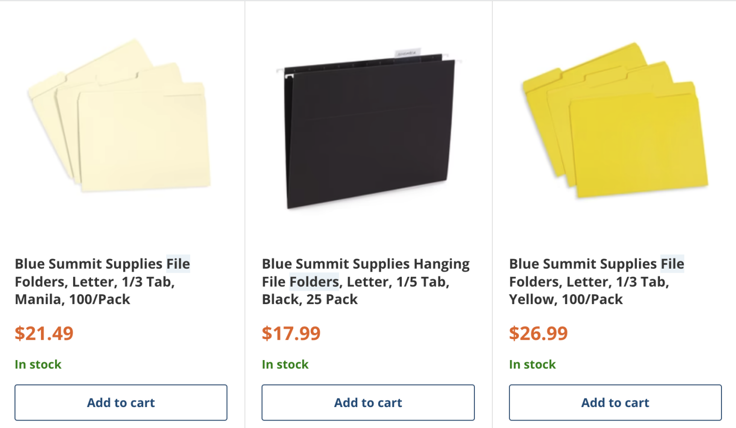
Reliable Offline Pricing in eCommerce Decision-Making: A Distributionally Robust Viewpoint

Jie Wang
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

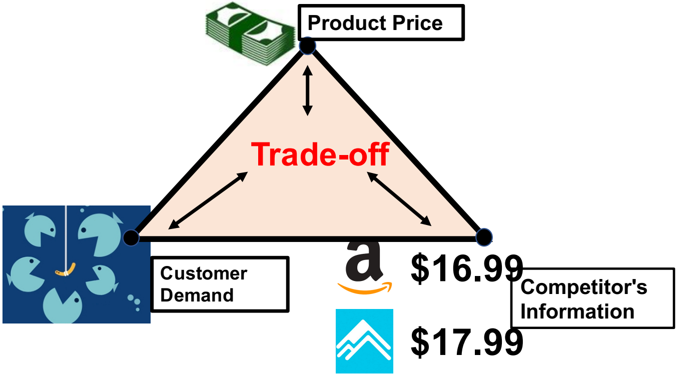
Date: October 14, 2023

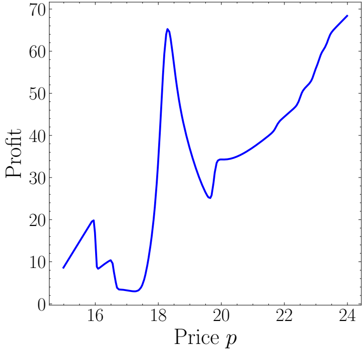
Data-Driven Pricing



- **Unknown**: Customer Demand
- **Decision**: Price offered
- **Information**: Historical data on demand, price, competitor's price, inventory level, etc.

Challenges in Data-Driven Pricing





- Inaccurate and Unreliable Prediction of Demand
- Nonconvex Optimization Due to Decision-Dependent Uncertainty

Profit Model

- Profit $=$ Sales $-$ (Cost-of-goods-sold $+$ eCommerence Fee $+$ Referral Fee $+$ Ad Spend).

Sales	$p(D \wedge y)$
Cost-of-goods-sold (COGs)	$a_1(D \wedge y)$
eCommerence Fee (FBA)	$a_2(D \wedge y)$
Referral Fee (REFFEE)	$15\% \cdot p(D \wedge y)$
Ad Spend	Constant ind. of any variable

  - y: inventory level
  - p: product price
  - z: competitor's price (Cprice)
  - D: customer demand, following demand distribution $f_D(p, z)$ for fixed (p, z).
- Reward function for given (y, D, p):

$$85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y).$$

Profit Model

- Profit = Sales − (Cost-of-goods-sold + eCommerence Fee + Referral Fee + Ad Spend).

Sales	$p(D \wedge y)$
Cost-of-goods-sold (COGs)	$a_1(D \wedge y)$
eCommerence Fee (FBA)	$a_2(D \wedge y)$
Referral Fee (REFFEE)	$15\% \cdot p(D \wedge y)$
Ad Spend	Constant ind. of any variable

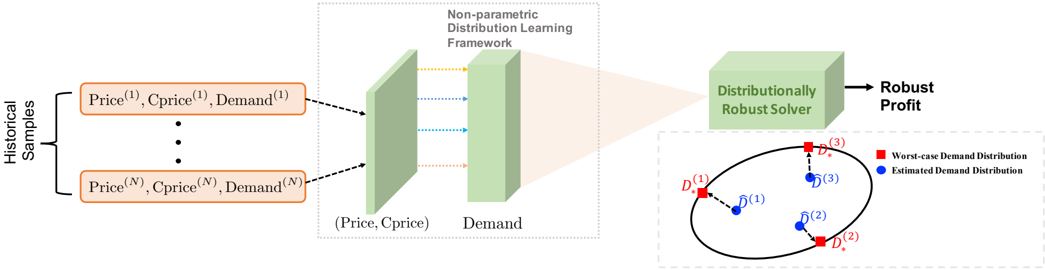
  - y: inventory level
  - p: product price
  - z: competitor's price (Cprice)
  - D: customer demand, following demand distribution $f_D(p, z)$ for fixed (p, z).
- **Goal**: find price p to maximize the expected reward function:

$$\mathbb{E}_{D \sim f_D(p,z)} \left[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \right].$$

Cons: $f_D(p, z)$ is unknown!

General Framework

$$\text{Optimal (Expected) Profit} = \max_p \left\{ \mathbb{E}_{D\sim f_D(p,z)}\Big[85\% \cdot p(D \wedge y) - (a_1 + a_2)\cdot (D\wedge y)\Big] \right\}.$$



- Step (I): **Non-parametric** prediction of $f_D(p,z)$ for given (p,z).
- Step (II): **Interval estimate** of profit via **distributionally robust optimization (DRO)**.

Step (I): Learning Demand Distribution

$$\text{Optimal (Expected) Profit} = \max_{p} \left\{ \mathbb{E}_{D \sim f_D(p,z)} \left[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \right] \right\}.$$

- **Given**: historical samples on $\{(p^i, z^i, D^i)\}_{i=1}^{N}$.

Step (I): Learning Demand Distribution

$$\text{Optimal (Expected) Profit} = \max_p \left\{ \mathbb{E}_{D \sim f_D(p,z)} \left[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \right] \right\}.$$

- **Given**: historical samples on $\{(p^i, z^i, D^i)\}_{i=1}^N$.
- Parzen-window density estimate of demand distribution:

$$\widehat{f}_D(p,z) = \sum_{i=1}^N \omega^i(p,z) \mathbf{1}\{D = D_i\},$$

$$\omega^i(p,z) \propto \exp\left(-\frac{\|(p,z) - (p^i, z^i)\|^2}{2\sigma^2} \right), \qquad \sum_{i=1}^N \omega^i(p,z) = 1$$

Under certain assumptions and with $\sigma^2 = O(N^{-\delta}), \delta \in (0, 1/\dim((p,z)))$, predicted profit based on $\widehat{f}_D(p,z)$ **converges** to the optimal profit as $N \to \infty$.[1]

[1] Ref: D. Bertsimas and N. Kallus, From Predictive to Prescriptive Analytics, Management Science (2019)

Step (II): Distributionally Robust/Optimistic Predictions

- **Motivation**: Demand estimator in Step (I) may be **unreliable**.
- **Distributionally Robust counterpart**:

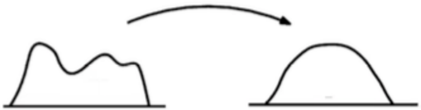
$$\max_{p} \min_{f_D(p,z)} \mathbb{E}_{D\sim f_D(p,z)}\Big[r(p,y;D) \triangleq 85\% \cdot p(D\wedge y) - (a_1+a_2)\cdot(D\wedge y)\Big]$$

$$\text{s.t.} \quad \mathrm{supp}(f_D(p,z)) \subseteq \mathrm{supp}(\widehat{f}_D(p,z))$$

$$\mathcal{W}\Big(f_D(p,z), \widehat{f}_D(p,z)\Big) \le \epsilon$$

$$\mathcal{W}(\mathbb{P},\mathbb{Q}) = \min_{\gamma} \ \left(\mathbb{E}[\|\omega-\omega'\|^2]\right)^{1/2}$$

$$\text{s.t.} \quad \gamma \text{ is a joint distribution of } (\omega,\omega') \text{ with mariginal distributions } \mathbb{P} \text{ and } \mathbb{Q}$$



Tractability of DRO

For fixed (p, z, y), **robust profit** is reformulated as

$$
\begin{aligned}
\min_{\gamma \in \mathbb{R}^{N\times N}, \nu \in \mathbb{R}_+^N} \quad & \sum_{i=1}^{N} \nu_i r(p, y; D^i) \\
\text{s.t.} \quad & \sum_{i=1}^{N} \gamma_{i,j} = \nu_j, \forall j, \sum_{j=1}^{N} \gamma_{i,j} = \omega^i(p, z), \forall i \\
& \sum_{i,j=1}^{N} \gamma_{i,j} |D^i - D^j|^2 \le \epsilon^2
\end{aligned}
$$

- γ: optimal transport map
- ν: worst-case demand distribution
- $r(p, y; D) = 85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y)$

- **Optimal robust profit prediction**: Enumerate all choices of p.

Regularization Effect of Robust Profit

$$\text{Optimal (Expected) Profit} = \max_{p} \left\{ \mathbb{E}_{D\sim f_D(p,z)} \Big[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \Big] \right\}.$$

With $\epsilon = O(1/\sqrt{N})$, the distributionally robust profit is equivalent to [2]

$$\max_{p} \left\{ \mathbb{E}_{D\sim \widehat{f}_D(p,z)} \Big[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \Big] \right.$$
$$\left. - \epsilon \left| 85\% \cdot p - (a_1 + a_2) \right| \cdot \mathbb{E}_{D\sim \widehat{f}_D(p,z)} [\mathbf{1}\{D \leq y\}] + \tilde{O}(1/N) \right\}.$$

Remark: regularization penalies aggresive pricing strategy.

[2]Ref: R. Gao, X. Chen, and A. Kleywegt, Wasserstein Distributionally Robust Optimization and Variation Regularization (2022)

Regularization Effect of Optimistic Profit

$$\text{Optimal (Expected) Profit} = \max_{p} \left\{ \mathbb{E}_{D\sim f_D(p,z)} \Big[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \Big] \right\}.$$

With $\epsilon = O(1/\sqrt{N})$, the distributionally optimistic profit is equivalent to [3]

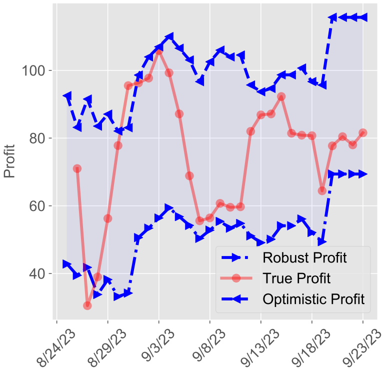
$$\max_{p} \Bigg\{ \mathbb{E}_{D\sim \widehat{f}_D(p,z)} \Big[85\% \cdot p(D \wedge y) - (a_1 + a_2) \cdot (D \wedge y) \Big] + \epsilon \left| 85\% \cdot p - (a_1 + a_2) \right| \cdot \mathbb{E}_{D\sim \widehat{f}_D(p,z)} [\mathbf{1}\{D \leq y\}] + \tilde{O}(1/N) \Bigg\}.$$

Remark: regularization balances the trade-off between **price maximization** and **customer demand satisfaction**.

[3]Ref: R. Gao, X. Chen, and A. Kleywegt, Wasserstein Distributionally Robust Optimization and Variation Regularization (2022)

Numerical Study on Profit Estimation

- **Training data**: historical data from 2/16/2022 to 8/23/2023 for `File Folders SKU 21`.
- **Testing data**: remaining historical data.

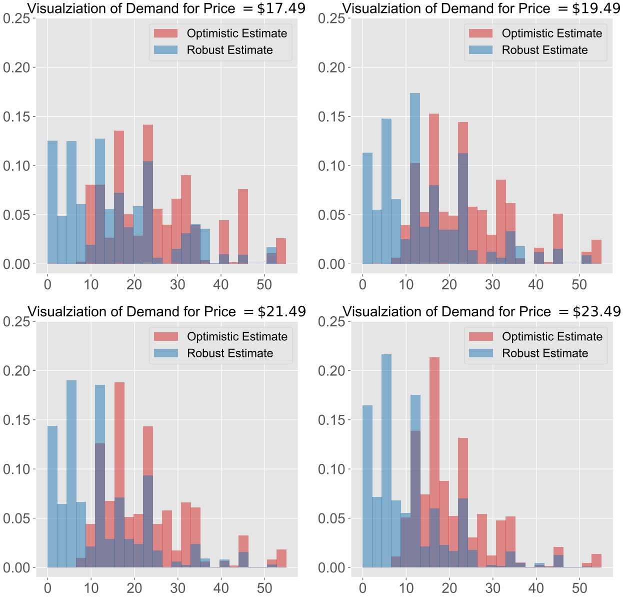


Interval profit estimation by distributionally robust/optimistic solver

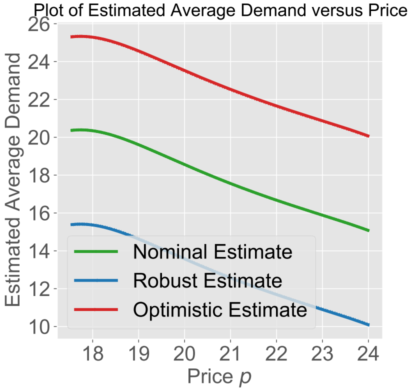
Take Away

Actual profit locates in our estimated interval with high probability and small interval width.

Visualization of Demand Distribution



Visualization of Demand Distribution

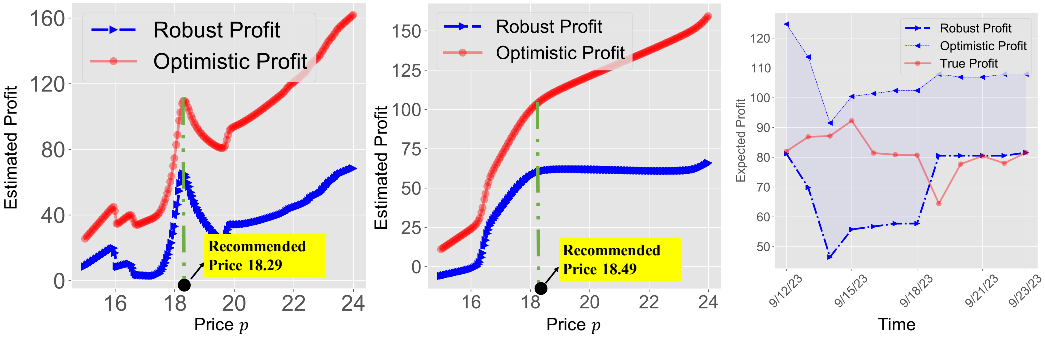


Mean of demand over all price choices

Take Away

A non-parametric, data-driven estimator of the relation between demand v.s. price

Optimal Pricing in Week1/Week2



Contributions

- **Contextual Optimization with Sample Average Approximation**

  Data-driven prediction on demand distribution, with theoretical guarantees

- **Distributionally Robust/Optimistic Optimization**

  Reliable interval estimate of optimal profit

  Regularization effects on robust profit

Source code and full manuscript:
https://github.com/WalterBabyRudin/BSS-Data-Challenge-Solution/tree/main