

Assignment 2 - Price Optimization Using Bandits

Deadline: 28 Mar, 2025

You have a service that you don't know how to price. Specifically, there is a “base price” P , and you can choose among the following options:

- $p_0 = \alpha^2 \cdot P$
- $p_1 = \alpha \cdot P$
- $p_2 = P$
- $p_3 = \beta \cdot P$
- $p_4 = \beta^2 \cdot P$

Assume $\alpha = 0.5, \beta = 1.5$.

You have a competitor for this service who gives a random price Q to each incoming user drawn from a uniform distribution in $[0, 4]$. However you **neither know** your competitor's policy beforehand, **nor** can you observe the price it offers at each round.

Question 1 - User Picks Randomly: Assume a user arrives at each round t (there are $T = 10000$ total rounds). We must show a price to this user among the above 5 choices. The user however is “naive” and picks randomly between the two sellers (it is as if the user **first** agrees to buy the service, and **then** hears the price) **You don't know that this is her policy though!!**. If that user picks you, you “earn” this price as a reward. If that user picks the competitor, you earn 0 reward.

Assuming that a total of $T = 10000$ users will arrive, one after the other, following the same random policy, apply Multiplicative Weights (MW) to try to optimize your pricing policy (i.e. maximize the amount of profit you'll make after T rounds.) What do you observe about the learned policy and the profits you make?

Question 2 - User Becomes Smart: Assume now that the average user start choosing randomly again, until $T/4$. However, at that point she realizes that one competitor might(?) have learned to overprice its service, and decides from now to ask both competitors **first** for a price, **then** chooses the one offering

the **lowest price!** Once more, you **do not know** that users choose competitors in this manner. Yet, you still have to figure out what's the “best” price to charge each incoming user over time using the MW algorithm. Do you expect to make better profits in this scenario or the previous one? Comment on your findings

For all questions you should choose again a random seed equal to the last two digits of your student ID and pick the base price P at the beginning, uniformly in $[1.5, 2.5]$.

As usual, you need to turn in your (Colab executable) notebook code, and appropriate plots and commentary (either in text cells, inline with your notebook) or in a separate pdf.