The Exponential Mechanism
(and maybe some mechanism design)

Frank McSherry and Kunal Talwar
Microsoft Research, Silicon Valley Campus
Context: A data set $d \in D^N$ and mechanism $M : D^N \rightarrow R$.

Evaluating $M(d)$ shouldn’t give specific info about tuples in $d$.

Source of much definitional anxiety for some 30-odd years. What is specific info? Can we prevent everything/anything?

Definition: A mechanism $M$ gives $\epsilon$-differential privacy if:
For $d, d' \in D^N$ differing on at most one datum, and any $S \subseteq R$,

$$Pr[M(d) \in S] \leq \exp(\epsilon) \times Pr[M(d') \in S].$$

Changing one tuple can not change the output distribution much. Relative change in the probability of any event (subset $S$ of $R$).
**Previous Constructions**

**Simple scheme:** Apply $f : \mathcal{D}^N \rightarrow \mathbb{R}$ to data, return noisy result.

$$\mathcal{K}_f(DB) \equiv f(DB) + \text{Noise}.$$  

**Theorem:** Using Laplace($\sigma, 0$) gives $(\Delta f/\sigma)$-differential privacy,

$$\Delta f = \max_{DB} \max_{Me} \| f(DB - Me) - f(DB + Me) \|.$$  

For many statistical properties: $\Delta f$ is small, small noise benign.
Problems with Perturbation

**Pricing**: Inputs are $n$ bids in $[0, 1]$. Output is a price $p \in [0, 1]$. Want to make lots of money, but we don’t want to reveal bids.

**Problem**: Perturbing the true answer by some noise may fail.

1. The function may have high sensitivity. (eg: Pricing)
2. Perturbations may not actually be useful. (eg: Pricing)

**Moreover**: Additive perturbations also fail when

3. Outputs are not numbers. (eg: strings, trees, etc...)
Previously a “query” was $f : \mathcal{D}^N \rightarrow \mathbb{R}$, mapping data to result. Implicit assumption that results $r$ near $f(d)$ are nearly as good.

Now, a query is $q : (\mathcal{D}^N \times \mathcal{R}) \rightarrow \mathbb{R}$. Score of result $r$ for data $d$.

**Eg:** Given bids and a price, revenue is $q(d, r) = r \times \#(i : d_i > r)$. 
A General Mechanism

Previously a “query” was $f : \mathcal{D}^N \rightarrow \mathbb{R}$, mapping data to result. Implicit assumption that results $r$ near $f(d)$ are nearly as good.

Now, a query is $q : (\mathcal{D}^N \times \mathcal{R}) \rightarrow \mathbb{R}$. Score of result $r$ for data $d$.

**Eg:** Given bids and a price, revenue is $q(d, r) = r \times \#(i : d_i > r)$.

**Definition:** Let $\mathcal{E}^\epsilon_q(d)$ output $r$ with probability $\propto \exp(\epsilon q(d, r))$. 

A General Mechanism

Previously a “query” was $f : \mathcal{D}^N \rightarrow \mathbb{R}$, mapping data to result. Implicit assumption that results $r$ near $f(d)$ are nearly as good.

Now, a query is $q : (\mathcal{D}^N \times \mathcal{R}) \rightarrow \mathbb{R}$. Score of result $r$ for data $d$.

**Eg:** Given bids and a price, revenue is $q(d, r) = r \times \#(i : d_i > r)$.

![Graph](graph.png)

**Definition:** Let $\mathcal{E}_q^\epsilon(d)$ output $r$ with probability $\propto \exp(\epsilon q(d, r))$. 
Two Exciting Properties

Privacy: $\mathcal{E}_q^\epsilon$ gives $(2\epsilon\Delta q)$-differential privacy, where we define

$$\Delta q = \max_r \max_{d \approx d'} |q(d, r) - q(d', r)| .$$

Proof: Density, normalization alter by factors of at most $\exp(\epsilon\Delta q)$. 
Two Exciting Properties

Privacy: $\mathcal{E}_q^\epsilon$ gives $(2\epsilon\Delta q)$-differential privacy, where we define

$$\Delta q = \max_r \max_{d \approx d'} |q(d, r) - q(d', r)|.$$ 

Proof: Density, normalization alter by factors of at most $\exp(\epsilon\Delta q)$.

Utility: For $S \subseteq \mathcal{R}$, write $\mu(S)$ for its base measure. (pre-$\mathcal{E}_q^\epsilon$).
Two Exciting Properties

**Privacy:** $\mathcal{E}_q^\epsilon$ gives $(2\epsilon \Delta q)$-differential privacy, where we define

$$\Delta q = \max_r \max_{d \approx d'} |q(d, r) - q(d', r)|.$$  

Proof: Density, normalization alter by factors of at most $\exp(\epsilon \Delta q)$.

**Utility:** For $S \subseteq \mathcal{R}$, write $\mu(S)$ for its base measure.  (pre-$\mathcal{E}_q^\epsilon$).

**Lem:** Let $S_t = \{r : q(d, r) > OPT - t\}$. $\Pr(\overline{S}_{2t}) \leq \exp(-t)/\mu(S_t)$.

Proof: $LHS \leq \Pr(\overline{S}_{2t})/\Pr(S_t) \leq \exp(-t)\mu(\overline{S}_{2t})/\mu(S_t) \leq RHS$. 


**Two Exciting Properties**

**Privacy:** $E_q^\epsilon$ gives $(2\epsilon\Delta q)$-differential privacy, where we define

$$\Delta q = \max_r \max_{d \approx d'} |q(d, r) - q(d', r)|.$$  

Proof: Density, normalization alter by factors of at most $\exp(\epsilon\Delta q)$.

**Utility:** For $S \subseteq \mathcal{R}$, write $\mu(S)$ for its base measure. (pre-$E_q^\epsilon$).

**Lem:** Let $S_t = \{r : q(d, r) > OPT - t\}$. $\Pr(\overline{S}_{2t}) \leq \exp(-t)/\mu(S_t)$.

Proof: $LHS \leq \Pr(\overline{S}_{2t})/\Pr(S_t) \leq \exp(-t)\mu(\overline{S}_{2t})/\mu(S_t) \leq RHS$.

**Thm:** $E[q(d, E_q^\epsilon(d))] \geq OPT - 3t$, for those $t \geq \ln(OPT/t\mu(S_t))$.

Proof: $\Pr(OPT - 2t) \geq 1 - \exp(-t)/\mu(S_t) \geq 1 - t/OPT$. Multiply.
Applications to Pricing

Every bidder gives a demand curve: \( d_i : [0, 1] \to \mathbb{R}^+ \). \((rd_i(r) \leq 1)\)

**Theorem:** Taking \( q(d, r) = r \sum_i d_i(r) \), then the mechanism \( \mathcal{E}_q^\epsilon \) gives \((2\epsilon)\)-differential privacy, and has expected revenue at least \( OPT - 3 \ln(e + \epsilon^2 OPTm)/\epsilon \),

where \( m \) is the number of items sold at the optimal price.

**Proof:** Grind \( t = \ln(e + \epsilon^2 OPTm) \) through the previous theorem. Argue that \( \mu(S_t) \) is not small. (near-opt \( r \) gives near-opt \( q(d, r) \)).
Game Theory Implications

Differential Privacy implies many game-theoretic properties:

\[ Pr[M(d) \in S] \leq \exp(\epsilon) \times Pr[M(d')] \in S]. \]

\(\epsilon\)-Dominance: For any "utility" function \(g : R \rightarrow \mathbb{R}^+\),

\[ E[g(M(d))] \leq \exp(\epsilon) \times E[g(M(d'))]. \]

Collusion Resilient: For \(d \approx_t d'\), (ie: differing on \(t\) data)

\[ Pr[M(d) \in S] \leq \exp(\epsilon t) \times Pr[M(d') \in S]. \]

Repeatability: For \(M = (M_1, M_2, \ldots M_t)\) with dependencies,

\[ Pr[M(d) \in S] \leq \exp\left(\sum_{i \leq t} \epsilon_i\right) \times Pr[M(d') \in S]. \]

Truthful whp [CKMT]: \(M\) can be implemented so that:
For all \(d, t\), with prob \(\exp(-2\epsilon t)\), \(M(d) = M(d')\) for all \(d' \approx_t d\).
Stuff we did:

General mechanism $\mathcal{E}_q^e$, more robust, awesome than previously. Applications to Auctions/Pricing of various and new flavors. Neat non-truthful solution concept. Cool consequences.

Stuff we didn’t do / did badly:

Computational questions of sampling from $\mathcal{E}_q^e$ efficiently. Going beyond auctions/pricing to other mechanism problems.

Thanks! Questions?