Submitted to *the INFORMS 2023 BSS Data Challenge Competition*

FOR INTERNAL USE ONLY (leave blank)
Submission Number:
Team Number:

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

Jie Wang
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, jwang3163@gatech.edu

Offline pricing is an essential problem in eCommerce decision-making, aiming to estimate an optimal price that maximizes the expected profit based on historical data. Due to the decision-dependent effect and lack of market environment information, it is imperative to develop a flexible and robust pricing strategy. In this study, we propose a novel two-step procedure to tackle this challenge. In the first step, we provide a data-driven and non-parametric distributional estimate of customer demand. In the second step, we provide a robust profit estimate that maintains satisfactory performance even with inaccurate demand estimators. Our proposed framework is also supported by empirical analysis. Notably, its numerical performance on the `File Folders SKU 21` product has the impressive rank of 2 among all File Folders SKU products, as per official rankings.

---

1. Introduction

Pricing for eCommerce products has been an important research topic in recent years. Compared with the conventional newsvendor problem that aims to decide about an order quantity for profit optimization, the pricing problem is more challenging due to the *decision-dependent effect*: The decision of order quantity for the former problem does not influence the distribution of random parameters, while the decision for the latter problem can affect the customer demand distribution. Such an effect causes difficulty in both computational traceability and model estimation. Even worse, due to the limited information, it is difficult to establish an accurate relation between the unit price and the customer demand distribution. If the retailer adopts the classical Predict-then-Optimize framework [3] that first estimates such a relation and then optimizes pricing strategy, the non-negligible estimation error in the first step will amplify the optimization error in the second step.

In this report, we propose a *distributionally robust contextual optimization model* with the *decision-dependent Wasserstein ambiguity set* to solve this problem. Specifically, we examine the performance among all possible choices of price in a two-step procedure and pick the price that achieves the best performance. In the first step, we adopt from [1] to develop a data-driven and non-parametric approximation of the customer demand distribution based on historical samples on price, demand, and Cprice (i.e., competitor's price). In the second step, we estimate the worst-case expected profit, where the worst-case means we take into account all candidate demand distributions that are close

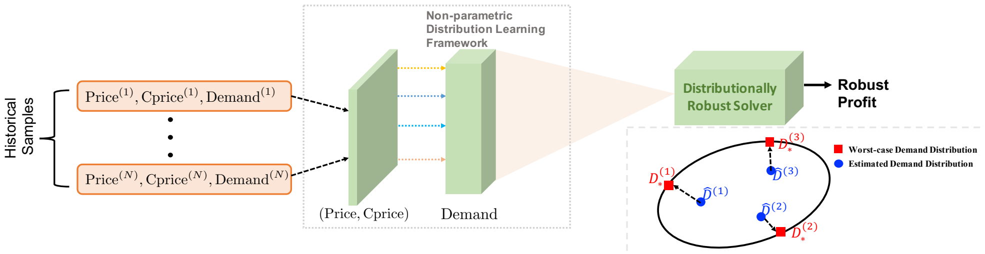

Figure 1 An overview of the proposed framework, which consists of two critical components: (a) a non-parametric framework that estimates the demand distribution based on historical samples; (b) a distributionally robust solver that estimates a robust profit based on the estimated demand distribution.

to the estimated demand distribution with respect to the Wasserstein metric. We summarize the general framework in Figure 1, and more details are provided in Section 2.

2. Methodology

Consider a retailer who sells a single product with finite amount of inventory. Assume the amount of inventory available for sale, denoted as $y \in \mathbb{R}_+$ is known to the retailer, and he/her determines the price $p \in [p^l, p^u]$ for sale. Besides, the retailer is provided with the cprice (i.e., competitors' price) $z \in \mathbb{R}^M$ with $M \in \mathbb{N}_+$, which can be viewed as a covariate variable. Here we assume that the price decision p and the cprice z will influence the customer demand D distribution, which follows the probability distribution $f_D(p,z)$. The goal is to select the optimal price such that the expected profit $\mathbb{E}_{D\sim f_D(p,z)}[c(p,D)]$ is maximized. In the following, we provide the detailed expression regarding the profit function $c(p,D)$.

Profit Model. It is known that the profit depends on *sales*, *Cost of Goods Sold* (COGs), *eCommerce Fee* (FBA), *Referral Fee* (REFFEE), and *Ad Spend* (ADSPEND). Given the price p, customer demand D, and inventory level y, we know sales $= p(D\wedge y)$. We make the following assumption regarding the other four variables:

Assumption 1. COGs $= a_1\cdot(D\wedge y)$; FBA $= a_2\cdot(D\wedge y)$; REFFEE $= 15\%\cdot p(D\wedge y)$; ADSPEND *is a constant independent of any other variable.*

Here, we validate these assumptions using the provided SKU dataset `File Folders SKU 21`. It turns out that these models fit the data with the high confidence level. Specifically, the coefficients $a_1 = 4.43965609, a_2 = 6.60097822$. Based on Assumption 1, we re-write the loss function $c(p,D) = 0.85\cdot p(D\wedge y) - (a_1+a_2)\cdot(D\wedge y) - \text{ADSPEND}$. Consider the ideal case where the demand distribution $f_D(p,z)$ is exactly known for any price p and cprice z, the profit maximization problem becomes

$$\max_{p\in[p^l,p^u]} \mathbb{E}_{D\sim f_D(p,z)}\Big[0.85\cdot p(D\wedge y) - (a_1+a_2)\cdot(D\wedge y)\Big] - \text{ADSPEND}. \tag{Ideal}$$

In a practical case, one may not have full information regarding the inventory level y, the cprice (i.e., competitors' prices) z, the variable ADSPEND, and the demand distribution $f_D(p,z)$. To tackle this challenge, we make the following simplifications:

- We find in most situations, the inventory level y is always greater than the units ordered, which motivates us to assume that $D \ll y$. Thereby one can replace $D \wedge y$ with y in problem (Ideal).
- The cprice z can be estimated using historical data based on time series prediction (e.g., based on python package `Skforecast`).
- From Figure 2, we find the variable ADSPEND seems to be correlated with variable unitsordered. Therefore, we estimate its value using time series prediction and treating variable unitsordered as an exogenous feature.
- The demand distribution $f_D(p,z)$ is difficult to obtain. In the following, we provide a non-parametric and data-driven way for estimating $f_D(p,z)$.

Demand Distribution. We provide a data-driven estimation of the demand distribution $f_D(p, z)$, in which only data samples $\{(p^i,z^i,D^i)\}_{i=1}^N$ are available. For a given price-cprice pair (p,z), we approximate the true demand distribution using the weighted discrete distribution sharing the same support of the training dataset:

$$\widehat{\mathbb{P}}_n^{(p,z)} = \sum_{i=1}^N \omega^i(p,z)\delta_{D^i},$$

where the weight function $\{\omega^i(p,z)\}_{i=1}^N$ are obtained using the (Gaussian) kernel regression model:

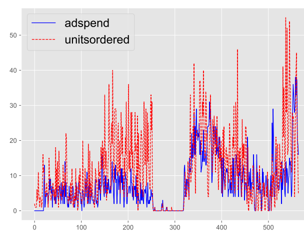

Figure 2. Plot of ADSPEND **and** Unitsordered **versus time index for product** `File Folders SKU 21`.

$$\omega^i(p,z) \propto \exp\left(\frac{-\|(p,z)-(p^i,z^i)\|_2^2}{2\sigma^2}\right), \quad \sum_{i=1}^N \omega^i(p,z) = 1. \tag{1}$$

Previous studies have proposed several modeling for customer demand, such as additive linear model [6, 2] and multiplicative model [4, 5]. However, it is vague whether those models result in an accurate estimation of the demand in our setting because the cprice z is involved. Instead, we use a data-driven estimator in this part, which is quite flexible as we *'let the data speak for itself.'*

Distributionally Robust Formulation. A natural pricing idea is to solve the sample average approximation (SAA) counterpart of problem (Ideal), i.e., by replacing $f_D(p,z)$ with the estimated distribution $\widehat{\mathbb{P}}_n$ and solve the approximated problem. However, SAA may not achieve satisfactory performance because the estimate $\widehat{\mathbb{P}}_n$ is not accurate enough. Instead, we solve the distributionally robust counterpart of the SAA problem using the 2-Wasserstein distance to model the ambiguity set:

$$\max_{p\in[p^l,p^u]} \left\{ \min_{\mathbb{P}:\ \mathrm{supp}(\mathbb{P})\subseteq \mathrm{supp}(\widehat{\mathbb{P}}_n^{(p,z)})} \mathbb{E}_{D\sim\mathbb{P}}[c(p,D)] :\ \mathcal{W}\left(\mathbb{P},\widehat{\mathbb{P}}_n^{(p,z)}\right) \le \epsilon \right\}. \tag{DRO-Pricing}$$

Here the 2-Wasserstein distance is defined as

$$\mathcal{W}(\mathbb{P},\mathbb{Q}) := \min_{\gamma} \left\{ \left(\mathbb{E}[\|\omega-\omega'\|^2]\right)^{1/2} :\ \begin{array}{l}\gamma \text{ is a joint distribution of } (\omega,\omega') \\ \text{with mariginal distributions } \mathbb{P} \text{ and } \mathbb{Q}\end{array} \right\}.$$

For fixed price p and cprice z, according to the definition of Wasserstein distance, the inner minimization problem can be reformulated as a linear program that can be solved efficiently:

$$\min_{\gamma\in\mathbb{R}^{N\times N},\nu\in\mathbb{R}_+^N}\left\{\sum_{i=1}^N \nu_i c(p,D^i):\ \left\{\begin{array}{l}\sum_{i=1}^N\gamma_{i,j}=\nu_j,\forall j,\sum_{j=1}^N\gamma_{i,j}=\omega^i(p,z),\forall i,\\ \sum_{i,j=1}^N\gamma_{i,j}|D^i-D^j|^2\le\epsilon^2\end{array}\right\}\right\}. \tag{2}$$

It is worth mentioning that (DRO-Pricing) is a nonconvex optimization problem because the decision variable p influences both the objective function and the demand distribution. However, it is required that the optimized pricing value should end in 0.05 or 0.09. Therefore, we solve problem (DRO-Pricing) by enumerating all possible choices of price p, whereas for each value p one only needs to solve a linear programming formulation (2). Similarly, one can obtain an optimistic pricing strategy by replacing the inner minimization in problem (DRO-Pricing) with maximization.

Remark 1 (The Spirit of Robustness). By solving problem (DRO-Pricing), the retailer tends to find a price to maximize his/her profit under the worst scenario among all customer demand distributions. Even if the true demand distribution is shifted slightly compared to the estimated demand distribution due to the dynamic market environment, a distributionally robust pricing strategy could still provide a lower bound on the actual profit. On the supply side, there is less room for the retailer to make a wrong pricing strategy with limited information on the demand distribution because of the marginal cost. Therefore, a distributionally robust strategy seems preferable.

3. Numerical Study

The bandwidth parameter σ^2 in (1) and the radius of ambiguity set ϵ in (DRO-Pricing) plays a critical impact on the performance of our model. In our numerical study, we tune these hyper-parameters using cross-validation: we split the given dataset into the training and validation sets. For each candidate pair (σ^2,ϵ), we formulate the estimated profit for each sample in the validation set based on the training set. Then, we take the optimal hyper-parameter as the one that achieves the smallest residual error between the estimated profit and the true profit in the validation set.

3.1. Profit Estimation

We use the historical data from the last 30 days as the testing data and use the remaining data as the training data to examine the performance of our profit estimation algorithm. Figure 3 reports the plot of true expected profit [1] together with estimated robust and optimistic profit. From the plot, we find our profit estimation algorithm works reasonably well: the true profit is generally covered by [robust profit,optimistic profit] with a reasonably tight interval width.

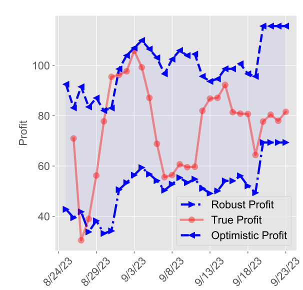

Figure 3. Plot of estimated profit versus true profit in last 30 days.

[1] We take the average of profit within the nearest 7 days as the true expected profit.

3.2. Visualization of Demand Distribution

Next, we visualize the estimated robust demand distribution for various price choices and fix the cprice $z = 22.99$ by treating the provided dataset as the training set. See the visualization in Figure 4. From the plot, we find that customer demand tends to decrease as the price value increases, which is consistent with our intuition.

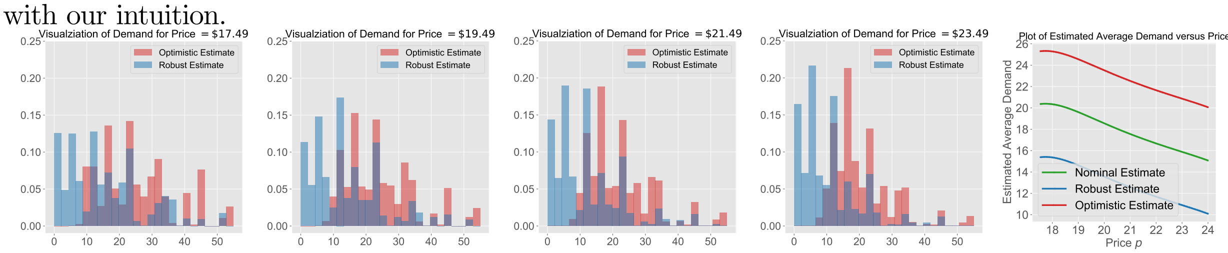

Figure 4 The four left plots visualize the estimated demand distribution for different prices. The rightmost plot reports the estimated *average* demand across different choices of price.

3.3. Optimal Pricing

Finally, we present the average of estimated profits across different price choices for Week 1 (Sep.12-18) and Week 2 (Sep.19-25) in Figure 5. From the plots, we realize that price 23.99 results in the best robust profit, but its interval width (optimistic profit minus robust profit) is also the largest, indicating it is not a stable and robust choice. Instead, we prefer to take a price that results in (nearly) the best robust profit but with a significantly small interval width. Therefore, we take the price 18.29 in Week 1 and price 18.49 in Week 2. Figure 5(c) plots the estimated profit and true profit during these two weeks, from which we can see that the numerical performance of our pricing and profit estimation strategy works reasonably well.

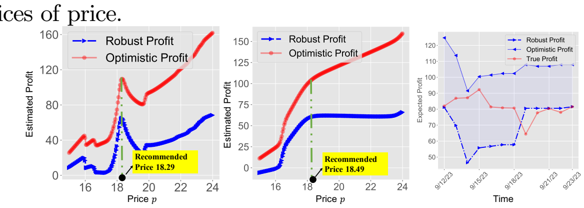

Figure 5. Left two: estimated profit for different prices during Week 1/Week 2 testing phase; Right: Estimated profit v.s. true profit.

4. Conclusion

We develop a data-driven pricing strategy that not only yields an optimal price but also provides interval estimation of the expected profit. Its practical performance on the `File Folders SKU 21` product has the second-highest rank among all File Folders SKU products, as validated by official rankings.

5. Team Member(s)

- Jie Wang, Georgia Institute of Technology, Atlanta, GA 30332, jwang3163@gatech.edu

References

[1] Bertsimas D, Kallus N (2020) From predictive to prescriptive analytics. *Management Science* 66(3):1025–1044.

[2] Bu J, Simchi-Levi D, Wang C (2022) Context-based dynamic pricing with partially linear demand model. *Advances in Neural Information Processing Systems* 35:23780–23791.

[3] Elmachtoub AN, Grigas P (2022) Smart predict, then optimize. *Management Science* 68(1):9–26.

[4] Kazaz B, Webster S (2015) Price-setting newsvendor problems with uncertain supply and risk aversion. *Operations Research* 63(4):807–811.

[5] Salinger M, Ampudia M (2011) Simple economics of the price-setting newsvendor problem. *Management Science* 57(11):1996–1998.

[6] Wang C, Chen X (2015) Optimal ordering policy for a price-setting newsvendor with option contracts under demand uncertainty. *International Journal of Production Research* 53(20):6279–6293.