A Bayesian Approach to Contextual Dynamic Pricing using the Proportional Hazards Model with Discrete Price Data

Dongguen Kim[1], Young-Geun Choi[2], and Minwoo Chae[1]

[1]Pohang University of Science and Technology, [2]Sungkyunkwan University

NeurIPS 2025

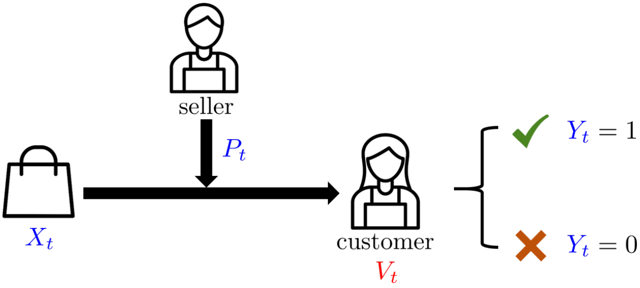
Introduction

- Recently, many firms have used sales data to determine prices that maximize revenue.
  - It is called contextual dynamic pricing.

Contextual dynamic pricing problem

At each time $t = 1, \ldots, T$,

- $X_t \in \mathbb{R}^d$: the context vector of the product and customer
- P_t: the price offered by the seller
- $V_t \in \mathbb{R}_{\geq 0}$: the customer's valuation of the product (**unknown**)
- $Y_t = \mathbb{1}\{V_t > P_t\}$: feedback on whether the customer purchases the product



Contextual dynamic pricing problem (cont.)

- π_t: (pricing) policy at time t
- **Goal**: Design a policy $\{\pi_t\}_{t=1}^{T}$ to minimize the cumulative regret over a given time horizon T:

$$R(T) = \sum_{t=1}^{T} \{ER(P_t^*, X_t) - ER(P_t, X_t)\},$$

where $ER(p, X_t)$ is the **expected revenue** from offered price p given X_t:

$$ER(p, X_t) := \mathbb{E}(p \cdot Y_t \mid X_t) = p \cdot \mathbb{P}(V_t > p \mid X_t) = p \cdot S(p \mid X_t),$$

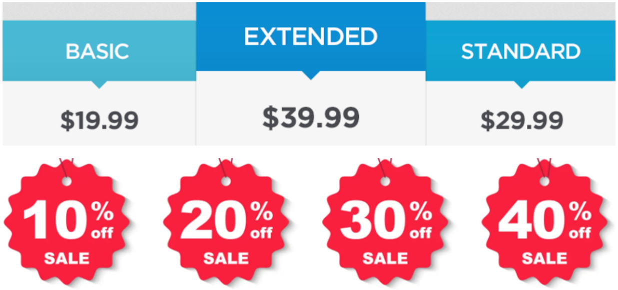
and $S(v \mid X_t)$ is the **survival function** of V_t given X_t:

$$S(v \mid X_t) := \mathbb{P}(V_t > v \mid X_t),$$

and P_t^* is the **optimal price** at time t.

Motivation: Discrete price

- In real-world applications, prices are often observed only on a discrete set.
  - Ex) magic number endings, fixed dicount levels
- However, much of the existing dynamic pricing literature focuses on continuous price spaces,
  - That is, $P_t \in \mathcal{P}$, where $\mathcal{P}$ is continuous subset of $\mathbb{R}_{\geq 0}$.



Summary of our contributions

- **Question**: *Can we improve dynamic pricing performance by leveraging discrete price information?*
- **Answer**: *Yes!*

Method	Model for V_t	Regret Upper Bound	Optimality in T	Adaptation to discrete support
Fan et al. (2024)	Linear	$\widetilde{O}((Td)^{\frac{2m+1}{4m-1}})$	×	×
Luo et al. (2024)	Linear	$\widetilde{O}(T^{\frac{2}{3}} + \Vert\widehat{\beta} - \beta^*\Vert_1 T)$	×	×
Luo et al. (2022)	Linear	$\widetilde{O}(T^{\frac{2}{3}} d^2)$	×	×
Luo et al. (2022)	Linear	$\widetilde{O}(T^{\frac{3}{4}} d)$	×	×
Shah et al. (2019)	Log-linear	$\widetilde{O}(T^{\frac{1}{2}} d^{\frac{11}{4}})$	×	×
Choi et al. (2023)	PH	$\widetilde{O}(T^{\frac{2}{3}} d)$	○	×
Our work	PH	$\widetilde{O}(T^{\frac{1+\gamma}{2}} + (dT)^{\frac{1}{2}})$ $(\gamma < 1/3)$ $\widetilde{O}(T^{\frac{2}{3}} + (dT)^{\frac{1}{2}})$ $(\gamma \geq 1/3)$	○	○

Discrete price setting

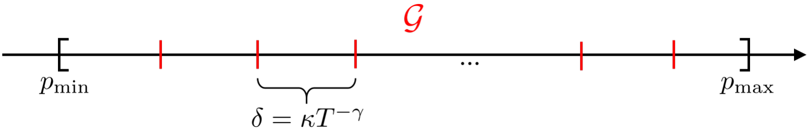
- Assume that P_t is supported on a **discrete grid**

$$\mathcal{G} := \{g_k : k = 1, \ldots, K\} \subset [p_{\min}, p_{\max}] \quad \forall t = 1, \ldots, T,$$

where

$$g_k - g_{k-1} = \delta, \qquad \delta := \kappa T^{-\gamma},$$

and γ is (unknown) **sparsity level** of the grid.
  - $\gamma = 0$: (fixed) discrete regime; $\gamma = 1$: continuous regime



- **Goal**: Propose the algorithm which adapts to γ.

Model

- $Y_t = \mathbb{1}\{V_t > P_t\}$ is called **case 1 interval-censored data** in survival analysis.
- We model the survival function of V_t given X_t using the **Cox proportional hazards (PH)** model:

$$S(v \mid X_t) = S_0(v)^{\exp(X_t^\top \beta)},$$

where $S_0 : [v_{\min}, v_{\max}] \to [0, 1]$ and $\beta \in \mathbb{R}^d$ are the parameters of interest.

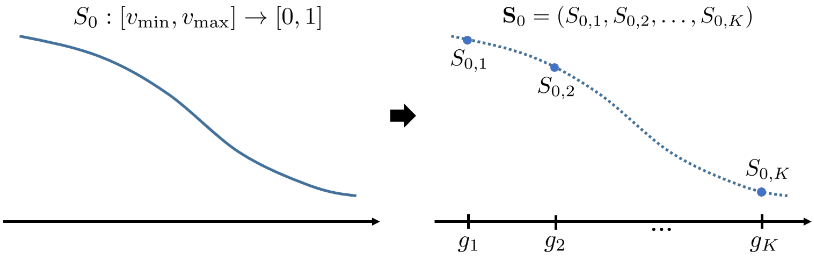
Key role of discrete set $\mathcal{G}$

- Since P_t is supported on $\mathcal{G}$, we can reduce the parameter space as follows:

$$\Theta = \{\theta = (\mathbf{S}_0, \beta) \in \mathcal{S}_0 \times \mathbb{R}^d\},$$

where

$$\mathcal{S}_0 = \{\mathbf{S}_0 = (S_{0,1}, \ldots, S_{0,K}) \in [0,1]^K : 1 > S_{0,1} \geq \cdots \geq S_{0,K} > 0\}.$$



Results

- We consider suitable prior Π on $\mathbf{S}_0$ and β.

THEOREM [Posterior convergence rate]
Let $\mathbf{D}_n = \{X_t, P_t, Y_t\}_{t=1}^n$ be i.i.d. samples and $\Pi(\cdot \mid \mathbf{D}_n)$ be the posterior. Let

$$\epsilon_n = \begin{cases} \widetilde{O}\left(n^{-\frac{1-\gamma}{2}} + \sqrt{d/n}\right) & \text{if } \gamma < 1/3, \\ \widetilde{O}\left(n^{-\frac{1}{3}} + \sqrt{d/n}\right) & \text{if } \gamma \geq 1/3. \end{cases}$$

Under some assumptions, we have

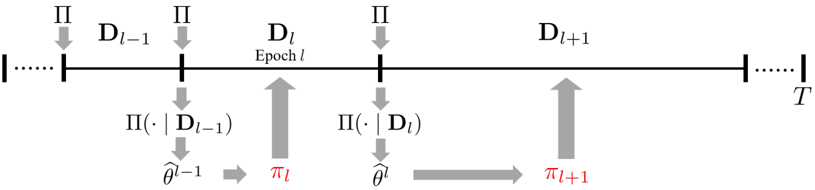
$$\Pi(\|\mathbf{S}_0 - \mathbf{S}_0^*\|_2 + \|\beta - \beta^*\|_2 \geq \epsilon_n \mid \mathbf{D}_n) \leq \exp(-n\epsilon_n^2),$$

with high probability.

Algorithm: BayesCoxCP

We employ an **epoch-based design** policy.

- $\mathcal{E}_l \subset [T]$: the set of time indices for epoch l
- $\hat{\theta}^l = (\hat{S}_0^l, \hat{\beta}^l)$: the point estimator derived from the posterior $\Pi(\cdot \mid \mathbf{D}_l)$
- π_l: the policy for epoch l using $\hat{\theta}^{l-1}$



Algorithm: BayesCoxCP (cont.)

- π_l is defined as:

$$\pi_l(X_t) = (1-\eta_l)\cdot \delta_{\widehat{P}_t^{l-1}} + \underbrace{\eta_l \cdot \mathbb{U}_{\mathcal{G}}}_{\text{provide necessary exploration}},$$

where $\widehat{P}_t^{l-1}$ is the **myopic price** determined by $\widehat{\theta}^{l-1}$:

$$\widehat{P}_t^{l-1} \in \operatorname*{argmax}_{p\in\mathcal{G}} \left\{ p \cdot \widehat{S}_0^{l-1}(p)^{\exp(X_t^\top \widehat{\beta}^{l-1})} \right\},$$

and η_l is the epoch-specific **exploration parameter**:

$$\eta_l = \min\left\{ \eta_1 \left(\eta_2 \sqrt{\frac{|\mathcal{G}|}{2^{l-1}}} \wedge 2^{-\frac{l-1}{3}} \right) \sqrt{\log 2^{l-1}}, 1 \right\}.$$

Regret analysis

THEOREM [Regret upper bound]
Under some assumptions, we have

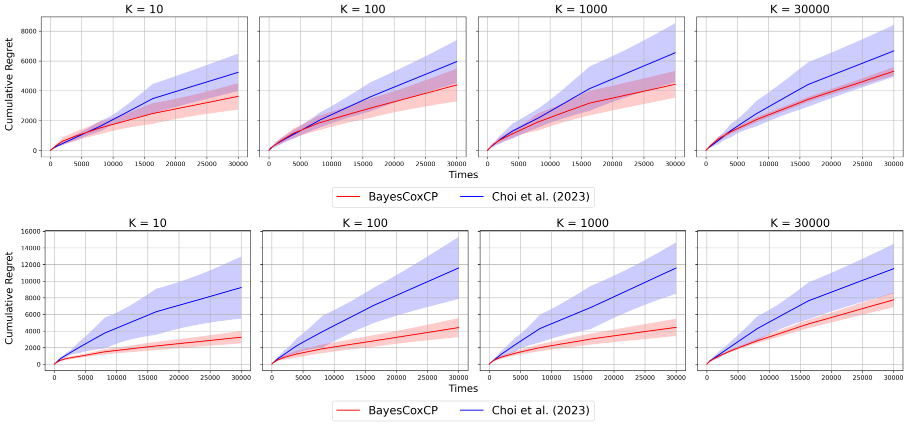
$$R(T) \leq \begin{cases} \widetilde{O}\left(T^{\frac{\gamma+1}{2}} + \sqrt{dT}\right) & \text{if } \gamma < 1/3, \\ \widetilde{O}\left(T^{\frac{2}{3}} + \sqrt{dT}\right) & \text{if } \gamma \geq 1/3, \end{cases}$$

with high probability.

THEOREM [Regret lower bound for non-contextual pricing]
In a non-contextual discrete pricing problem, for any $\eta > 0$, no pricing policy (algorithm) can achieve expected regret $O(T^{\frac{1+\gamma}{2}-\eta})$ if $\gamma < 1/3$, and $O(T^{\frac{2}{3}-\eta})$ if $\gamma \geq 1/3$.

Numerical experiments

- To highlight the benefits of leveraging discrete support information, we compare our method with [Choi et al. 2023][1], which considers the PH model in the continuous setting.



[1] Young-Geun Choi et al. (2023). “Semi-parametric contextual pricing algorithm using Cox proportional hazards model”. In: *Proc. International Conference on Machine Learning*, pp. 5771–5786.

Thank You!