# Joint Learning & Dynamic Pricing



- Dynamic pricing without knowing the demand function - Besbes & Zeevi
- Dynamic pricing with limited supply - Babaioff, Dughmi, Kleinberg, Slivkins

## Model

- $k$ items, $n$ agents, sequential arrival $(k < n)$

  (Alt: time horizon $T$, arrival rate $\lambda$ Poisson process)

- Posted price $P_t$ for agent $t$, $p \in [0,1]$
- Agent $t$ has value $V_t \sim F$. $F$ unknown, $\in [0,1]$
- $S(p) = 1 - F(p)$ - Sales rate / quantile. $S(p)$ strictly dec

  $R(p) = p\, S(p)$ - Revenue function

  $F$ regular $\Rightarrow R''(p) \leq 0 \quad \forall p \in [0,1]$

  $F$ strictly regular $\Rightarrow R''(p) < 0$

- $P^*_{MP} \equiv$ Monopolist price $\left(\arg\max_p R(p)\right)$
- Fixed price benchmark - $A^n_k(p)$. If $k = \infty$, $\cancel{E[A^n_k(p)]}$ $Rev(A^n_k(p)) = R(p)\, n$

$\boxed{\text{Lemma 1}}$ let $\nu(p) = p \min(k, nS(p)) = \min(kp, nR(p))$

then $\nu(p) - O(p\sqrt{k \log k}) \leq \underbrace{Rev(A^n_k(p))}_{E[\text{Revenue by } A^n(p)]} \leq \nu(p)$

thus the best choice of fixed price is $p^* = \arg\max_p \partial(p)$

$$= \max\left[P_{MP}, S^{-1}\left(\frac{k}{n}\right)\right]$$

Pf - The upper bound is obvious. For the lower bound, let

$$X_t \equiv \mathbb{1}\{\text{Sale to } t^{th} \text{ agent}\}, \quad X = \sum_{t=1}^{n} X_t, \quad \mu = \mathbb{E}[X] \quad (= n S(p))$$

By Chernoff bounds, $\mathbb{P}\left[X - \mu \leq -O(\sqrt{\mu \log k})\right] \leq \frac{1}{k}$

$$\Rightarrow \# \text{ of sales} = \mathbb{E}\left[\min[k, X]\right] \geq \min\left(k, \mu - O(\sqrt{\mu \log k})\right)$$

$$\geq \min(k, \mu) - O(\sqrt{k \log k})$$

---

Capped UCB $(n, k)$

$$\boxed{\text{Regret} = \partial(p^*) - \mathbb{E}[Rev(\mathcal{A})]}$$

- Choose parameter $\delta \in (0,1)$
- Set 'active prices' set $\mathcal{P} \equiv \{\underbrace{\delta(1+\delta)^i}_{\in[0,1]}; i \in \mathbb{N}\}$
- While $\exists$ unsold item
  - Pick $p \in \arg\max_{p \in \mathcal{P}} \underbrace{I_t(p)}_{\text{index for price } p}$
- Else set $p = \infty$ (close shop!)

**Thm** - With $\delta = k^{-1/3}(\log n)^{2/3}$, Capped UCB has regret $= O((k \log n)^{2/3})$ for <u>any</u> distribution (regular)

- For any distrib (regular), $\exists$ $S_F, C_F$ s.t Capped UCB with $\delta = \sqrt{k} \log n$ achieves regret $O(C_F \sqrt{k} \log n)$ when $\frac{k}{n} \leq S_F$. For MMR,

# Notes

- The choice of index $I_t$ depends on $(n, k)$, not instantaneous state
- Benchmark is $\gamma(p)$, not $\mathbb{E}\left[\mathrm{Rev}\left(A_n^k(p)\right)\right]$ – the previous lemma shows this is justified
- Need a refined UCB to handle 'rare' prices
- Index should incorporate 'stock-out price' $S^{-1}\left(\frac{k}{n}\right)$ in addition to $p_{MP}$

- More details for CappedUCB$(n,k)$

  – $$I_t(p) \triangleq p \cdot \min\left[k, \underbrace{n\, S_t^{UB}(p)}_{\text{UCB}}\right]$$

  $$S_t^{UB}(p) \triangleq \underbrace{\hat{S}_t(p)}_{\text{empirical rate}} + \underbrace{r_t(p)}_{\text{confidence radius}}$$

  $$\hat{S}_t(p) = \min\left\{\frac{R_t(p)}{N_t(p)}, \underbrace{1}_{\text{when } N_t(p)=0}\right\},$$

  $R_t(p) = \#$ of sales at price $p$

  $N_t(p) = \#$ of agents offered price $p$

  $$r_t(p) = \frac{c \log n}{N_t(p) + 1} + \sqrt{\frac{c \log n\, \hat{S}_t(p)}{N_t(p) + 1}}$$

  Henceforth, we define $\alpha = c \log n$

# Pf of $O((k \log n)^{2/3})$ (worst-case) regret

- $X_t = \mathbb{1}\{\text{Sale to } t^{th} \text{ agent}\} \sim \text{Bin}(S(p_t))$

  $X = \sum_{t=1}^{n} X_t \, , \quad S \triangleq E[X] = \sum_{t=1}^{n} S(p_t)$

  Suppose we ignore the capacity $k$ to define $X$ (i.e., we continue to sell after running out of items)

  Then $Rev = \sum_{t=1}^{N} p_t X_t$, where $N = \max\{N \leq n \mid \sum_{t=1}^{N} X_t \leq k\}$

- Pf outline - First we define a set of 'good events' and analyze a deterministic algo under these. Then we bound the probability of 'not good events' to show small loss in regret.

- $\boxed{\text{Lemma 2}}$ With probability at least $1-n^{-2}$, the following are true: for all $t \in \{1,2,\ldots,T\}$, $p \in \mathcal{P}$

  i) $|S(p) - \hat{S}_t(p)| \leq r_t(p) \leq 3\left(\frac{\alpha}{N_t(p)+1} + \sqrt{\frac{\alpha S_t(p)}{N_t(p)+1}}\right)$

  And ii) $|X - S| < O(\sqrt{S \log n} + \log n)$

  iii) $\left|\sum_{t=1}^{n} p_t (X_t - S(p_t))\right| < O(\sqrt{S \log n} + \log n)$

Note - the last two bounds depend on $S$, not $n$

- Henceforth, assume i, ii and iii are TRUE

- Define $p^*_{act} \equiv \arg\max_{p \in P} \left[ \nu(p) \right]$ (best active price)

$$\Delta(p) \equiv \max\Big(0, \underbrace{\frac{1}{n}\nu(p^*_{act}) - p\,S(p)}\Big)$$

similar to $\mu^* - \mu_i$ in UCB

$$N(p) \equiv N_{n+1}(p) = \text{total } \# \text{ of agents offered } p$$

---

**Lemma 3** - $\forall p \in P$, $N(p)\Delta(p) \leq O(\log n)\left[1 + \frac{k}{n} \cdot \frac{1}{\Delta(p)}\right]$

Pf - by def$^n$, $\forall t, p \in P$ - $|S(p) - \hat{S}_t(p)| \leq r_t(p)$

$$\Rightarrow \nu(p) \leq I_t(p) \leq p \cdot \min\left[k, n\left(S(p) + 2r_t(p)\right)\right]$$

Thus - $\left[\begin{array}{l} I_t(p_t) \geq \nu(p^*_{act}) \\ I_t(p_t) \leq p_t \cdot \min\left[k, n\left(S(p_t) + 2r_t(p_t)\right)\right] \end{array}\right.$ (Choose highest $I_t(p)$!)

$$\Rightarrow \frac{\nu(p^*_{act})}{n} \leq p_t \min\left[\frac{k}{n}, S(p_t) + 2r_t(p_t)\right]$$

$\Rightarrow$ i) $p_t \geq \frac{\nu^*_{act}}{k}$, ii) $\Delta(p_t) \geq 2p_t r_t(p_t)$, iii) $\Delta(p_t) > 0$
$\Rightarrow S(p_t) < \frac{k}{n}$

Now we use the form of $r_t(\cdot)$

Consider $t$, the last time price $p$ was chosen (for any $p \in P$)

$\Rightarrow$ - $N(p) = N_t(p) + 1$ ] by defn

- $\Delta(p) \leq 2p\, r_t(p)$
- $\Delta(p) > 0 \Rightarrow S_t(p) < \frac{k}{n}$

] from above

- $r_t(p) \leq 3\left(\frac{O(\log n)}{N(p)} + \sqrt{\frac{S_t(p)}{N(p)}}\right)$

Combining, we get $\Delta(p) \leq O(p) \cdot \max\left[\frac{\log n}{N(p)}, \sqrt{\frac{k}{n}\frac{\log n}{N(p)}}\right]$

---

- Next, instead of Rev, we want to analyze $\sum_{t=1}^{n} p_t S(p_t)$ (i.e., ignore capacity constraints)

<u>Lemma</u> - Let $\beta(s) = O(\sqrt{s \log n} + \log n)$

then $\text{Rev} \geq \min\left(\mathcal{D}(p^*_{act}), \sum_{t=1}^{n} p_t S(p_t)\right) - \beta(k)$

Pf - $\because p_t \geq \frac{\mathcal{D}(p^*_{act})}{k}\ \forall t \Rightarrow \text{Rev} \geq \mathcal{D}(p^*_{act})$ if $\sum_{t=1}^{n} x_t > k$

- If $\sum_{t=1}^{n} x_t < k$, then $\text{Rev} = \sum_{t=1}^{n} p_t x_t \geq \underbrace{\sum_{t=1}^{n} p_t S(p_t) - \beta(s)}_{\text{from the assumed high prob events}}$

- Now let's consider $\sum p_t S(p_t)$

$$\sum_{t=1}^{n} p_t S(p_t) \geq \sum_{t=1}^{n} \left[\frac{\nu(p^*_{act})}{n} - \Delta(p_t)\right]$$

$$= \nu(p^*_{act}) - \sum_{p \in P} \Delta(p) N(p)$$

Define $P_{sel} = \{p \in P \mid N(p) \geq 1\}$ (prices selected at least once)

$P_\epsilon = \{p \in P_{sel} \mid \Delta(p) > \epsilon\}$ (will choose $\epsilon$ later)

$$\Rightarrow \sum_{p \in P} N(p)\Delta(p) = \sum_{p \in P(\epsilon)} \Delta(p) N(p) + \sum_{p \notin P(\epsilon)} \Delta(p) N(p)$$

$$\leq O(\log n)\left[|P_\epsilon| + \frac{k}{n}\sum_{p \in P_\epsilon} \frac{1}{\Delta(p)}\right] + \epsilon n$$

- Thus, for any set $P$, any $\epsilon > 0$

$$\boxed{\nu(p^*_{act}) - \mathbb{E}[Rev] \leq \epsilon n + O(\log n)\left[|P_\epsilon| + \frac{k}{n}\sum_{p \in P_\epsilon} \Delta(p)^{-1}\right] + \beta(k)}$$

- For $P \equiv \{\delta(1+\delta)^i\}$, $\boxed{\nu(p^*_{act}) - \nu(p^*) \geq -\delta k}$
  - If $p^* < \delta$, $\nu(p^*) \leq \delta k$. Else let $p_0 \equiv \max\{p \in P, p \leq p^*\}$

$$\Rightarrow \frac{p_0}{p} \geq \frac{1}{1+\delta} \geq 1-\delta \Rightarrow \nu(p^*_{act}) \geq \nu(p_0) \geq \frac{p_0}{p^*}\nu(p^*) \geq \nu^* - \delta k$$

- Putting everything together, we get

$$\text{Regret} = \nu^* - E[\text{Rev}]$$

$$\leq \epsilon n + O(\log n)\left[|P_\epsilon| + \frac{k}{n}\sum_{p \in P_\epsilon} \Delta(p)^{-1}\right] + \beta(k) + \delta k$$

$$\leq O(\log n)\left[|P_\epsilon|\left(1 + \frac{k}{\epsilon n}\right)\right] + O\left(\sqrt{k \log n} + \log n\right) + \delta k + \epsilon n$$

Also $|P_\epsilon| \leq \frac{1}{\delta} \log n$. assume $\delta \geq \frac{1}{n}$, $\epsilon = \delta \frac{k}{n}$

$$\Rightarrow \quad \text{Regret} \leq O\left(\delta k + \frac{1}{\delta^2}(\log n)^2 + \sqrt{k \log n}\right)$$

Choosing $\delta = k^{-1/3} (\log n)^{2/3} \Rightarrow \text{Regret} = O\left((k \log n)^{2/3}\right)$

# Concentration Bounds in Babaioff et al.

* Standard Chernoff bounds - For $X_1, X_2, \ldots, X_n \in [0,1]$, iid, $X = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X]$. Then

  i) $\mathbb{P}[|X-\mu| \geq \epsilon\mu] \leq 2e^{-n\mu\epsilon^2/3}$ $\quad \forall\, 0<\epsilon<1$

  ii) $\mathbb{P}[X > a] < 2^{-an}$ for $a > 6\mu$

* <u>Lemma</u> - In the above setting, let $r(\alpha, x) = \frac{\alpha}{n} + \sqrt{\frac{\alpha x}{n}}$.

  Then $\mathbb{P}\left[|X-\mu| < r(\alpha, X) < 3r(\alpha,\mu)\right] > 1 - e^{-\Omega(\alpha)}$

<u>Pf</u> - The main idea is to separately deal with small and large $\mu$.

i) Consider $\mu > \alpha/6n$. Let $\epsilon = \frac{1}{2}\sqrt{\frac{\alpha}{6\mu n}}$. Now by (i)

$$\mathbb{P}[|X-\mu| \geq \epsilon n] \leq 2\exp\left(\frac{-n\mu\alpha}{72\mu n}\right) = 2e^{-c\alpha}$$

$$\Rightarrow |X-\mu| < \mu\epsilon < \mu/2 \quad \text{w.p } 1-e^{-\Omega(\alpha)}$$

Also by choice of $\epsilon$, we have w.p $1-e^{-\Omega(\alpha)}$

$$|X-\mu| < \frac{\mu}{2}\sqrt{\frac{\alpha}{6\mu n}} \leq \sqrt{\frac{\alpha X}{n}} \leq r(\alpha, X) \leq 1.5\, r(\alpha,\mu)$$

ii) Consider $\mu < \alpha/6n$. Let $a = \alpha/n$; by (ii)

$$X < \alpha/n \quad \text{wp} \geq 1 - e^{-\Omega(\alpha)}$$

$$\Rightarrow |X - \mu| < \alpha/n < r(\alpha, X)$$

Finally $r(\alpha, X) = \frac{\alpha}{n} + \sqrt{\frac{\alpha X}{n}} < (1+\sqrt{2})\frac{\alpha}{n} < 3 r(\alpha, \mu)$

---

Now we return to our 'bad event' bounds

i) $|S(p) - \hat{S}_t(p)| \leq r_t(p) \leq 3\left(\frac{\alpha}{N_t(p)+1} + \sqrt{\frac{\alpha S_t(p)}{N_t(p)+1}}\right)$

Pf - For any $p \in P$, let $\{Z_{i,p}\}_{i \leq n} \equiv$ ~~Ber~~ $\text{Ber}(S(p))$ r.v.s

$\Rightarrow$ Sale to $i^{th}$ agent who sees $p$ $\equiv \{Z_{i,p} = 1\}$. Now we can use our lemma

$$\Rightarrow \mathbb{P}\left[|S(p) - \hat{S}_t(p)| \leq r_t(p) \leq 3\left(\frac{\alpha}{N_t(p)+1} + \sqrt{\frac{\alpha S_t(p)}{N_t(p)+1}}\right)\right] \geq 1 - n^{-4}$$

(if $\alpha = c \log n$)

Also $|P| \leq n$

$\Rightarrow$ By union bound over $t \in \{1, \ldots, n\}$, $p \in P$, we get the result.

---