cdots \times \bar{E}_n$

$$M_{\bar{E}}(D) \quad \text{vs.} \quad OPT(D)$$

- What's the best conceivable lower bound for $M_{\bar{E}}(D)$ given $\bar{E} \preceq D$?

$$M_{\bar{E}}(D) \geq OPT(\bar{E}) \qquad \text{(strong monotonicity)}$$

- What's the best conceivable lower bound for $OPT(\bar{E})$ given $\bar{E} \succeq \bar{D}$?

$$OPT(\bar{E}) \geq OPT(\bar{D}) \qquad \text{(weak monotonicity)}$$

- Compare $OPT(\bar{D})$ and $OPT(D)$

Strong (Revenue) Monotonicity

Theorem

For any value distributions $D \succeq \bar{E}$*, and the optimal auction* $M_{\bar{E}}$ *for* $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Strong (Revenue) Monotonicity

Theorem

For any value distributions $D \succeq \bar{E}$*, and the optimal auction* $M_{\bar{E}}$ *for* $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Here we only prove weak monotonicity, i.e., $OPT(D) \geq OPT(\bar{E})$, via coupling

Strong (Revenue) Monotonicity

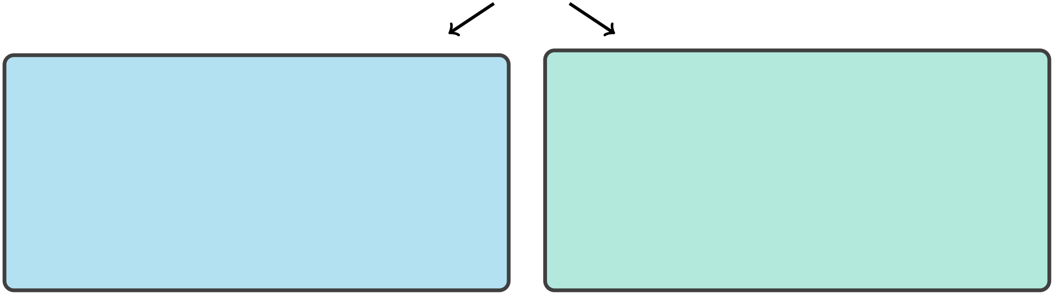
Theorem

For any value distributions $D \succeq \bar{E}$, and the optimal auction $M_{\bar{E}}$ for $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Here we only prove weak monotonicity, i.e., $OPT(D) \geq OPT(\bar{E})$, via coupling

quantiles $q_1, q_2, \ldots, q_n$



Strong (Revenue) Monotonicity

Theorem

For any value distributions $D \succeq \bar{E}$, and the optimal auction $M_{\bar{E}}$ for $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Here we only prove weak monotonicity, i.e., $OPT(D) \geq OPT(\bar{E})$, via coupling

quantiles $q_1, q_2, \ldots, q_n$

1. Values $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \sim \bar{E}$
2. Allocate to bidder i with highest non-negative $\bar{\varphi}_{\bar{\boldsymbol{E}}_i}(\bar{\boldsymbol{v}}_i)$
3. Winner pays threshold bid

Strong (Revenue) Monotonicity

Theorem

For any value distributions $D \succeq \bar{E}$*, and the optimal auction* $M_{\bar{E}}$ *for* $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Here we only prove weak monotonicity, i.e., $OPT(D) \geq OPT(\bar{E})$, via coupling

quantiles $q_1, q_2, \ldots, q_n$

1. Values $v_1, v_2, \ldots, v_n \sim D$

1. Values $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \sim \bar{E}$
2. Allocate to bidder i with highest non-negative $\bar{\varphi}_{\bar{E}_i}(\bar{v}_i)$
3. Winner pays threshold bid

Strong (Revenue) Monotonicity

Theorem

For any value distributions $D \succeq \bar{E}$, *and the optimal auction* $M_{\bar{E}}$ *for* $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Here we only prove weak monotonicity, i.e., $OPT(D) \geq OPT(\bar{E})$, via coupling

quantiles $q_1, q_2, \dots, q_n$

1. Values $v_1, v_2, \dots, v_n \sim D$
2. Allocate to bidder i with highest non-negative $\bar{\varphi}_{\bar{E}_i}(\bar{v}_i)$

1. Values $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n \sim \bar{E}$
2. Allocate to bidder i with highest non-negative $\bar{\varphi}_{\bar{E}_i}(\bar{v}_i)$
3. Winner pays threshold bid

Strong (Revenue) Monotonicity

Theorem

For any value distributions $D \succeq \bar{E}$, and the optimal auction $M_{\bar{E}}$ for $\bar{E}$

$$M_{\bar{E}}(D) \geq OPT(\bar{E})$$

Here we only prove weak monotonicity, i.e., $OPT(D) \geq OPT(\bar{E})$, via coupling

quantiles $q_1, q_2, \ldots, q_n$

1. Values $v_1, v_2, \ldots, v_n \sim D$
2. Allocate to bidder i with highest non-negative $\bar{\varphi}_{\bar{E}_i}(\bar{v}_i)$
3. Winner pays threshold bid, which is **at least as large**

1. Values $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \sim \bar{E}$
2. Allocate to bidder i with highest non-negative $\bar{\varphi}_{\bar{E}_i}(\bar{v}_i)$
3. Winner pays threshold bid

Comparing $OPT(D)$ and $OPT(\bar{D})$

Reminder

$$F_{\bar{D}}(v) - F_D(v) \approx \sqrt{\frac{F_D(v)(1-F_D(v))\log\frac{m}{\delta}}{m}} + \frac{\log\frac{m}{\delta}}{m}$$

Lemma

If we have $m \gtrsim \frac{n \cdot (\log \frac{m}{\varepsilon\delta})^2}{\varepsilon^2}$ *samples, then the auxiliary distribution* $\bar{D}$

$$\mathrm{H}(D, \bar{D}) \le \frac{\varepsilon}{\sqrt{2}}$$

Comparing $OPT(D)$ and $OPT(\bar{D})$

Reminder

$$F_{\bar{D}}(v) - F_D(v) \approx \sqrt{\frac{F_D(v)(1-F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

Lemma

If we have $m \gtrsim \frac{n \cdot (\log \frac{m}{\varepsilon\delta})^2}{\varepsilon^2}$ samples, then the auxiliary distribution $\bar{D}$

$$\mathrm{H}(D, \bar{D}) \leq \frac{\varepsilon}{\sqrt{2}}$$

$\Rightarrow$

$$\mathrm{TV}(D, \bar{D}) \leq \varepsilon$$

Comparing $OPT(D)$ and $OPT(\bar{D})$

Reminder

$$F_{\bar{D}}(v) - F_D(v) \precsim \sqrt{\frac{F_D(v)(1-F_D(v))\log\frac{m}{\delta}}{m}} + \frac{\log\frac{m}{\delta}}{m}$$

Lemma

If we have $m \gtrsim \frac{n \cdot (\log \frac{m}{\varepsilon\delta})^2}{\varepsilon^2}$ samples, then the auxiliary distribution $\bar{D}$

$$\mathrm{H}(D, \bar{D}) \leq \frac{\varepsilon}{\sqrt{2}}$$

$$\Rightarrow \quad \mathrm{TV}(D, \bar{D}) \leq \varepsilon$$

$$\Rightarrow \quad \text{for any auction } A, \quad A(\bar{D}) \geq A(D) - \varepsilon$$

Comparing $OPT(D)$ and $OPT(\bar{D})$

Reminder

$$F_{\bar{D}}(v) - F_D(v) \lesssim \sqrt{\frac{F_D(v)(1-F_D(v))\log\frac{m}{\delta}}{m}} + \frac{\log\frac{m}{\delta}}{m}$$

Lemma

If we have $m \gtrsim \frac{n\cdot(\log\frac{m}{\varepsilon\delta})^2}{\varepsilon^2}$ samples, then the auxiliary distribution $\bar{D}$

$$\mathrm{H}(D, \bar{D}) \le \frac{\varepsilon}{\sqrt{2}}$$

$$\Rightarrow \quad \mathrm{TV}(D, \bar{D}) \le \varepsilon$$

$$\Rightarrow \quad \text{for any auction } A, \quad A(\bar{D}) \ge A(D) - \varepsilon$$

$$\Rightarrow \quad OPT(\bar{D}) \ge OPT(D) - \varepsilon$$

Summary

Distributions	Sample Complexity
$[0,1]$-Bounded	$\frac{n}{\varepsilon^2}$
Regular distributions	$\frac{n}{\varepsilon^3}$
MHR distributions	$\frac{n}{\varepsilon^2}$
$[1,H]$-bounded distributions	$\frac{Hn}{\varepsilon^2}$

- **Upper Bound:**
  Learnability of product distribution + strong (revenue) monotonicity
- **Lower Bound:**
  Assouad's method

Recap

Two Different Viewpoints

Learnability of Product Distributions

Strong (Revenue) Monotonicity

Further Extensions and Open Questions

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization
  - Single-parameter auctions (e.g., multiple homogeneous items)

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization
  - Single-parameter auctions (e.g., multiple homogeneous items)
  - Multi-parameter auctions (e.g., multiple heterogeneous items)
    Optimal sample complexity is still open

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization
  - Single-parameter auctions (e.g., multiple homogeneous items)
  - Multi-parameter auctions (e.g., multiple heterogeneous items)
    Optimal sample complexity is still open
- Sequential decision-making in stochastic models

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization
  - Single-parameter auctions (e.g., multiple homogeneous items)
  - Multi-parameter auctions (e.g., multiple heterogeneous items)
    Optimal sample complexity is still open
- Sequential decision-making in stochastic models
  - Prophet inequality
    Optimal sample complexity is still open

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization
  - Single-parameter auctions (e.g., multiple homogeneous items)
  - Multi-parameter auctions (e.g., multiple heterogeneous items)
    Optimal sample complexity is still open
- Sequential decision-making in stochastic models
  - Prophet inequality
    Optimal sample complexity is still open
  - Pandora's box

Sample Compelxity of Optimization Problems in Stochastic Models

- Revenue maximization
  - Single-parameter auctions (e.g., multiple homogeneous items)
  - Multi-parameter auctions (e.g., multiple heterogeneous items)
    Optimal sample complexity is still open
- Sequential decision-making in stochastic models
  - Prophet inequality
    Optimal sample complexity is still open
  - Pandora's box
  - Online stochastic matching
    Optimal sample complexity is still open

Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0,1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0, 1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

- Dvoretzky–Kiefer–Wolfowitz (DKW) inequality

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{\log \frac{1}{\delta}}{m}}$$

Underestimating Value Distribution

- Value distribution D, e.g., uniform on $[0, 1]$
- Empirical distribution E over m samples
- Bernstein Inequality + Union Bound

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{F_D(v)(1 - F_D(v)) \log \frac{m}{\delta}}{m}} + \frac{\log \frac{m}{\delta}}{m}$$

- Dvoretzky–Kiefer–Wolfowitz (DKW) inequality

$$\left|F_E(v) - F_D(v)\right| \lesssim \sqrt{\frac{\log \frac{1}{\delta}}{m}}$$

- **Open question:** Is there a Bernstein-style DKW inequality?

Bidders' Strategic Behaviors in Data-Driven Auction Design

- Bidders may underbid today in order to get a lower price tomorrow

Bidders' Strategic Behaviors in Data-Driven Auction Design

- Bidders may underbid today in order to get a lower price tomorrow
- Can we learn optimal auctions despite of bidders' strategic behaviors?

Bidders' Strategic Behaviors in Data-Driven Auction Design

- Bidders may underbid today in order to get a lower price tomorrow
- Can we learn optimal auctions despite of bidders' strategic behaviors?
  - Impossible if bidders are patient

Bidders' Strategic Behaviors in Data-Driven Auction Design

- Bidders may underbid today in order to get a lower price tomorrow
- Can we learn optimal auctions despite of bidders' strategic behaviors?
  - Impossible if bidders are patient
  - Possible for relatively simple auctions, and impatient bidders (with slower convergence rate than learning form non-strategic bidders)

Bidders' Strategic Behaviors in Data-Driven Auction Design

- Bidders may underbid today in order to get a lower price tomorrow
- Can we learn optimal auctions despite of bidders' strategic behaviors?
  - Impossible if bidders are patient
  - Possible for relatively simple auctions, and impatient bidders
    (with slower convergence rate than learning form non-strategic bidders)
  - **Open question:** Is the slower convergence rate avoidable?

References

1. Chenghao Guo, Zhiyi Huang, and Xinzhi Zhang. *"Settling the sample complexity of single-parameter revenue maximization."* In Proceedings of the 51st Annual ACM Symposium on Theory of Computing, ACM, pp. 662–673, 2019.
2. Chenghao Guo, Zhiyi Huang, Zhihao Gavin Tang, and Xinzhi Zhang. *"Generalizing complex hypotheses on product distributions: auctions, prophet inequalities, and Pandora's problem."* In Proceedings of the 34th Annual Conference on Learning Theory, PMLR, pp. 2248–2288, 2021.
3. Ziyun Chen, Zhiyi Huang, Dorsa Madji, Zipeng Yan. *"Strong revenue (non-)monotonicity of single-parameter auctions."* In Proceedings of the 24th ACM Conference on Economics and Computation, ACM, pp. 452–471, 2023.