Prior Independent Mechanism Design

Yingkai Li

EC4501/EC4501HM Semester 2, AY2024/25

Leading Example: Auctions

Selling a single item to n buyers to maximize expected revenue.

- each buyer i has private value v_i drawn independently from F_i;
- linear utilities: $u_i = v_i x_i - p_i$.

Common prior assumption: F is commonly known by both the seller and the buyers.

Optimal mechanism: virtual value maximization [Myerson '81] or equivalently marginal revenue maximization [Bulow and Robert '89].

Robustness to Distributional Knowledge

Relaxing the knowledge of the seller:

1. seller only has sample access to the valuation distributions;
2. seller has no information about the distributions except knowing that they are i.i.d.;
3. there is no underlying distributional structure for the valuations.

Focus on dominant strategy incentive compatible (DSIC) mechanisms.

- with loss of generality;
- requires no assumption on the buyers' distributional knowledge.

If buyers have common knowledge about the valuation distribution:

- Implementation theory [Caillaud and Robert '05]: implements the Bayesian optimal mechanism when there are multiple agents.

Robust Mechanism Design Frameworks

Let $\mathbb{M}$ be the set of possible mechanisms (e.g., DSIC mechanisms). For any $M \in \mathbb{M}$

- $M(v)$ is the performance of mechanism M given realized value profile v;
- $M(F) \triangleq \mathbf{E}_{v\sim F}[M(F)]$.

Bayesian Optimal

$$\mathrm{OPT}_F = \operatorname*{argmax}_{M\in\mathbb{M}} M(F).$$

Prior-independent Optimal

$$\beta = \min_{M\in\mathbb{M}} \max_{F\in\mathbb{F}} \frac{\mathrm{OPT}_F(F)}{M(F)}$$

Prior-free Optimal (for Benchmark B)

$$\gamma_B = \min_{M\in\mathbb{M}} \max_{v\in\mathbb{V}} \frac{B(v)}{M(v)}$$

Mechanism Design with Samples

In auction environment, a sample is a valuation profile $v = (v_1, \dots, v_n)$ drawn independently from prior F.

Mechanism design with K samples:

- the seller does not have prior knowledge about the valuation distribution F;
- the seller has access to K samples $S^K = (s_1, \dots, s_K)$ from the valuation distribution F, e.g, with access to historical data from the agents.

A sampling mechanism $\widehat{M} : \mathbb{V}^K \to \Delta(\mathbb{M})$.

- $\widehat{M}(F) \triangleq \mathbf{E}_{v^K \sim F}\left[\widehat{M}(v^K)(F)\right]$.

Prior-independent approximation:

$$\beta = \min_{\widehat{M} \in \widehat{\mathbb{M}}} \max_{F \in \mathbb{F}} \frac{\mathrm{OPT}_F(F)}{\widehat{M}(F)}$$

Sample Complexity

With $K = \infty$ samples, the seller can perfectly learn the valuation distributions, and design the Bayesian optimal mechanisms.

- In reality, sample size is finite.

Question: what is the minimum K that guarantees a $(1 + \epsilon)$-approximation to the optimal.

- depends on the set $\mathbb{F}$ of possible distributions.

Necessity of Distributional Assumptions

No mechanism can guarantee any approximation to the optimal without any additional assumption on the set of possible distribution.

Example: a single buyer with value distribution as follows:

- with probability $\epsilon > 0$, the value is v;
- with probability $1 - \epsilon$, the value is 0.

Given any finite sample size K, $\exists \epsilon > 0$ s.t. the seller only sees K samples with value 0 with high probability.
$\Rightarrow$ seller cannot infer the value v from the sample with high probability.
$\Rightarrow$ given any mechanism, there exists a v such that the revenue loss is large.

The revenue contribution concentrates too heavily on the tail events.

Sample Complexity

Definition (Regularity)

A distribution F is regular if $\phi(v) = v - \frac{1-F(v)}{f(v)}$ is non-decreasing in v.

Assumes regularity (overkill for $(1+\epsilon)$-approximation, only need small tail assumption)

Theorem (Guo, Huang and Zhang '19)

For any $\epsilon, \delta \in (0,1)$, if the valuation distribution is regular for all agents, there exists a sampling mechanism $\widehat{M}$ such that with $O(n\epsilon^{-3} \cdot \ln^2(\frac{n}{\epsilon\delta}))$ samples, with probability at least $1-\delta$, the multiplicative revenue loss from $\widehat{M}$ is at most ϵ.

Sample Complexity

Setting	Lower Bound (Sec. 4)	Upper Bound (Sec. 3)
Regular	$\Omega(n\epsilon^{-3})$	$\tilde{O}(n\epsilon^{-3})$
MHR	$\tilde{\Omega}(n\epsilon^{-2})$	$\tilde{O}(n\epsilon^{-2})$
$[1, H]$	$\Omega(nH\epsilon^{-2})$	$\tilde{O}(nH\epsilon^{-2})$
$[0, 1]$-additive	$\Omega(n\epsilon^{-2})$	$\tilde{O}(n\epsilon^{-2})$

Figure: Various sample complexity bounds in [Guo, Huang and Zhang '19].

Proof of Sample Complexity

Given samples $S^K = (s_1, \dots, s_K)$, let $\hat{F}(S^K)$ be the empirical distribution.

- uniform distribution over samples.

Dominated Empirical Distributions: Let $\tilde{F}(S^K)$ be the distribution that shifts the quantiles of $\hat{F}(S^K)$ down by $\tilde{O}(\frac{1}{\sqrt{K}})$.

- overestimation given sample realizations: potentially no sale; avoid
- underestimation given sample realizations: slightly lower price of sale.

Run Myerson's optimal auction based on $\tilde{F}(S^K)$.

Proof of Sample Complexity

Lemma

With probability at least $1-\delta$, F_i first order stochastically dominates $\tilde{F}_i(S^K)$ for all i.

Application of concentration inequalities: Bernstein's inequality [Bernstein '24].

- with high probability, the estimation error of the quantiles based on the empirical distribution is at most $\tilde{O}(\frac{1}{\sqrt{K}})$.

Lemma

With probability at least $1-\delta$, $\mathrm{OPT}_{\tilde{F}_i(S^K)}(F) \geq \mathrm{OPT}_{\tilde{F}_i(S^K)}(\tilde{F}_i(S^K))$.

Lemma

If $K = O(n\epsilon^{-3} \cdot \ln^2(\frac{n}{\epsilon\delta}))$, w.p. at least $1-\delta$, $\mathrm{OPT}_{\tilde{F}_i(S^K)}(\tilde{F}_i(S^K)) \geq \frac{1}{1+\epsilon} \cdot \mathrm{OPT}_F(F)$.

Consider $\tilde{F}$ that is obtained by shifting the quantiles of F down by $\tilde{O}(\frac{1}{\sqrt{K}})$.

- $\tilde{F}$ is close to F in KL-distance;
- optimal revenue of $\tilde{F}$ and F must be close [Pinsker '60].

Strong Revenue Monotonicity

Theorem (Devanur, Huang and Psomas '16)

For any distributions F and F' such that F'_i first order stochastically dominates F_i for all i, letting OPT_F be the optimal mechanism for F, we have $\mathrm{OPT}_F(F') \geq \mathrm{OPT}_F(F)$.

Remark: the strong revenue monotonicity does not hold for arbitrary mechanisms.

A weaker version: $\mathrm{OPT}_{F'}(F') \geq \mathrm{OPT}_F(F)$.

Idea:

- Amortized analysis;
- Coupling argument.

See illustration for the weaker version on board.

A Few Samples

Seller may not have sufficient samples to learn the distributions for a $(1+\epsilon)$-approximation.

Question: what is the best approximation ratio fixing a small K, e.g., $K = 1, 2$?

- focus on even simpler setting with a single buyer.

Even characterizing the robust performance of an arbitrary mechanism is challenging with small sample sizes.

- consider mechanisms with specific forms that are "easy" to analyze.

A Simple Example: A Single Sample

Theorem (Dhangwatnotai, Roughgarden and Yan '15)

For regular valuation distribution, posting a price equal to the sample guarantees a 2-approximation.

Prove by graphical illustration.

Can we do better?

- Not for deterministic mechanisms;
- Yes for using lotteries.

Randomized Sample-based Pricing

Focus on mechanisms that (randomly) markup/markdown the sample.

Pricing mechanism ψ:

- draw markup r from distribution ψ;
- post price $r \cdot s$ where s is the sample.

Further simplification: focus on ψ with binary support $\{\underline{r}, \bar{r}\}$ with $\underline{r} < 1 < \bar{r}$.

- still no closed-form characterization for worst case distribution given the mechanism;
- a general numerical procedure (using discretization and dynamic programming) for approximately computing the worst case performance [Allouah, Bahamou and Besbes '22].

Intuition on why random mechanisms improve the worst case performance

- see graphical illustration.
- markdown the sample by a small factor significantly improves the worst case performance for large q^*;
- markup the sample improves the worst case performance for small q^*.

Empirical Revenue Maximization

Given samples $S^K = (s_1, \ldots, s_K)$, let $\hat{F}(S^K)$ be the empirical distribution.

Empirical Revenue Maximization: post price $p(S^K)$ to the agent

- $p(S^K)$ is the monopoly optimal price for empirical distribution $\hat{F}(S^K)$.

Theorem (Daskalakis and Zampetakis '20)

When $K = 2$, the empirical revenue maximization mechanism achieve at least 55.8% of the optimal revenue for regular distributions.

Not optimal: ERM cannot guarantee more than 61.2% of the optimal revenue [Allouah, Bahamou and Besbes '22].

- intuitive reason: samples may overestimates the distributions.

Alternative ideas: shift the empirical distributions.

- difficult to analyze its worst case performance for this class of mechanism.

A Few Samples

To provide a tractable framework for analyzing the performance of given mechanisms, focus on mechanisms that (randomly) markup/markdown the order statistics from the samples.

Pricing mechanism (i, ψ) [Allouah, Bahamou and Besbes '22]:

- draw markup r from distribution ψ;
- post price $r \cdot s_{[i]}$ where $s_{[i]}$ is the ith order statistics from K samples.

Easier to analyze (compared to ERM):

- price posted is monotone with respect to first order stochastic dominance of the distribution.

A general numerical procedure (using discretization and dynamic programming) for approximately computing the worst case performance [Allouah, Bahamou and Besbes '22].

- the analysis turn out to be pretty tight.

A Few Samples

Allouah, Bahamou, and Besbes: *Pricing with Samples*
Operations Research, 2022, vol. 70, no. 2, pp. 1088–1104, © 2022 INFORMS 1103

Table 4. Lower Bounds on the Maximin Ratio $\mathcal{R}(\mathcal{P}_N, \mathcal{F}_\alpha)$

			Performance of mechanism		Distribution parameters	
Class	N	Pricing mechanism (i, ψ)	Lower bound	Upper bound	q_0	q_1
mhr	1	$(1, \delta_{0.76})$	64.4%	64.8%	0.448	0.079
	2	$(2, \delta_{0.73})$	71.6%	72.3%	0.999	0.154
	5	$(3, \delta_{0.85})$	79.1%	79.9%	0.41	0.07
	10	$(6, \delta_{0.81})$	80.4%	81.0%	1	0
Regular	1	$(1, 0.9483\,\delta_{0.98} + 0.0517\,\delta_{2.1})$	50.2%	50.4%	0.016	0.002
	2	$(2, \delta_{0.75})$	61.5%	61.9%	0.003	0.001
	5	$(4, \delta_{0.80})$	62.4%	62.5%	0.002	0.001
	10	$(8, \delta_{0.70})$	65.3%	66.0%	0.002	0.001

Note. The table reports mechanisms and an interval in which its performance is guaranteed to belong for various number of the number of samples N. The table also reports near-worst-case distributions for the mechanisms proposed (for all of these, we fixed $\overline{v} = +\infty$).

Revelation Principle

Revelation principle: in Bayesian environments with common knowledge, it is without loss to focus on revelation mechanisms [Myerson '81].

In general, in robust environments, the principal cannot simulate arbitrary mechanisms using revelation mechanisms due to the lack of information.

Definition (Revelation Gap (Informal))

The revelation gap of an robust mechanism design environments is defined as the multiplicative gap between the worst case approximation ratio of non-revelation mechanisms and revelation mechanisms.

Question: is there a setting where the revelation gap is strictly larger than 1.

Revelation Gap for Pricing from Samples

Pricing from samples: a single buyer

- private value $v \sim F$;
- seller has access to a sample $s \sim F$;
- buyer knows both v and F.

Remark: Requires distributional knowledge of the buyer.

In this setting, revelation mechanism is equivalent to sample-based pricing:

- post a price $p \sim G(s)$ to the agent as take-or-leave-it offer.

Non-revelation Mechanisms

Definition (Sample-bid Mechanism)

For any $\alpha > 0$, sample-bid mechanism SB_α

1. solicits a non-negative bid $b \geq 0$;
2. charges the agent $\alpha \cdot \min\{b, s\}$;
3. allocates the item to the agent if $b \geq s$.

Sample-bid mechanism SB_α shares similar format as Becker-DeGroot-Marschak (BDM) method in experimental economics [Becker, DeGroot, Marschak '64].

Remark: SB_α is individual rational: the utility of bidding 0 is 0.

Performance of Sample-bid Mechanisms

Consider the simpler case where F is MHR.

Agent's utility for bidding $b \geq 0$ in SB_α:

$$u(b; v, F) = v \cdot \underbrace{F(b)}_{\mathbf{Pr}_{s\sim F}[s \leq b]} - \underbrace{\alpha b \cdot (1 - F(b))}_{\text{payment when } s \geq b} - \underbrace{\alpha \int_0^b t dF(t)}_{\text{payment when } s \leq b}$$

$u(b; v, F)$ is convex in b given any v and F that is MHR.

- agent maximizes his utility by bidding 0 or ∞.

Let $w = \mathbf{E}_{v\sim F}[v]$ be the expected welfare.
SB_α is equivalent to posting $\alpha \cdot w$ when F is MHR.

- perfectly infer the welfare using a single sample!

Setting $\alpha = 0.824$ achieves an approximation of 1.296.

- the probability of sale is at least $\frac{1}{e}$ when selling at price w if F is MHR [Barlow and Marshall '65].

Approximation Guarantees

	Class of revelation mechanisms [Allouah, Bahamou and Besbes '22]		Class of all mechanisms [Feng, Hartline and Li '21]	
	Regular dists.	MHR dists.	Regular dists.	MHR dists.
Upper bound	1.996	1.575	1.835	1.296
Lower bound	1.957	1.543	1.073	

Theorem (Revelation Gap [Feng, Hartline and Li '21])

In single-item single-agent auction with single sample access, for the revenue maximization problem, the revelation gap is

- *For MHR distributions* $\mathbb{F}_M$, $\Gamma(\mathbb{F}_M) \in [1.190, 1.467]$;
- *For regular distributions* $\mathbb{F}_R$, $\Gamma(\mathbb{F}_R) \in [1.066, 1.859]$.

Take away: importance of analyzing non-revelation mechanisms in robust settings.

Non-truthful Samples

In practical applications, non-truthful auctions are widely adopted [Hartline and Taggart '19]

- the seller may be restricted to only adopt mechanisms with all-pay format or winner-pays-bid format;
- seller collects strategic bids instead of valuation profile of the buyers from historical data.

Question: how to optimize the allocation rule based on the restriction on payment formats and non-truthful bidding data.

Informal Statement of Result: polynomial number of samples are sufficient to guarantee an $(1+\epsilon)$-approximation [Hartline and Taggart '19] .

Prior Independent Mechanism Design

Prior-independent Optimal

$$\beta = \min_{M \in \mathbb{M}} \max_{F \in \mathbb{F}} \frac{\mathrm{OPT}_F(F)}{M(F)}$$

Prior Independent Mechanism Design

Single-item auction, n agents, revenue maximization

- $\mathbb{F}$: the set of all i.i.d. regular distributions.

Naive idea: partition the agents randomly, and use half of the agents as samples.

- robust but not optimal for the prior-independent analysis.

Better idea: second-price auction [Bulow and Klemperer '96; Allouah and Besbes '20]
random markup mechanisms [Fu, Immorlica, Lucier and Strack '15; Hartline, Johnson and Li '20]

- avoid revenue loss by just using part of agents as samples.

Auctions vs Negotiations

Theorem (Bulow and Klemperer '96)

For any $n \geq 1$, assuming that all agents have i.i.d. regular value distributions, the expected revenue from second-price auction with $n+1$ agents is at least the optimal revenue with n agents.

Proof: [Hartline '20] Given i.i.d. regular value distributions, the optimal mechanism allocates the item to the agent with highest non-negative (virtual) value.

Let M_S be the optimal mechanism that always sells the item.

- M_S allocates the item to the agent with highest (virtual) value;
- M_S is the second-price auction.

Revenue from OPT with n agents is at most the revenue from M_S with $n+1$ agents.

$$\mathrm{SPA}(F^{n+1}) = M_S(F^{n+1}) \geq \mathrm{OPT}(F^n).$$

One mechanism that always sell with $n+1$ agents: run optimal with n agents, give the item to the additional agent for free if no sale.

Prior Independent Approximations

Corollary

When there are $n \geq 2$ agents, assuming that all agents have i.i.d. regular value distributions, the expected revenue from second-price auction (SPA) is at least $\frac{n-1}{n}$ fraction of the optimal revenue.

Optimal revenue is submodular.

- Given any $n' < n$ agents, simulate the values for $n - n'$ agents and run the optimal mechanism for n agents on n' real agents and $n - n'$ simulated agents.

$$\mathrm{SPA}(F^n) \geq \mathrm{OPT}(F^{n-1}) \geq \frac{n-1}{n} \cdot \mathrm{OPT}(F^n).$$

SPA is asymptotically optimal as $n \to \infty$.

- 2-approximation to the optimal when $n = 2$.

Prior Independent Optimal Auction

Is SPA optimal when $n = 2$?

- Depends on the family of possible distributions.

Theorem (Allouah and Besbes '20)

For $n = 2$, if $\mathbb{F}$ is the set of i.i.d. MHR distributions, second-price auction is prior-independent optimal with approximation ratio 1.398.

- SPA and truncated exponential distributions are mutual best response.

If $\mathbb{F}$ is the set of i.i.d. regular distributions, SPA is not prior-independent optimal.

Randomization helps for improving the prior-independent approximation when $n = 2$ [Fu, Immorlica, Lucier and Strack '15].

Prior Independent Optimal Auction

Definition (Random Markup Mechanism with Scale Distribution G)

- Draw $\alpha \sim G$
- Offer agent 1 price $\alpha\, v_2$, and vice versa.

Consider a specific distribution $G^* : \alpha \sim \begin{cases} 1 & \text{w.p. } 0.806 \\ 2.447 & \text{otherwise.} \end{cases}$

Theorem (Hartline, Johnson and Li '20)

Random markup mechanism with scale distribution G^* is prior-independent optimal with $\beta \approx 1.91.^*$

*lower bound holds under a technical restriction on the family of mechanisms.

Prior Independent Optimal Auction

A mechanism is scale-invariant if for any $\alpha > 0$, the outcome distributions given (v_1, v_2) is the same as $(\alpha v_1, \alpha v_2)$.

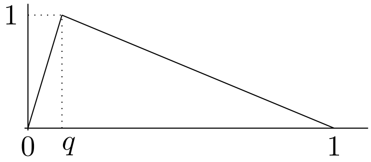
Lemma (Allouah and Besbes '20)

*It is without loss to focus on scale-invariant mechanism.**

*This holds under a technical restriction on the family of mechanisms.

Triangle distributions: Tri_q with cumulative distribution function

$$\mathrm{Tri}_q(v) = \begin{cases} 1 - \frac{1}{1+v(1-q)} & v \leq 1/q, \\ 1 & \text{otherwise.} \end{cases}$$



Triangle distributions are first order stochastically dominated by other regular distributions.

Lemma (Hartline, Johnson and Li '20)

It is without loss to focus on worst case distributions that are triangle distributions.

Prior Independent Optimal Auction

Scale invariance mechanisms are essentially random markup mechanism [Hartline, Johnson and Li '20].

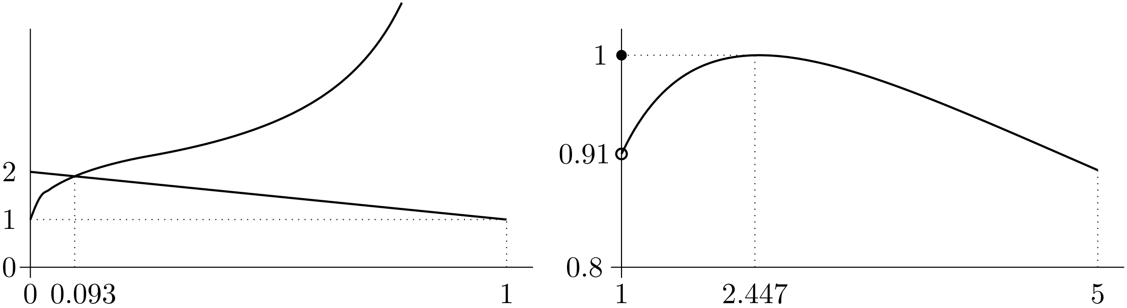


Figure: Left: approximation of second-price and non-trivial markup mechanism for distribution Tri_q.

Right: revenue of the markup mechanisms for triangle distribution Tri_{q^*} with $q^* \approx 0.093$.