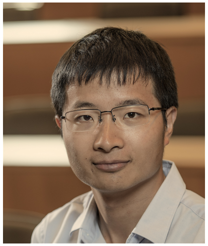
A Central Limit Theorem for Differentially Private Query Answering

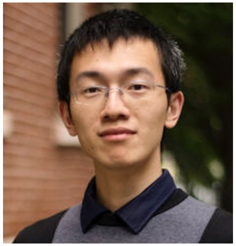
Jinshuo Dong

Northwestern University / IDEAL

Joint work with Weijie Su[†] and Linjun Zhang[§]

[†]University of Pennsylvania [§]Rutgers University





Our goal

- Differential Privacy:

  Hide individual details in the noise.
  Keep population information clean.

- Great success in recent years:

- Core question:

  Privacy-accuracy trade-off

- Many statistics/ML tasks:
  - Exists (ε, δ)-DP algorithm with error $\leqslant C \cdot \frac{\sqrt{\log \delta^{-1}}}{\varepsilon} \cdot \frac{d}{n}$
  - Any (ε, δ)-DP algorithm has error $\geqslant c \cdot \frac{\sqrt{\log \delta^{-1}}}{\varepsilon} \cdot \frac{d}{n}$
- Our goal: understand the constant, for the simplest problem

Query answering

Privacy-accuracy trade-off

- Query $f : D \mapsto \mathbb{R}$ or $\mathbb{R}^d$ where D is a dataset.
- Query answering: evaluate $f(D)$ privately.
- Noise addition mechanisms:
  - Generate r.v. X
  - $M(D) = f(D) + X$
- more privacy $\leftarrow$ larger $X \rightarrow$ less accuracy
- (Constant-sharp) Optimal noise under given privacy constraint?

Quiz: 1-dim

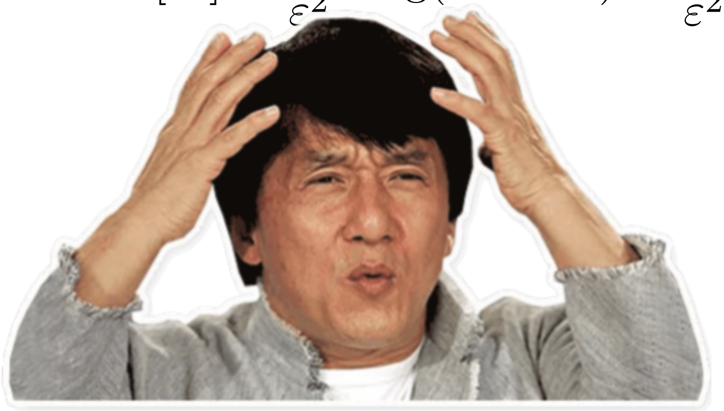
$$M(D) = f(D) + X.$$

- Accuracy is measured by $\mathrm{Var}[X]$.
- Question: What noise for $(\varepsilon, 0)$-DP?
- Textbook: Laplace noise [DMNS 06]

$$\mathrm{Var}[X] = \frac{2}{\varepsilon^2}$$

- Question: What if we relax by δ?
- Textbook: Gaussian noise [DKMMN 06]

$$\mathrm{Var}[X] = \frac{2}{\varepsilon^2} \cdot \log(1.25\delta^{-1}) > \frac{2}{\varepsilon^2}$$



Quiz: 1-dim

$$M(D) = f(D) + X.$$

- Accuracy is measured by $\mathrm{Var}[X]$.
- Question: What noise for $(\varepsilon, 0)$-DP?
- Everyone: Laplace noise [DMNS 06]

$$\mathrm{Var}[X] = \frac{2}{\varepsilon^2}$$

- Question: What if we relax by δ?
- Everyone: Gaussian noise [DKMMN 06]

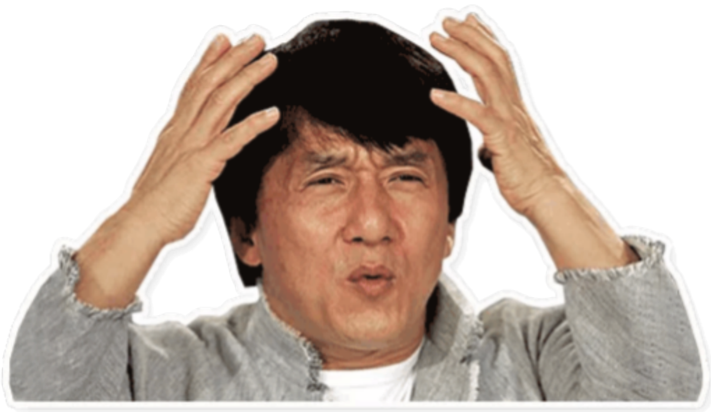
$$\mathrm{Var}[X] = \frac{2}{\varepsilon^2} \cdot \log(1.25\delta^{-1}) > \frac{2}{\varepsilon^2}$$

- (ε, δ) done right: truncated Laplace [GDGK 18]

  Truncate at $\pm h$ with $h = \log(1 + \frac{e^{\varepsilon}-1}{2\delta})$.

$$\mathrm{Var}[X] = \frac{2}{\varepsilon^2} \cdot \left(1 - \frac{\varepsilon^2 h(h+2)}{e^h - 1}\right) < \frac{2}{\varepsilon^2}$$

It took 12 years...



- This is a fundamental problem.
- We need a mindset that makes it simple.
- Here's how I visualize and reason about it.
- But we need a slightly more advanced perspective.

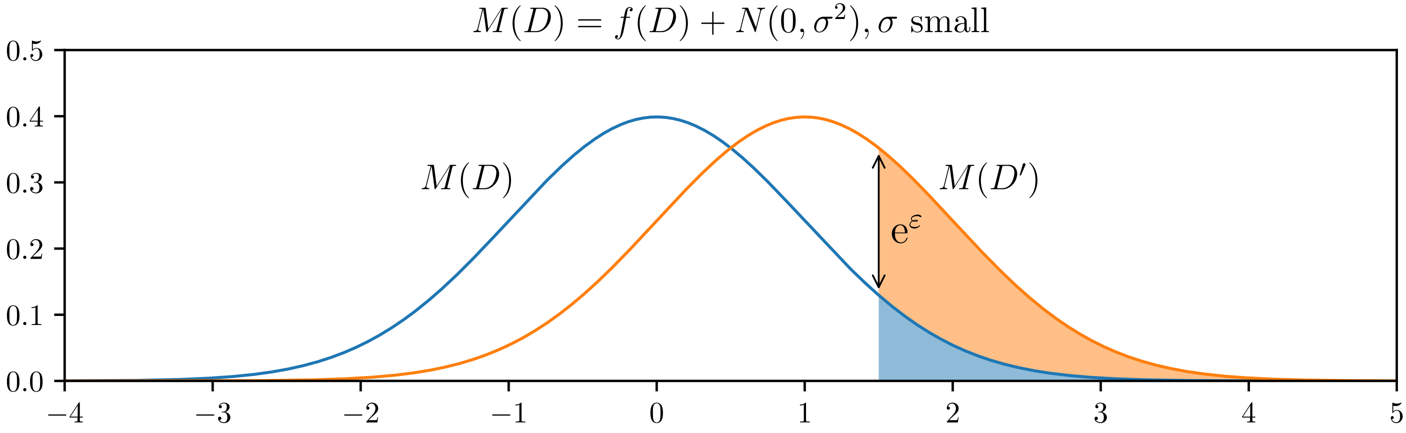
Recall: What is Differential Privacy?

Definition (DMNS 06, DKMMN 06)

A randomized algorithm $M : X \to Y$ is (ε, δ)-DP if

$$\mathbb{P}[M(D') \in E] \leqslant \mathrm{e}^{\varepsilon}\mathbb{P}[M(D) \in E] + \delta$$

- $E \subseteq Y$ is any event.
- D and D' are arbitrary neighboring databases that differ by one person



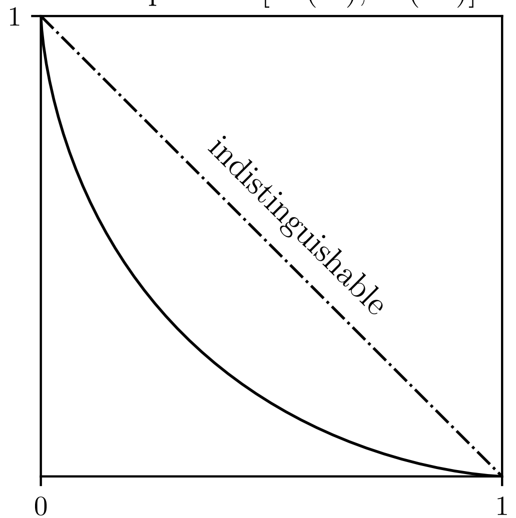
"Functional" perspective

$$\text{"}M(D) \approx M(D')\text{"}$$

- "Functional" perspective: A "true" δ for each ε

$$\delta(\varepsilon) = H_{e^\varepsilon}\big(M(D) \| M(D')\big) : \mathbb{R}_{>0} \to [0,1]$$

Example of $T[M(D), M(D')]$



- Equivalent via primal-dual
- Interpretation: FP vs FN in binary classification D vs D'
- Larger = more privacy
- [WZ 10, KOV 15, DRS 19]: M is (ε, δ)-DP iff $\underbrace{T\big[M(D), M(D')\big]}_{\text{ROC function}} \geqslant f_{\varepsilon,\delta}$

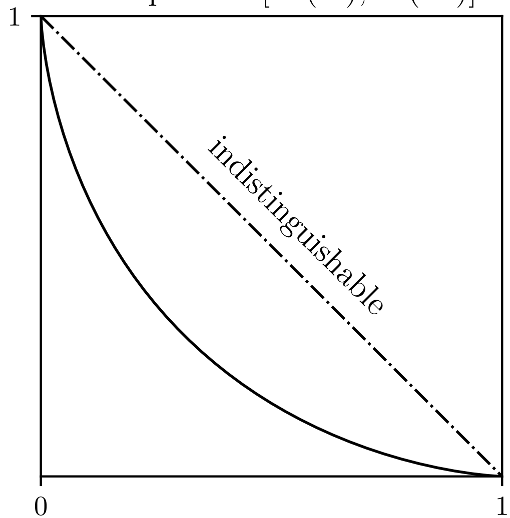
"Functional" perspective

$$\text{“}M(D) \approx M(D')\text{”}$$

- "Functional" perspective: A "true" δ for each ε

$$\delta(\varepsilon) = H_{e^\varepsilon}(M(D) \| M(D')) : \mathbb{R}_{>0} \to [0,1]$$
$$T[M(D), M(D')] : [0,1] \to [0,1] \qquad \text{([DRS 19])}$$

Example of $T[M(D), M(D')]$



- Equivalent via primal-dual
- Interpretation: FP vs FN in binary classification D vs D'
- Larger = more privacy
- [WZ 10, KOV 15, DRS 19]: M is (ε, δ)-DP iff $\underbrace{T[M(D), M(D')]}_{\text{ROC function}} \geqslant f_{\varepsilon,\delta}$

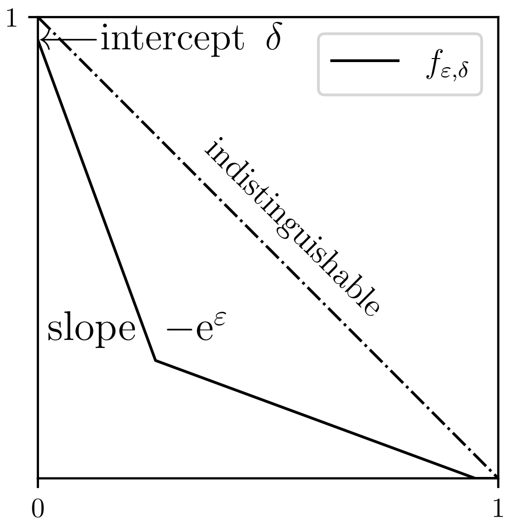
"Functional" perspective

"$M(D) \approx M(D')$"

- "Functional" perspective: A "true" δ for each ε

$$\delta(\varepsilon) = H_{e^\varepsilon}(M(D)\|M(D')) : \mathbb{R}_{>0} \to [0,1]$$

$$T[M(D), M(D')] : [0,1] \to [0,1] \qquad \text{([DRS 19])}$$



- Equivalent via primal-dual
- Interpretation: FP vs FN in binary classification D vs D'
- Larger = more privacy
- [WZ 10, KOV 15, DRS 19]: M is (ε, δ)-DP iff $\underbrace{T[M(D), M(D')]}_{\text{ROC function}} \geqslant f_{\varepsilon,\delta}$

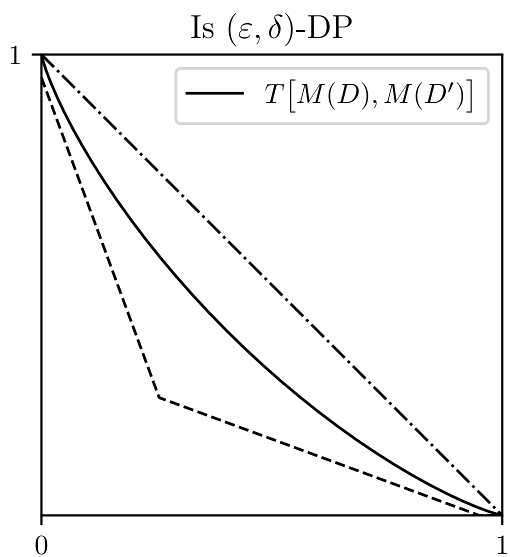
"Functional" perspective

"$M(D) \approx M(D')$"

- "Functional" perspective: A "true" δ for each ε

$$\delta(\varepsilon) = H_{e^\varepsilon}(M(D) \| M(D')) : \mathbb{R}_{>0} \to [0,1]$$

$$T[M(D), M(D')] : [0,1] \to [0,1] \qquad \text{([DRS 19])}$$



- Equivalent via primal-dual
- Interpretation: FP vs FN in binary classification D vs D'
- Larger = more privacy
- [WZ 10, KOV 15, DRS 19]: M is (ε, δ)-DP iff $\underbrace{T[M(D), M(D')]}_{\text{ROC function}} \geqslant f_{\varepsilon,\delta}$

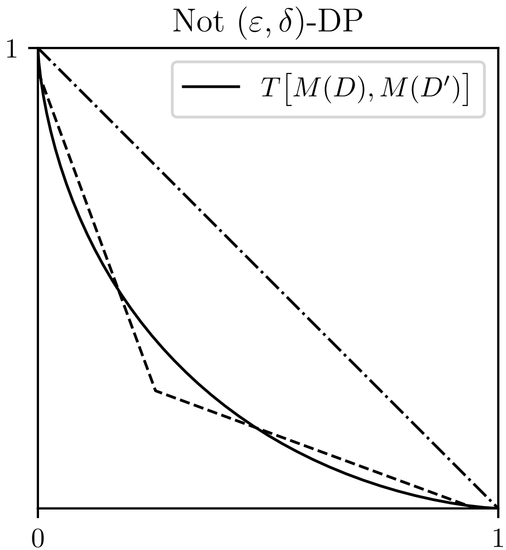
"Functional" perspective

"$M(D) \approx M(D')$"

- "Functional" perspective: A "true" δ for each ε

$$\delta(\varepsilon) = H_{e^\varepsilon}\left(M(D) \| M(D')\right) : \mathbb{R}_{>0} \to [0,1]$$

$$T\left[M(D), M(D')\right] : [0,1] \to [0,1] \qquad \text{([DRS 19])}$$



- Equivalent via primal-dual
- Interpretation: FP vs FN in binary classification D vs D'
- Larger = more privacy
- [WZ 10, KOV 15, DRS 19]: M is (ε, δ)-DP iff $\underbrace{T\left[M(D), M(D')\right]}_{\text{ROC function}} \geqslant f_{\varepsilon,\delta}$

Back to the quiz

$$M(D) = f(D) + X$$

X	Privacy	Var$[X]$
Laplace	$(\varepsilon, 0)$	$2/\varepsilon^2$
Gaussian	(ε, δ)	$> 2/\varepsilon^2$
Truncated Laplace	(ε, δ)	$< 2/\varepsilon^2$

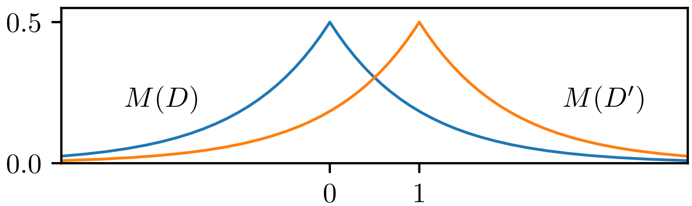
- Understand this by comparing

$$\begin{aligned} f_{\varepsilon,\delta} &= \text{budget of privacy} \\ \text{ROC} &= \text{actual spend by mechanism} \end{aligned}$$

- Which X makes good use of the budget?

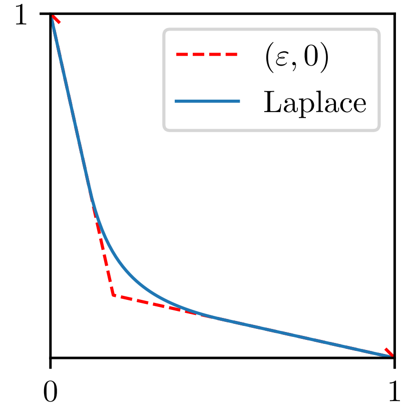
Back to the quiz

$M(D) = f(D) + \text{Lap}(0, \varepsilon^{-1})$

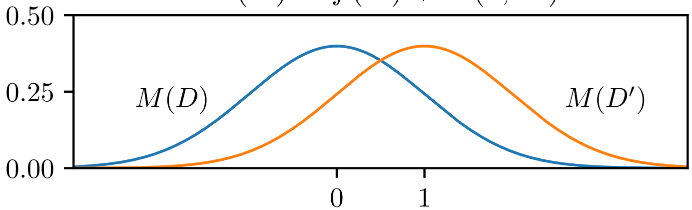


$\text{Var}[X] = \frac{2}{\varepsilon^2}$

$\longleftrightarrow$

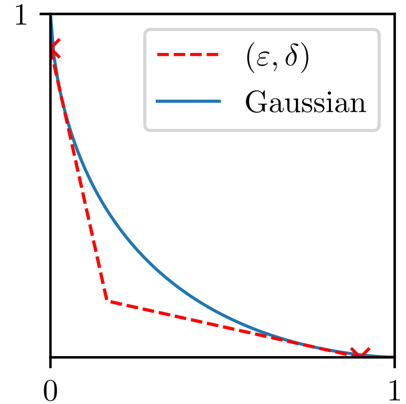


$M(D) = f(D) + N(0, \sigma^2)$

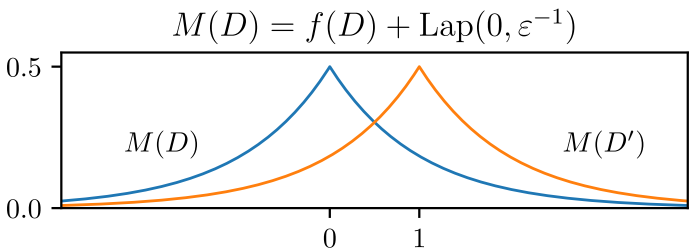


$\text{Var}[X] > \frac{2}{\varepsilon^2}$

$\longleftrightarrow$

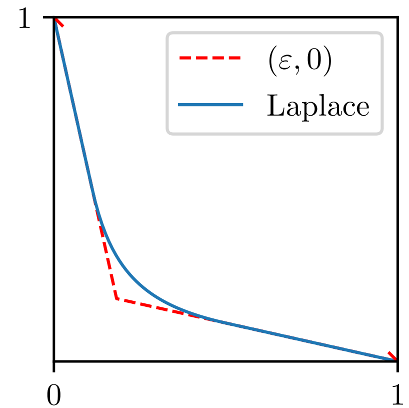


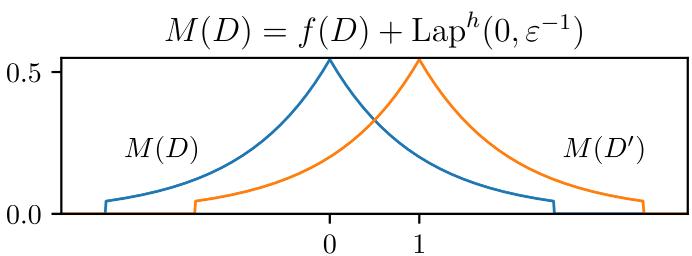
Truncation creates a δ



$\mathrm{Var}[X] = \frac{2}{\varepsilon^2}$

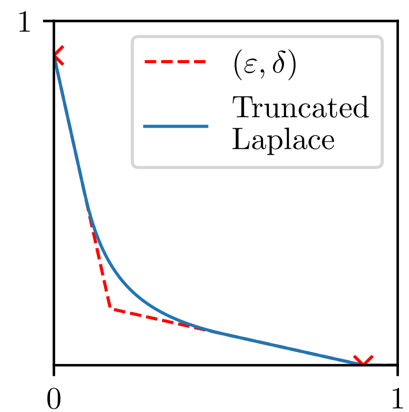
$\longleftrightarrow$



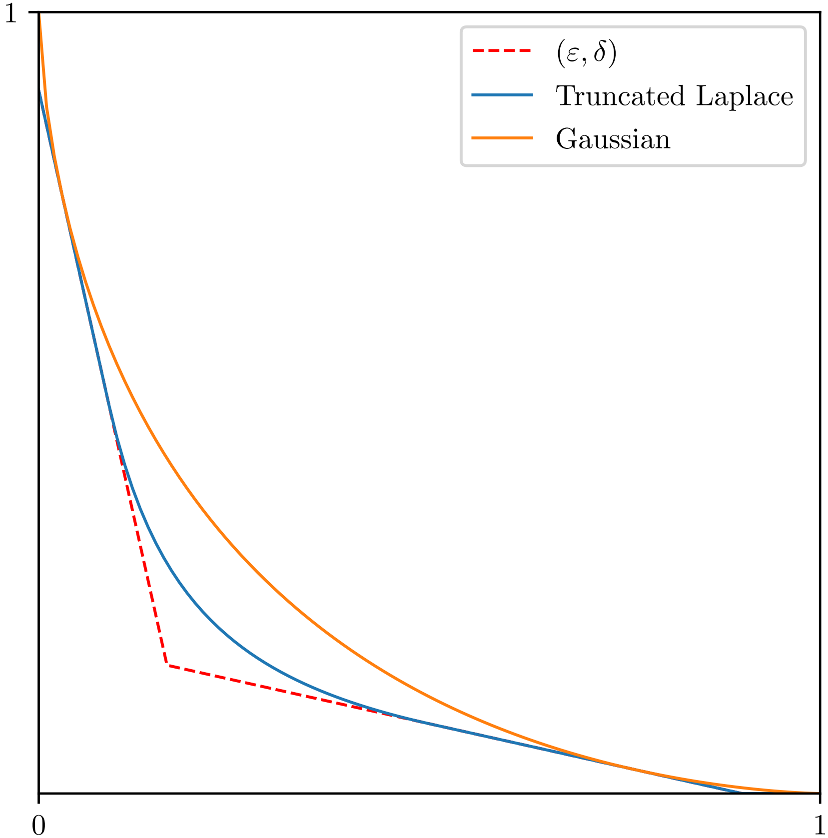


$\mathrm{Var}[X] < \frac{2}{\varepsilon^2}$

$\longleftrightarrow$



Zoom in comparision



- Want to achieve better accuracy?
- Try to make good use of your privacy budget.

What if $d \gg 1$?

Consider noise-addition mechanisms in $\mathbb{R}^d$

$$M(D) = f(D) + X.$$

- Q: How to choose noise X to fit (ε, δ) budget?
- A: No way!

Theorem (Informal CLT, this work)

When $d \gg 1$, for many X,

$$\text{ROC of } M \approx \text{ROC of Gaussian} \neq f_{\varepsilon,\delta}$$

Details of the statement of CLT

- Consider the mechanism $M(D) = f(D) + X$ where X is log-concave with density $\propto \mathrm{e}^{-\varphi(x)}$ where φ is convex.
- WLOG f has ℓ_2 sensitivity 1, i.e. $\|f(D) - f(D')\| \leqslant 1$
- WLOG $f(D) = 0, f(D') = v$ where $\|v\| = 1$, hence

$$T[M(D), M(D')] = T[X, X+v].$$

- "ROC of $M \approx$ ROC of Gaussian"

$$T[X, X+v] \approx T[G, G+v]$$

  where $G = N(0, \Sigma)$ is some Gaussian.
- Normalization:
  - Textbook CLT: $\sum X_i \approx \sum G_i$ if $\mathbb{E}X = \mathbb{E}G$ and $\mathrm{Var}[X] = \mathrm{Var}[G]$.
  - Our CLT:

$$T[X, X+v] \approx T[G, G+v] \quad \text{if} \quad \mathcal{I}_X = \mathcal{I}_G$$

    where $\mathcal{I}_X = \mathbb{E}\nabla\varphi(X)\nabla\varphi(X)^T$ is the $d \times d$ Fisher information matrix.

Details of the statement of CLT, cont'd.

$$T[X, X+v] \approx T[G, G+v] \quad \text{if} \quad \mathcal{I}_X = \mathcal{I}_G \tag{1}$$

- Remember this is a high-dimensional phenomenon
- Unfortunately, high-dimensional DP algorithm can exhibit 1-d behavior
- When $v = (1, 0, \ldots, 0)$, $T[X, X+v] = T[X_1, X_1+1]$ where $X_1 \in \mathbb{R}$.
- Solution: exclude a small fraction of v, i.e. (1) holds w.h.p over $v \sim S^{d-1}$.
- For what X?

$$\text{density} \ \propto \exp(-\|Ux\|_p^\alpha) \quad \text{where } p, \alpha \in [1, +\infty), U \text{ orthogonal}$$

Call this class of densities $\mathcal{F}$.

Statement and proof idea

Theorem (CLT, this work)

For X with densities in $\mathcal{F}$ and $\mathcal{I}_X = I_{d\times d}$, w.p. $\geqslant 1 - o(1)$ over $v \sim S^{d-1}$,

$$\|T[X, X+v] - T[G, G+v]\|_\infty \leqslant o(1),$$

where G is Gaussian such that $\mathcal{I}_G = \mathcal{I}_X = I_{d\times d}$.

Proof idea:

Theorem ([V.N.Sudakov 1978])

If X is an isotropic r.v. in $\mathbb{R}^d$ and satisfies "thin shell" condition, then w.p. $1 - o(1)$ over $v \sim S^{d-1}$, $\langle X, v\rangle \approx N(0,1)$.

- We show that an analog of Sudakov's theorem holds for a nonlinear projection of X that we call "likelihood projection"
- Our CLT follows easily from this "nonlinear Sudakov".
- Conjectured to be extendable to general log-concave distributions with proper regularity.

Some numerical results

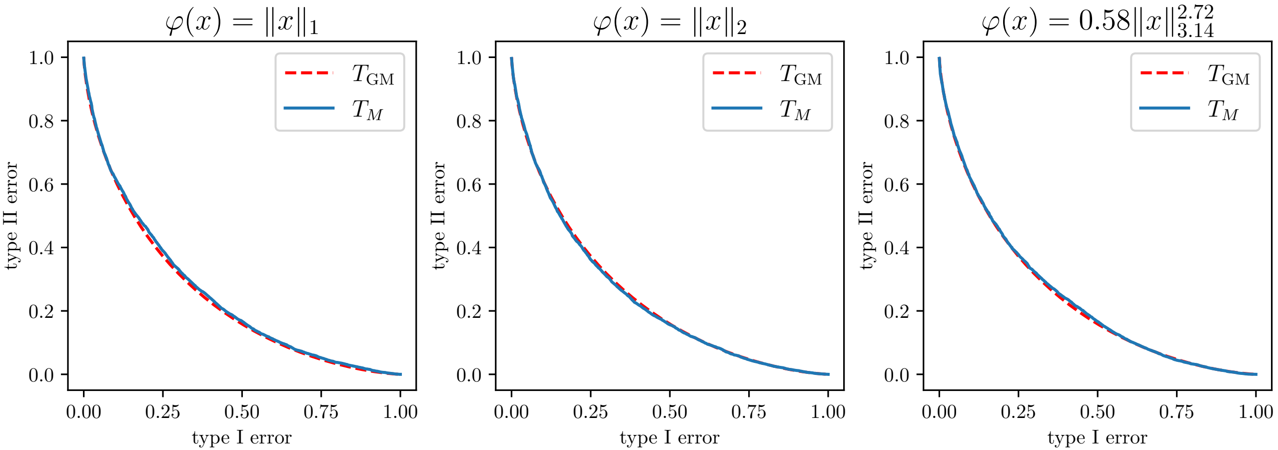
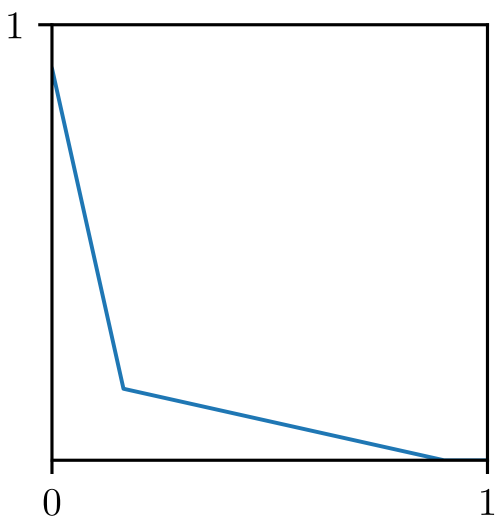


Figure 1: Numerical evaluation of ROC functions for noise addition mechanism $M(D) = f(D) + X$
X has density $\propto \mathrm{e}^{-\varphi(x)}$
Dimension $d = 30$.

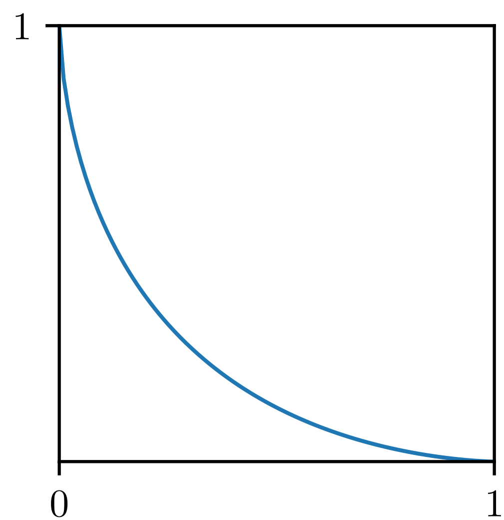
So far...

- For $d = 1$, truncated Laplace fits (ε, δ) budget better and has smaller variance than Gaussian.
- For $d \gg 1$, no hope to fit (ε, δ). Everything works like Gaussian.

$d = 1$



$d \gg 1$



Privacy-Accuracy Trade-off

- (ε, δ) and $d \gg 1$ don't really work together.
- Why not use Gaussian instead of (ε, δ) to measure privacy?
- Exactly what [D-Roth-Su 19] did
- μ-GDP $\Leftrightarrow T[X, X+v] \geqslant \cancel{f_{\varepsilon,\delta}}\, T[N(0,1), N(\mu,1)]$
- By CLT, $T[X, X+v] \approx T[G, G+v]$
- By linear algebra, $T[G, G+v] = T[N(0,1), N(\mu,1)]$ with $\mu^2 = v^T \mathcal{I}_G v$.
- Worst case over $v \in S^{d-1}$: $\mu^2 = \|\mathcal{I}_G\| = \|\mathcal{I}_X\|$.
- That is, adding X is roughly μ-GDP with $\mu^2 = \|\mathcal{I}_X\|$.
- By Cramer–Rao,
$$\mathbb{E}\|X\|_2^2 \cdot \|\mathcal{I}_X\| \geqslant d.$$
- i.e. mean-squared error satisfies
$$\mathrm{err}_M \cdot \mu^2 \geqslant d$$
- $=$ holds for Gaussian mechanism.

Privacy-Accuracy Trade-off, Cont'd

CLT + Cramer–Rao yields

$$\mathrm{err}_M \cdot \mu^2 \geqslant d$$

Compared to previously known lower bounds, e.g. [Steinke-Ullman 17]

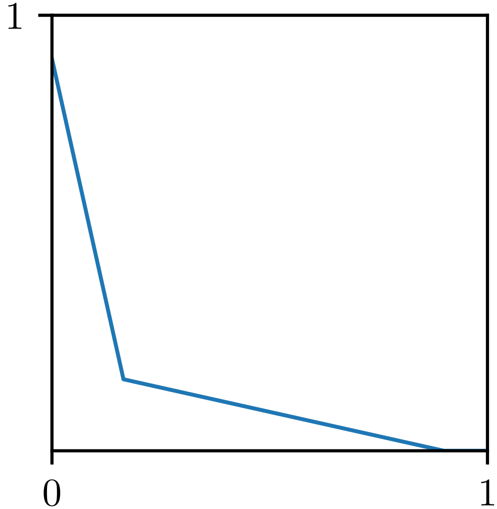
$$\mathrm{err}_M \cdot \frac{\varepsilon^2}{\log \delta^{-1}} = \Omega\left(d\right)$$

- No mysterious constant.
- Equality is precisely achievable by Gaussian mechanism.
- Privacy parameter makes more sense, e.g. avoids "$\delta \to 0$ blowing-up" problem

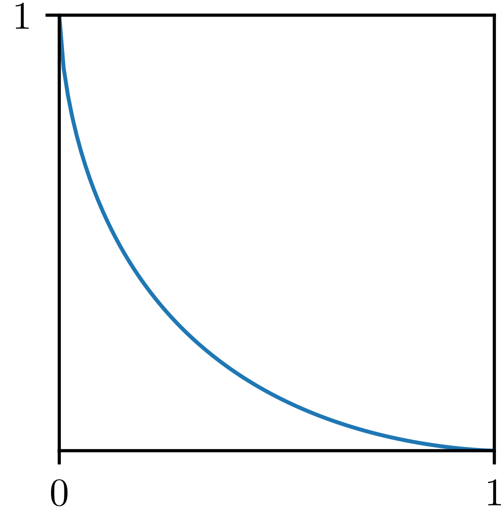
Summary

- CLT: $d = 1$ and $d \gg 1$ are drastically different.
- CLT + Cramer–Rao: $\mathrm{err}_M \cdot \mu^2 \geqslant d$

$d = 1$



$d \gg 1$



Future directions

- Generalize CLT to log-concave distributions?
- To distributions with bounded support?
- Gap between "almost all v" and "all v"?
- Other high-dimensional phenomenon in DP? Constant-sharp lower bound there?
- In particular, what if we consider ℓ_∞ error instead of ℓ_2 error? Constant-sharp optimality of [Dagan-Kur 20]?

Thank you!

- More on [DRS 19]: my blog at **`dongjs.github.io`**

@JinshuoD
@zlj11112222
@weijie444