Lecture 7

Exponential Mechanism — McSherry-Talwar 2007

Intro

Digital Goods auction

Seller: has movie, ebook, game ∞ copies

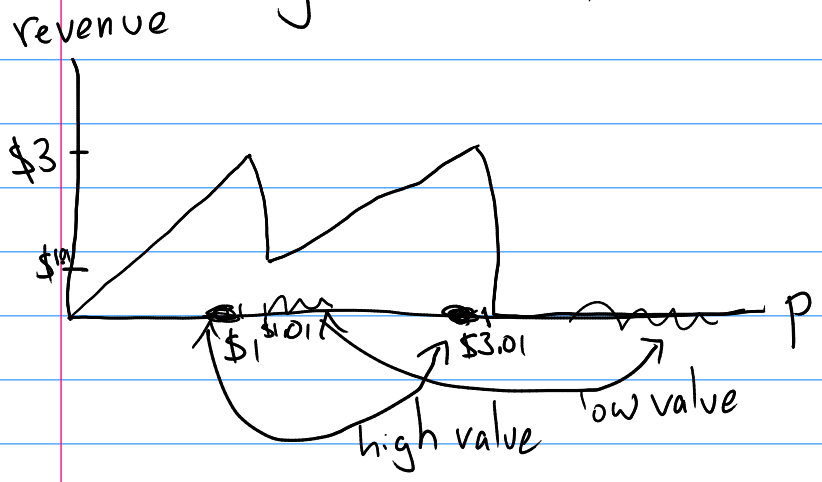
n Buyers: Value v_i for item

Choose price p → Buy if $v_i \geq p$

Else, not

$$\text{Revenue} = \sum_{i:\, p \leq v_i} p = p \cdot |\{i : p \leq v_i\}|$$

$n = 3$ buyers, $\vec{v} = (1, 1, 3.01)$



Setup

- Dataset: $X \in \mathcal{X}^n$ ← Private
- A set of objects: $\mathcal{H}$ ← Public
- Score function: $s: \mathcal{X}^n \times \mathcal{H} \to \mathbb{R}$ ← Public

X, X' nbs

$$\Delta s = \max_{h \in \mathcal{H}} \max_{X, X'} |s(X,h) - s(X',h)|$$

Δ

Exponential Mech: M_E, given inputs $X, \mathcal{H}, s$, output $h \in \mathcal{H}$ w.p. $\propto \exp\left(\frac{\varepsilon s(X,h)}{2\Delta}\right)$

$$\Pr[M_E(X) = h] = \frac{\exp\left(\frac{\varepsilon s(X,h)}{2\Delta}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(X,h')}{2\Delta}\right)}$$

- Computational

Privacy

M_E is ε-DP.

Fix $X, X', h \in \mathcal{H}$.

$$\frac{\Pr[M_E(X)=h]}{\Pr[M_E(X')=h]} = \frac{\left(\frac{\exp\left(\frac{\varepsilon s(X,h)}{2\Delta}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(X,h')}{2\Delta}\right)}\right)}{\left(\frac{\exp\left(\frac{\varepsilon s(X',h)}{2\Delta}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(X',h')}{2\Delta}\right)}\right)} \leq \exp\left(\frac{\varepsilon s(X,h)}{2\Delta}\right) \cdot \exp\left(\frac{\varepsilon}{2}\right)$$

$$= \exp\left(\frac{\varepsilon}{2\Delta}\left(s(X,h) - s(X',h)\right)\right)\left(\frac{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(X',h')}{2\Delta}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(X,h')}{2\Delta}\right)}\right)$$

$$\leq \exp\left(\frac{\varepsilon}{2}\right) \cdot \exp\left(\frac{\varepsilon}{2}\right) \cdot 1$$

$$= \exp(\varepsilon). \quad \square$$

Utility

Theorem 3. *Let X be a dataset, and $OPT(X) = \max_{h \in \mathcal{H}} s(X,h)$ be the score attained by the best object h with respect to the dataset X. For a dataset X, let $\mathcal{H}^* = \{h \in \mathcal{H} : s(X,h) = OPT(X)\}$ be the set of objects which achieve this score. Then*

$$\Pr\left[s(M_E(X)) \leq OPT(X) - \frac{2\Delta}{\varepsilon}\left(\ln\left(\frac{|\mathcal{H}|}{|\mathcal{H}^*|}\right) + t\right)\right] \leq \exp(-t).$$

$|\mathcal{H}^*| \geq 1$

Corollary 4.

$$\Pr\left[s(M_E(X)) \leq OPT(X) - \frac{2\Delta}{\varepsilon}\left(\ln\left(|\mathcal{H}|\right) + t\right)\right] \leq \exp(-t).$$

$$\Pr[s(M_E(X)) \leq c] = \sum_{h : s(X,h) \leq c} \frac{\exp\left(\frac{\varepsilon s(X,h)}{2\Delta}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(X,h')}{2\Delta}\right)}$$

$$\leq |\mathcal{H}| \cdot \frac{\exp\left(\frac{\varepsilon c}{2\Delta}\right)}{|\mathcal{H}^*| \exp\left(\frac{\varepsilon OPT(X)}{2\Delta}\right)} = \frac{|\mathcal{H}|}{|\mathcal{H}^*|} \exp\left(\frac{\varepsilon}{2\Delta}(c - OPT(X))\right) \square$$

Applications

Laplace Mechanism

X, $f(X)$, f is sens Δ

- Dataset: X
- $\mathcal{H} = \mathbb{R}$
- $s(x,h) = -|f(x) - h|$

$$\Pr[M_\varepsilon(x) = h] \propto \exp\left(-\frac{\varepsilon |f(x) - h|}{2\Delta}\right)$$

Selling one digital good

- Seller: ∞ copies, 1 item, choose $p \in [0,1]$
- n Buyers: $v_i \in [0,1]$

Sec 10.1 DR14

$$\text{Rev}(p) = p \cdot |\{i : v_i \geq p\}|$$

$$\gamma = \max_p p\,|\{i : v_i \geq p\}|$$

~~$\mathcal{H} = [0,1]$~~ $\mathcal{H} = \{\alpha, 2\alpha, \ldots, 1\}$, $|\mathcal{H}| = 1/\alpha$.

$\text{OPT}(v) \geq \gamma - \alpha n$

Data: v_i's,

$s(v,p) = p\,|\{i : v_i \geq p\}|$. $\Delta \leq 1$

$$\approx \text{OPT} - \frac{2\Delta}{\varepsilon}\log|\mathcal{H}| \geq (\gamma - \alpha n) - \frac{2}{\varepsilon}\log(1/\alpha) \geq \gamma - O\left(\frac{\log n}{\varepsilon}\right)$$

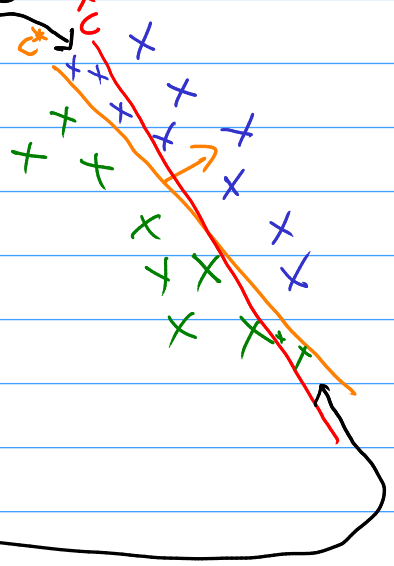
$$\alpha = \frac{\log n}{n\varepsilon}$$

Private PAC Learning Valiant '84

$\mathcal{C} = \{c : \{0,1\}^d \to \{0,1\}\}$

n elts: $(X_i, Y_i = c^*(X_i))$ unknown $c^* \in \mathcal{C}$, $X_i \sim D$

Goal: Output $\hat{c}$ s.t. $\Pr_{X \sim D}[\hat{c}(x) \neq c^*(x)]$ is small



Lemma: If $n \gg \frac{\log|\mathcal{C}|}{\alpha^2}$, then $\Pr_{X \sim D}[\hat{c}(x) \neq c^*(x)] \leq \frac{\alpha}{2}$.

No errs on train $\to \leq \frac{\alpha}{2}$ err on "population"

Fix $h \in \mathcal{C}$

$$\Pr_{X_1,\dots,X_n \sim D}\left[\left|\frac{|\{i : h(X_i) = c^*(X_i)\}|}{n} - \Pr_{X \sim D}[h(X_i) = c^*(X_i)]\right| \geq \frac{\alpha}{2}\right] \leq \exp(-\log|\mathcal{C}|) \leq \frac{1}{|\mathcal{C}|^{10}}$$

$\forall h \in \mathcal{C}$ at same time, emp frac of mistake $\approx$ "pop" $\pm \frac{\alpha}{2}$

w.p. $\geq 1 - \frac{1}{|\mathcal{C}|^9}$. ☒

$(X_i, Y_i) \to (X_i', Y_i')$, $X_i \sim D$

(X, Y), $\mathcal{H} = \mathcal{C}$, $s((X,Y), h) = -\frac{|\{i : h(X_i) \neq Y_i\}|}{n}$, $\Delta = \frac{1}{n}$

$s((X,Y), \hat{c}) \geq -\frac{2\Delta}{\varepsilon}\log|\mathcal{H}| = -\frac{2}{\varepsilon n}\log|\mathcal{C}| \geq -\alpha/2$

$n \gg \frac{\log|\mathcal{C}|}{\alpha\varepsilon}$

$\frac{|\{i : h(X_i) \neq Y_i\}|}{n} \leq \frac{\alpha}{2}$.

Thm: If $n \gg \frac{\log|\mathcal{C}|}{\alpha^2} + \frac{\log|\mathcal{C}|}{\alpha\varepsilon}$, then ε-DP alg.

$\Pr_{X \sim D}[\hat{c}(x) \neq c^*(x)] \leq \frac{\alpha}{2} + \frac{\alpha}{2} \leq \alpha$. ☒