CS 396: Online Markets

Lecture 6: Online Learning (Cont)

Last Time:

- online learning (cont)
- warmup up: be the leader
- perturbed follow the leader

Today:

- online learning (cont)
- warmup: geometric random variables
- perturbed follow the leader (analysis)
- multi-armed bandit learning

Exercise: Exponential Random Variables

Setup:

- X is exponentially distributed with rate $\lambda = 1$

Question: Evaluate

- its expectation $\mathbf{E}[X]$?
- its conditional expectation $\mathbf{E}[X \mid X > 1]$?
- its conditional expectation $\mathbf{E}[X \mid X > 100]$?

Geometric Distribution

"number of successful trials before success"

Model:

- coin with bias p (of heads)

  (a.k.a., "probability of success")

- flip coin until heads

  (a.k.a., "until success")

- geometric random variable X = "number of tails"

  (a.k.a., "number of failures")

Properties:

- possible values: $\{0, 1, \ldots, \}$
- probability mass function:

  $\mathbf{P}r[X = \ell] = (1-p)^\ell p$

- expectation: $\mathbf{E}[X] = {}^{1-p}/_{p}$

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{\ell=0}^{\infty} \ell (1-p)^\ell p \\
&= (1-p)\, p \sum_{\ell=0}^{\infty} \ell (1-p)^{\ell-1} \\
&= (1-p)\, p \sum_{\ell=0}^{\infty} \left[-\tfrac{d}{dp} (1-p)^\ell \right] \\
&= (1-p)\, p \, \frac{d}{dp} \left[- \sum_{\ell=0}^{\infty} (1-p)^\ell \right] \\
&= (1-p)\, p \, \tfrac{d}{dp} (-{}^{1}/_{p}) \\
&= (1-p)\, p \, ({}^{1}/_{p^2}) \\
&= {}^{1-p}/_{p}
\end{aligned}$$

- memoryless: $\mathbf{E}[X | X \geq \ell] = \ell + \mathbf{E}[X]$

Lemma 0: expected maximum of k geometric random variables is

$\mathbf{E}[\max_i X_i] = O(\frac{1}{p} \log k)$

Proof sketch:

- flip coins in rounds.
- discard failures
- stop when all coins are discarded
- how many rounds?

Example: $p = 1/2$, $k = k^0 = 124$

- first round: about half discarded $\Rightarrow k^1 \approx 64$
- second round: about half discarded $\Rightarrow k^2 \approx 32$
- third round: about half discarded $\Rightarrow k^3 \approx 16$

  ...

- $(\log_2 k)$th round: about none remaining.

Exercise: FTPL

Setup:

- FTPL (with $h = 1$ and $\epsilon = 1/2$)
  - let $V_j^i = \sum_{r=1}^{i} v_j^r$
  - hallucinate: $v_j^0 \sim$ "geometric with rate $1/2$''
  - in round i choose $j^i = \text{argmax}_j\, v_j^0 + V_j^{i-1}$
- **Input:**

	1	2	3
Action 1	$1/2$	0	0
Action 2	0	1	1

Question: On this input, what is probability of action 1:

- in round 1?
- in round 2?

(Hint for round 2: let p be "probability of choosing action i when i is ahead by $1/2$''; write recurrence for p; and solve.)

Analysis of FTPL

Alg: FTPL

- hallucinate: $v_j^0 = h\times$ "geometric with rate ϵ"
- let $V_j^i = \sum_{r=1}^{i} v_j^i$
- in round i choose: $j^i = \text{argmax}_j\, v_j^0 + V_j^{i-1}$

Thm: for payoffs in $[0, h]$,

$$\mathbf{E}[\text{FTPL}] \geq (1-\epsilon)\,\text{OPT} - \frac{h}{\epsilon}\ln k.$$

Lemma 1: (Stabilty) FTPL $\geq (1-\epsilon)$ BTPL

Lemma 2: (Small Perturbation)

BTPL $\geq$ OPT $-O(\frac{h}{\epsilon}\log k)$

Proof of Lemma 2: (intuition)

- maximum hallucination:

  MAXPTRB $= \mathbf{E}[\max_j v_j^0]$

- $\widetilde{\text{BTPL}}$ = BTPL payoff with perturbation
- $\widetilde{\text{OPT}}$ = OPT payoff with perturbation

$$\begin{aligned}\mathbf{E}[\text{BTPL}] + \mathbf{E}[\text{MAXPTRB}] &= \widetilde{\text{BTPL}}\\ &\geq \widetilde{\text{OPT}}\\ &\geq \text{OPT}\end{aligned}$$

- $\mathbf{E}[\text{BTPL}] \geq \text{OPT} - \mathbf{E}[\text{MAXPTRB}]$

  $\geq \text{OPT} - O(\frac{h}{\epsilon}\log k)$

Proof of Lemma 1:

- coupling argument
- start with raw scores
- add perturbation as:
  - pick action with lowest total score
  - flip coin:
    - heads: discard
    - tails: add h to score.
  - repeat until one action j^* left
- flip j^*'s coin:
  - tails: (w.p. $1-\epsilon$)
    - best action score $>$ h + second-best score
    - FTPL and BTPL pick j^*
  - heads:
    - doesn't matter.
- conclusion:

  FTPL picks same action as BTPL with probability at least $1-\epsilon$
- $\mathbf{E}[\text{FTPL}] \geq (1-\epsilon)\mathbf{E}[\text{BTPL}]$