Statistical Efficiency in Local Differential Privacy

Lukas Steinberger
(University of Vienna)

joint with Nikita Kalinin, ISTA

Institute for Statistics and Mathematics, WU
January 8, 2025

research supported by Austrian Science Fund (FWF): I 5484-N

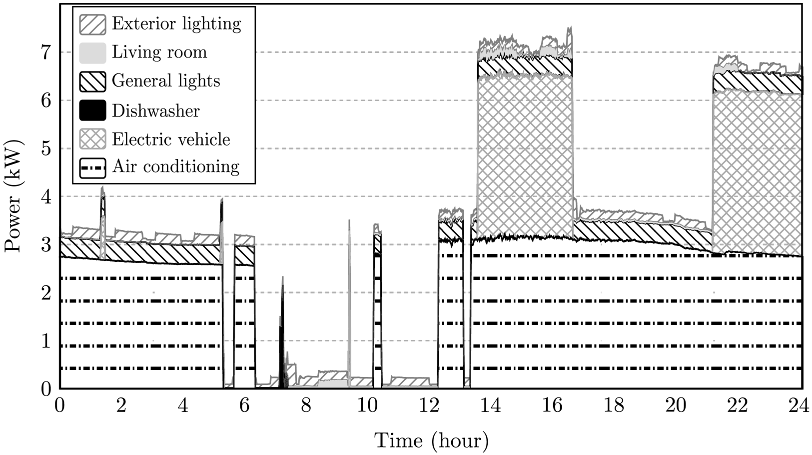
universität wien

Issues of data privacy protection

This is an old problem with increasing relevance in the modern era of big data. For instance:

- official statistics: statistical disclosure control
- large scale medical research
- smart phone user data
- social media data
- social or psychological surveys: *evasive answer bias*
- IoT
- etc.

Example: Data from smart meter



(from Giaconi et al., 2018)

Definition of Differential Privacy

Dwork et al. (2006) proposed the following.

individuals		database		public data
Alice	$\xrightarrow{x_1}$			
Bob	$\xrightarrow{x_2}$	x	$\xrightarrow[Q]{\text{randomization}}$	Z
$\vdots$				
Zoey	$\xrightarrow{x_n}$			

Distribution of Z should not depend too much on any individual contribution x_i.

Differential privacy

Dwork et al. (2006) proposed the following.

- For a given original data set $X = (X_1, \ldots, X_n)$ in $\mathcal{X}^n$, randomly generate sanitized data Z in $\mathcal{Z}$, with conditional distribution

$$Q(A|x) \;=\; P(Z \in A|X = x).$$

- The conditional distribution (Markov kernel) $Q \in \mathcal{M}(\mathcal{X}^n \to \mathcal{Z})$ is called a ***privacy mechanism*** or a ***channel***.
- The distribution of the sanitized data Z is given by

$$QP := \int_{\mathcal{X}^n} Q(\cdot|x)\, dP(x).$$

Differential Privacy

For $x, x' \in \mathcal{X}^n$, consider the Hamming distance

$$d_0(x, x') = \#\{i : x_i \neq x'_i\}.$$

Definition (Dwork et al., 2006)

Fix a privacy parameter $\varepsilon \in (0, \infty)$. The Markov kernel $Q \in \mathcal{M}(\mathcal{X}^n \to \mathcal{Z})$ is called ε**-differentially private** if for all $x, x' \in \mathcal{X}^n$ with $d_0(x, x') \leq 1$, we have

$$Q(A|x) \leq e^\varepsilon Q(A|x'), \quad \forall A \in \mathcal{G},$$

ε-Differential privacy

$$\forall A, \forall x, x' : d_0(x, x') \leq 1 :$$
$$e^{-\varepsilon} \ \leq \ \frac{Q(A|x)}{Q(A|x')} \ \leq \ e^{\varepsilon}$$

- **Idea:** The conditional distribution of Z given $X = x$ does not depend too much on the data of the i-th individual in the database, thereby protecting its privacy.
- The smaller $\varepsilon \in (0, \infty)$, the stronger is the privacy protection.

ε-Differential privacy

$$\forall A, \forall x, x' : d_0(x, x') \leq 1 :$$
$$e^{-\varepsilon} \;\leq\; \frac{Q(A|x)}{Q(A|x')} \;\leq\; e^{\varepsilon}$$

- **Idea:** The conditional distribution of Z given $X = x$ does not depend too much on the data of the i-th individual in the database, thereby protecting its privacy.
- The smaller $\varepsilon \in (0, \infty)$, the stronger is the privacy protection.

Example - Laplace noise for mean estimation

- Let $X_1, \ldots, X_n \overset{iid}{\sim} P \in \mathcal{P}(\mathcal{X})$ with $\mathcal{X} = [-M, M]$.
- We want to release an estimate of $\theta := \mathbb{E}[X_1]$ while respecting ε-DP.
- Publish $Z = \bar{X}_n + Lap(n\varepsilon/(2M))$, where

$$f_{Lap(\gamma)}(z) = \frac{\gamma}{2} \exp(-\gamma|z|).$$

Example - Laplace noise for mean estimation

- Let $X_1, \ldots, X_n \overset{iid}{\sim} P \in \mathcal{P}(\mathcal{X})$ with $\mathcal{X} = [-M, M]$.
- We want to release an estimate of $\theta := \mathbb{E}[X_1]$ while respecting ε-DP.
- Publish $Z = \bar{X}_n + Lap(n\varepsilon/(2M))$, where

$$f_{Lap(\gamma)}(z) = \frac{\gamma}{2} \exp(-\gamma|z|).$$

$$\begin{aligned}\frac{q(z|x)}{q(z|x')} &= \exp\left(-\frac{n\varepsilon}{2M}\left[|z-\bar{x}_n| - |z-\bar{x}'_n|\right]\right) \\ &\leq \exp\left(\frac{n\varepsilon}{2M}|\bar{x}_n - \bar{x}'_n|\right) \\ &= \exp\left(\frac{n\varepsilon}{2M}\left|\frac{x_{i_0} - x'_{i_0}}{n}\right|\right) \leq e^{\varepsilon}.\end{aligned}$$

Example - Laplace noise for mean estimation

- Let $X_1, \dots, X_n \overset{iid}{\sim} P \in \mathcal{P}(\mathcal{X})$ with $\mathcal{X} = [-M, M]$.
- We want to release an estimate of $\theta := \mathbb{E}[X_1]$ while respecting ε-DP.
- Publish $Z = \bar{X}_n + Lap(n\varepsilon/(2M))$, where

$$f_{Lap(\gamma)}(z) = \frac{\gamma}{2} \exp(-\gamma|z|).$$

- This requires a trusted third party who collects $X_1, \dots, X_n$, computes $\bar{X}_n$ and adds the Laplace noise.
  $\Rightarrow$ *local* differential privacy

LOCAL DIFFERENTIAL PRIVACY

We say that an ε-DP channel $Q \in \mathcal{M}(\mathcal{X}^n \to \mathcal{Z}^n)$ provides **local privacy**, if individual i can generate its sanitized data Z_i on its 'local machine', without ever giving away its original data X_i.

- No trusted third party needed

Local privacy - non-interactive case

Definition

We say that a channel $Q \in \mathcal{M}(\mathcal{X}^n \to \mathcal{Z}^n)$ is **non-interactive (NI)**, if there exist channels $Q_i \in \mathcal{M}(\mathcal{X} \to \mathcal{Z})$, such that

$$Q(dz|x) = \bigotimes_{i=1}^{n} Q_i(dz_i|x_i).$$

$$\begin{array}{ccccc} X_1 & & X_2 & & X_n \\ Q_1 \big\downarrow & & Q_2 \big\downarrow & \cdots\cdots & Q_n \big\downarrow \\ Z_1 & & Z_2 & & Z_n \end{array}$$

Q is ε-DP $\iff Q_i(A_i|x_i) \le e^{\varepsilon} Q_i(A_i|x_i'), \ \forall i, A_i, x_i, x_i'$

Example: Central vs. local mean estimation

- Let $X_1, \dots, X_n \overset{iid}{\sim} P \in \mathcal{P}(\mathcal{X})$ with $\mathcal{X} = [-M, M]$.
- Estimate $\theta := \mathbb{E}[X_1]$ while respecting ε-DP.

With a central data curator: $\hat{\theta}_n = \bar{X}_n + Lap(n\varepsilon/(2M))$

- $\mathbb{E}[\hat{\theta}_n] = \theta$
- $\text{Var}[\hat{\theta}_n] = \frac{Var[X_1]}{n} + \frac{8M^2}{n^2\varepsilon^2}$

With local privacy: $Z_i = X_i + Lap(\varepsilon/(2M))$, $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$

- $\mathbb{E}[\hat{\theta}_n] = \theta$
- $\text{Var}[\hat{\theta}_n] = \frac{Var[X_1]}{n} + \frac{8M^2}{n\varepsilon^2}$

Additional noise is non-negligible for $n \to \infty$.

EXAMPLE: LOCALLY PRIVATE MEAN ESTIMATION

- Let $X_1, \ldots, X_n \overset{iid}{\sim} P \in \mathcal{P}(\mathcal{X})$ with $\mathcal{X} = [-M, M]$.
- Estimate $\theta := \mathbb{E}[X_1]$ while respecting ε-DP.

With local privacy: $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n Z_i$

- $Z_i = X_i + Lap(\varepsilon/(2M))$
  - $\mathbb{E}[\hat{\theta}_n] = \theta$
  - $\mathrm{Var}[\hat{\theta}_n] = \frac{Var[X_1]}{n} + \frac{8M^2}{n\varepsilon^2} = \frac{1}{n}\left(\sigma^2 + \frac{8M^2}{\varepsilon^2}\right)$
- $Z_i = \pm z_0$, w.p. $\frac{1}{2}\left(1 \pm \frac{X_i}{z_0}\right)$, where $z_0 := M\frac{e^\varepsilon+1}{e^\varepsilon-1}$.
  - $\mathbb{E}[\hat{\theta}_n] = \mathbb{E}[\mathbb{E}[Z_1|X_1]] = \mathbb{E}[X_1] = \theta$
  - $\mathrm{Var}[\hat{\theta}_n] = \frac{1}{n}\left(z_0^2 - \theta^2\right)$

Most of the literature deals with minimax rates of convergence.
Can't distinguish mechanisms!

Asymptotic efficiency

- **classical parametric estimation problem:** (Hájek, 1970; Le Cam, 1960)

Given data $X_1, \dots, X_n \overset{iid}{\sim} P_\theta$, $\theta \in \Theta \subseteq \mathbb{R}^p$, and a regular estimator $\hat{\theta}_n : \mathcal{X}^n \to \Theta$ of θ with

$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{P_\theta^n}{\rightsquigarrow} D_\theta,$$

then $\mathrm{Cov}(D_\theta) \succcurlyeq I_\theta^{-1}$ and the MLE achieves this minimal asymptotic covariance matrix.

Asymptotic efficiency

Differentiability in Quadratic Mean (DQM)

The model $(P_\theta)_{\theta\in\Theta}$ with $\Theta \subseteq \mathbb{R}^p$ is called *differentiable in quadratic mean* at the point $\theta \in \Theta$, if θ is an interior point of Θ and there exists a (σ-finite) dominating measure μ such that the corresponding μ-densities $p_\theta = \frac{dP_\theta}{d\mu}$ satisfy

$$\int_{\mathcal{X}} \left(\sqrt{p_{\theta+h}(x)} - \sqrt{p_\theta(x)} - \frac{1}{2} h^T s_\theta(x) \sqrt{p_\theta(x)} \right)^2 d\mu(x) \;=\; o(\|h\|^2)$$

as $h \to 0$, for some measurable vector valued function $s_\theta : \mathcal{X} \to \mathbb{R}^p$. The function s_θ is called the *score function* at θ.

Define $\dot{p}_\theta := s_\theta p_\theta$.

Asymptotic efficiency

Regular Estimator

An estimator $\hat{\theta}_n : \mathcal{X}^n \to \Theta$ in a DQM model is called regular at $\theta \in \Theta$ if

$$\sqrt{n}\Big(\hat{\theta}_n - (\theta + h/\sqrt{n})\Big) \overset{P^n_{\theta+h/\sqrt{n}}}{\rightsquigarrow} D_\theta, \quad \forall h \in \mathbb{R}^p,$$

where the limiting distribution D_θ does not depend on h.

Asymptotic efficiency with LDP

$$\begin{array}{cccc} X_1 & X_2 & & X_n \\ Q_0 \big\downarrow & Q_0 \big\downarrow & \cdots\cdots & Q_0 \big\downarrow \\ Z_1 & Z_2 & & Z_n \end{array}$$

- **private estimation problem:**

Given sanitized data $Z_1, \ldots, Z_n \overset{iid}{\sim} Q_0 P_\theta$, $\theta \in \Theta \subseteq \mathbb{R}^p$ and a regular estimator $\hat{\theta}_n : \mathcal{Z}^n \to \Theta$ of θ with

$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{[Q_0 P_\theta]^n}{\rightsquigarrow} D_\theta,$$

then $\mathrm{Cov}(D_\theta) \succcurlyeq I_\theta(Q_0)^{-1}$ and the MLE achieves this asymptotic covariance matrix.

$$\sup_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q) \quad (\leq I_\theta) \qquad \Theta \subseteq \mathbb{R}$$

$$\mathcal{Q}_\varepsilon(\mathcal{X}) = \bigcup_{(\mathcal{Z},\mathcal{G})} \left\{ Q \in \mathcal{M}(\mathcal{X} \to \mathcal{Z}) \middle| Q(A|x) \leq e^\varepsilon Q(A|x'),\ \forall A, x, x' \right\}$$

- infinite dimensional domain $\mathcal{Q}_\varepsilon(\mathcal{X})$
- maximizing a convex function on a convex set (local optima!)
- maximizer depends on θ?!

If $|\mathcal{X}| = k \in \mathbb{N}$

$$\sup_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q) = \sup_{Q \in \mathcal{Q}_{\varepsilon,k}} I_\theta(Q)$$

$$\mathcal{Q}_{\varepsilon,k} = \bigcup_{\mathcal{Z}:|\mathcal{Z}|=k} \left\{ Q \in \mathcal{M}(\mathcal{X} \to \mathcal{Z}) \middle| Q(A|x) \le e^\varepsilon Q(A|x') \; \forall A, x, x' \right\}$$

$$\triangleq \left\{ Q \in [0,1]^{k \times k} \middle| \sum_{i=1}^k Q_{ij} = 1, Q_{ij} \le e^\varepsilon Q_{ij'} \; \forall i, j, j' \right\}$$

- Notice that we went from all possible measurable spaces $(\mathcal{Z}, \mathcal{G})$ to $\mathcal{Z} = \{1, \dots, k\}$.
- Kairouz et al. (2016) provide an equivalent LP with time complexity $O(2^k)$.

If $|\mathcal{X}| = k \in \mathbb{N}$

$$\sup_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q) = \sup_{Q \in \mathcal{Q}_{\varepsilon,k}} I_\theta(Q)$$

$$\mathcal{Q}_{\varepsilon,k} = \bigcup_{\mathcal{Z}:|\mathcal{Z}|=k} \left\{ Q \in \mathcal{M}(\mathcal{X} \to \mathcal{Z}) \middle| Q(A|x) \leq e^\varepsilon Q(A|x') \; \forall A, x, x' \right\}$$

$$\triangleq \left\{ Q \in [0,1]^{k \times k} \middle| \sum_{i=1}^{k} Q_{ij} = 1, Q_{ij} \leq e^\varepsilon Q_{ij'} \; \forall i, j, j' \right\}$$

- Notice that we went from all possible measurable spaces $(\mathcal{Z}, \mathcal{G})$ to $\mathcal{Z} = \{1, \ldots, k\}$.
- Kairouz et al. (2016) provide an equivalent LP with time complexity $O(2^k)$.

If $|\mathcal{X}| = k \in \mathbb{N}$

$$\sup_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q) = \sup_{Q \in \mathcal{Q}_{\varepsilon,k}} I_\theta(Q)$$

$$\begin{aligned}\mathcal{Q}_{\varepsilon,k} &= \bigcup_{\mathcal{Z}:|\mathcal{Z}|=k} \left\{ Q \in \mathcal{M}(\mathcal{X} \to \mathcal{Z}) \middle| Q(A|x) \le e^\varepsilon Q(A|x') \; \forall A, x, x' \right\} \\ &\triangleq \left\{ Q \in [0,1]^{k \times k} \middle| \sum_{i=1}^k Q_{ij} = 1, Q_{ij} \le e^\varepsilon Q_{ij'} \; \forall i, j, j' \right\}\end{aligned}$$

- Notice that we went from all possible measurable spaces $(\mathcal{Z}, \mathcal{G})$ to $\mathcal{Z} = \{1, \ldots, k\}$.
- Kairouz et al. (2016) provide an equivalent LP with time complexity $O(2^k)$.

$\max_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$

- Bernoulli(θ):
  $p_\theta(x) = \theta^x(1-\theta)^{1-x}$, $\theta \in (0,1)$, $x \in \mathcal{X} = \{0,1\}$

$$Q^* = \frac{1}{1+e^\varepsilon}\begin{pmatrix} e^\varepsilon & 1 \\ 1 & e^\varepsilon \end{pmatrix}$$

See Kairouz et al. (2016)

$\max_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$

- Binomial$(2, \theta)$:
  $p_\theta(x) = \binom{2}{x}\theta^x(1-\theta)^{2-x}$, $\theta \in (0,1)$, $x \in \mathcal{X} = \{0,1,2\}$

$$Q^* \quad = \quad ?$$

$\max_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$

- Binomial$(2, \theta)$:
  $p_\theta(x) = \binom{2}{x} \theta^x (1-\theta)^{2-x}$, $\theta \in (0,1)$, $x \in \mathcal{X} = \{0, 1, 2\}$

$$Q^* = \frac{1}{2 + e^\varepsilon} \begin{pmatrix} e^\varepsilon & 1 & 1 \\ 1 & e^\varepsilon & 1 \\ 1 & 1 & e^\varepsilon \end{pmatrix} \qquad ?$$

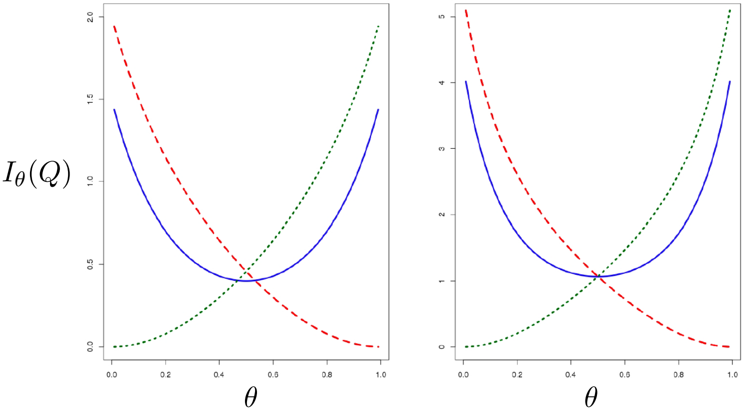
$\max_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$

- Binomial$(2, \theta)$:
  $p_\theta(x) = \binom{2}{x}\theta^x(1-\theta)^{2-x}$, $\theta \in (0,1)$, $x \in \mathcal{X} = \{0, 1, 2\}$

$$Q^*_\theta = \begin{cases} \frac{1}{1+e^\varepsilon}\begin{pmatrix} e^\varepsilon & 1 & 1 \\ 1 & e^\varepsilon & e^\varepsilon \\ 0 & 0 & 0 \end{pmatrix}, & 0 < \theta \le \frac{1}{2} - c_\varepsilon \\ \frac{1}{2+e^\varepsilon}\begin{pmatrix} e^\varepsilon & 1 & 1 \\ 1 & e^\varepsilon & 1 \\ 1 & 1 & e^\varepsilon \end{pmatrix}, & \frac{1}{2} - c_\varepsilon < \theta < \frac{1}{2} + c_\varepsilon \\ \frac{1}{1+e^\varepsilon}\begin{pmatrix} e^\varepsilon & e^\varepsilon & 1 \\ 1 & 1 & e^\varepsilon \\ 0 & 0 & 0 \end{pmatrix}, & \frac{1}{2} + c_\varepsilon \le \theta < 1 \end{cases}$$

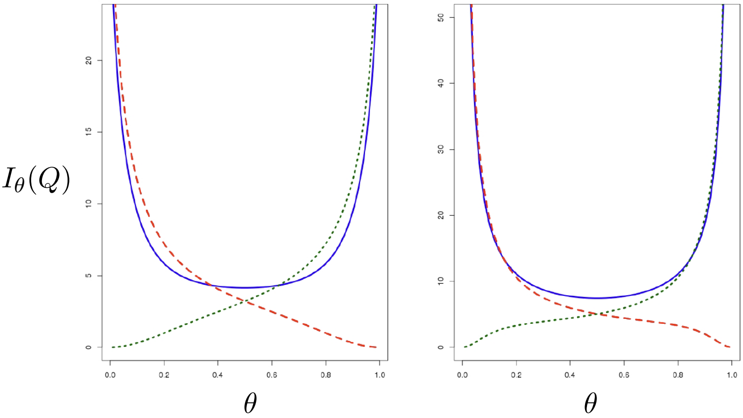
See Hucke (2019)

$\max_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$



(a) $\varepsilon = \log(2)$ (b) $\varepsilon = \log(3)$

$\max_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$



(c) $\varepsilon = \log(10)$ (d) $\varepsilon = \log(100)$

$\max_{Q\in\mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$

Fortunately we have continuity at $\theta \in \Theta$:

$$I_\theta(Q^*_{\theta_0}) \xrightarrow[\theta_0\to\theta]{} \max_{Q\in\mathcal{Q}_\varepsilon} I_\theta(Q).$$

Thus, we only need to solve

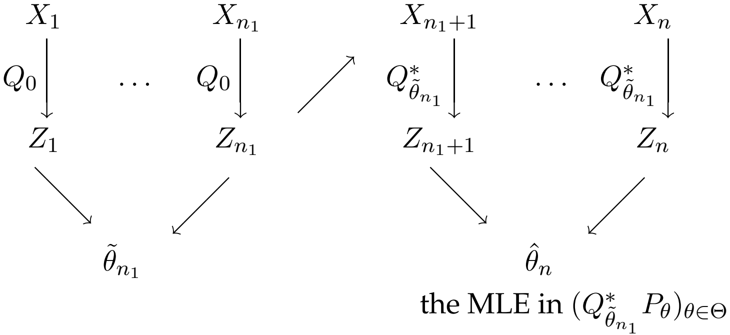
$$\max_{Q\in\mathcal{Q}_\varepsilon(\mathcal{X})} I_{\tilde{\theta}_{n_1}}(Q),$$

for a consistent estimator $\tilde{\theta}_{n_1}$.

In general, for regular parametric models, we have

$$\sup_{Q\in\mathcal{Q}_\varepsilon(\mathcal{X})} |I_\theta(Q) - I_{\theta'}(Q)| \xrightarrow[\theta\to\theta']{} 0.$$

A two-step procedure



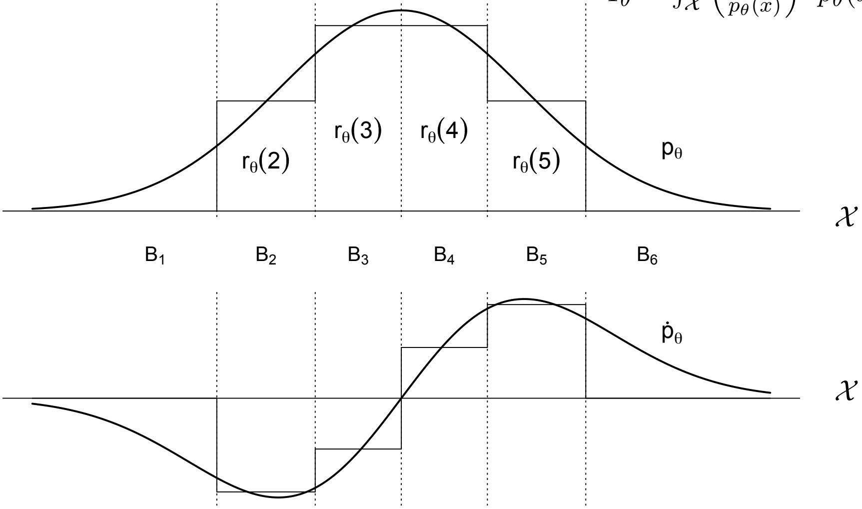
$$\tilde{\theta}_{n_1} \xrightarrow[n_1 \to \infty]{} \theta$$

$$I_\theta(Q^*_{\tilde{\theta}_{n_1}}) \xrightarrow[n_1 \to \infty]{} I_\theta(Q^*_\theta) = \max_{Q \in \mathcal{Q}_\varepsilon} I_\theta(Q)$$

But notice that for efficiency of $\hat{\theta}_n$ we need $\frac{n-n_1}{n} \to 1$.

General sample space $\mathcal{X}$

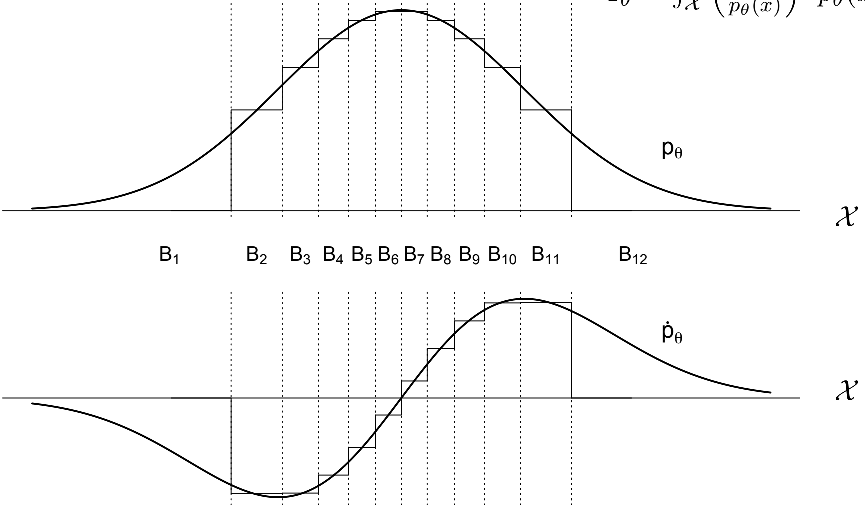
$$I_\theta = \int_{\mathcal{X}} \left(\frac{\dot p_\theta(x)}{p_\theta(x)}\right)^2 p_\theta(x)dx$$



$r_\theta(j) := P_\theta(B_j(\theta_0))$ is the pmf of a regular model on finite sample space $\bar{\mathcal{X}} = \{1, \ldots, 6\}$ $k = 6$

General sample space $\mathcal{X}$

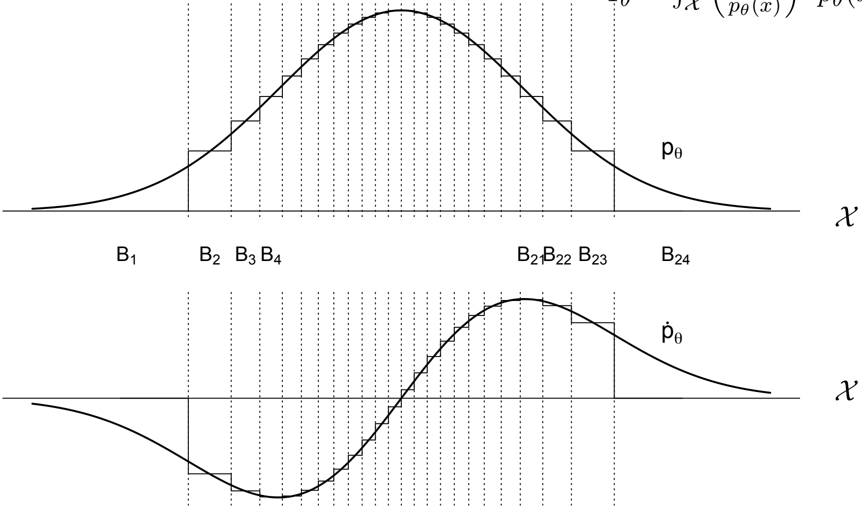
$$I_\theta = \int_{\mathcal{X}} \left(\frac{\dot{p}_\theta(x)}{p_\theta(x)}\right)^2 p_\theta(x)dx$$



$r_\theta(j) := P_\theta(B_j(\theta_0))$ is the pmf of a regular model on finite sample space $\bar{\mathcal{X}} = \{1, \ldots, 12\}$ $k = 12$

General sample space $\mathcal{X}$

$$I_\theta = \int_{\mathcal{X}} \left(\frac{\dot{p}_\theta(x)}{p_\theta(x)}\right)^2 p_\theta(x)dx$$



$r_\theta(j) := P_\theta(B_j(\theta_0))$ is the pmf of a regular model on finite sample space $\bar{\mathcal{X}} = \{1, \ldots, 24\}$ $k = 24$

APPROXIMATION BY DISCRETE MODELS

$$T_{k,\theta} : \mathcal{X} \to \{1, \dots, k\}, \quad T_{k,\theta}(x) = j \iff x \in B_j(\theta)$$

$$Y_i = T_{k,\theta}(X_i), \quad Z_i \sim Q^*_\theta(dz|Y_i)$$

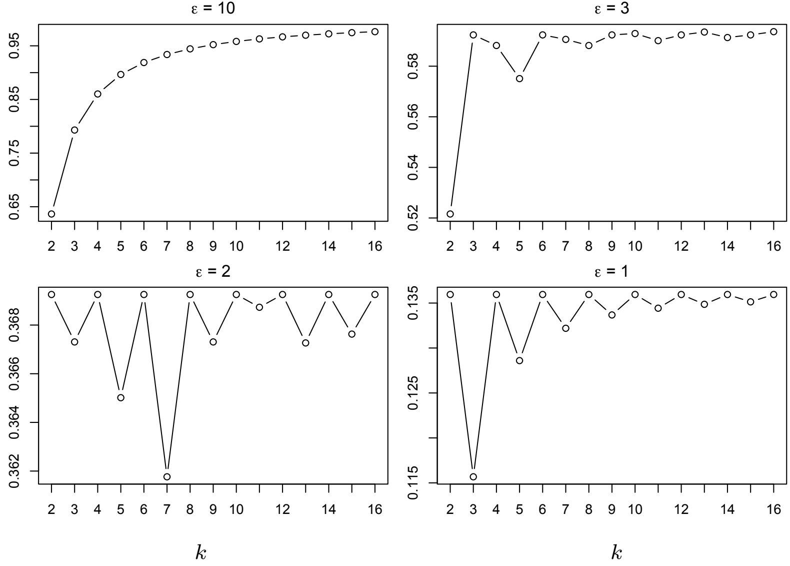
$$\max_{Q \in \mathcal{Q}_{\varepsilon,k}} I_\theta(QT_{k,\theta}) \xrightarrow[k \to \infty]{} \sup_{Q \in \mathcal{Q}_\varepsilon(\mathcal{X})} I_\theta(Q)$$

- Use with $\theta = \tilde{\theta}_{n_1}$.
- Need to solve the LP of Kairouz et al. (2016) for large k.
- Efficient numerical procedures are needed.

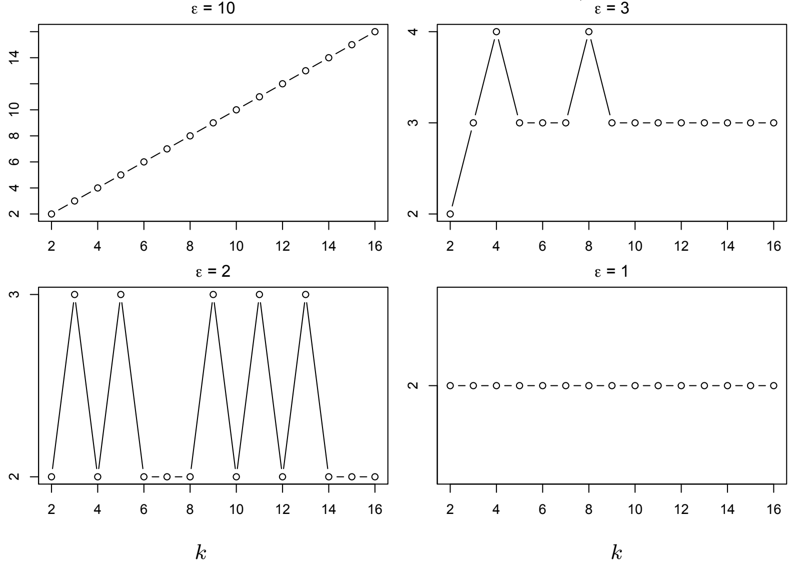
Example: Gaussian Location Model

$$P_\theta = N(\theta, 1),\ \theta \in \mathbb{R}$$

$P_\theta = N(\theta, 1), \theta \in \mathbb{R}$, $\max_{Q \in \mathcal{Q}_{\varepsilon,k}} I_0(QT_{k,0}) \in \mathbb{R}$



$$P_\theta = N(\theta, 1), \theta \in \mathbb{R},\ Q^* \in \mathrm{argmax}_{Q \in \mathcal{Q}_{\varepsilon,k}} I_0(QT_{k,0}) \in \mathbb{R}^{k \times k}$$



Theorem (Kalinin and S. (2024))

In the Gaussian location model with unit variance, if $\varepsilon \leq 1.04$ the sign-mechanism Q_θ^{sgn} that generates

$$Z_i = \begin{cases} \text{sgn}(X_i - \theta), & \text{with probability } \frac{e^\varepsilon}{1+e^\varepsilon} \\ -\text{sgn}(X_i - \theta), & \text{with probability } \frac{1}{1+e^\varepsilon}, \end{cases}$$

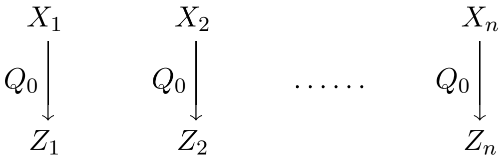
satisfies

$$I_\theta(Q) \leq I_\theta(Q_\theta^{sgn}) = \frac{2}{\pi}\left(\frac{e^\varepsilon - 1}{e^\varepsilon + 1}\right)^2,$$

for **all** ϵ-DP mechanisms Q and all $\theta \in \mathbb{R}$.

cf. Duchi and Rogers (2019)

ASYMPTOTIC EFFICIENCY WITH **NON-INTERACTIVE** LDP

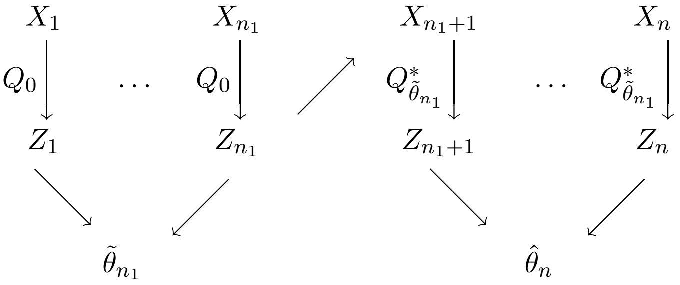


Given sanitized data $Z_1, \ldots, Z_n \overset{iid}{\sim} Q_0 P_\theta$, $\theta \in \Theta \subseteq \mathbb{R}^p$ and a regular estimator $\hat{\theta}_n : \mathcal{Z}^n \to \Theta$ of θ with

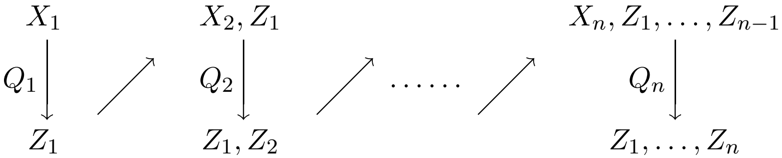
$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{[Q_0 P_\theta]^n}{\rightsquigarrow} D_\theta,$$

then $\mathrm{Cov}(D_\theta) \succcurlyeq I_\theta(Q_0)^{-1}$ and the MLE achieves this asymptotic covariance matrix.

A TWO-STEP PROCEDURE **(INTERACTIVE)**



Asymptotic efficiency with interaction



$$Q^{(n)}(dz|x) = Q_n(dz_n|x_n, z_1, \ldots, z_{n-1}) \cdots Q_2(dz_2|x_2, z_1)Q_1(dz_1|x_1)$$

Given sanitized data $(Z_1, \ldots, Z_n) \sim Q^{(n)} P_\theta^n$, $\theta \in \Theta \subseteq \mathbb{R}$ and a regular estimator $\hat{\theta}_n : \mathcal{Z}^n \to \Theta$ of θ with

$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{[Q_0 P_\theta]^n}{\rightsquigarrow} D_\theta,$$

then $\mathrm{Var}_\theta(D_\theta) \geq [\sup_{Q \in \mathcal{Q}_\varepsilon} I_\theta(Q)]^{-1}$ and the two-step procedure achieves this asymptotic variance.

Asymptotic efficiency with interaction

Given sanitized data $(Z_1, \ldots, Z_n) \sim Q^{(n)} P_\theta^n$, $\theta \in \Theta \subseteq \mathbb{R}$ and a regular estimator $\hat{\theta}_n : \mathcal{Z}^n \to \Theta$ of θ with

$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{[Q_0 P_\theta]^n}{\rightsquigarrow} D_\theta,$$

then $\mathrm{Var}_\theta(D_\theta) \geq [\sup_{Q \in \mathcal{Q}_\varepsilon} I_\theta(Q)]^{-1}$ and the two-step procedure achieves this asymptotic variance.

- We proof LAMN of $(\mathcal{Z}^n, \mathcal{G}^n, (Q^{(n)} P_\theta^n)_{\theta \in \Theta})$, $n \in \mathbb{N}$, along subsequences.
- We need DQM, and separability of the σ-Algebras of $(\mathcal{X}, \mathcal{F})$ and $(\mathcal{Z}, \mathcal{G})$.
- For efficiency of the two-step MLE we use more classical differentiability conditions on the density $\theta \mapsto p_\theta(x)$.

Summary

- We develop a theory of asymptotic efficiency for (sequentially) interactive local differential privacy.
- We provide a numerical procedure that identifies a nearly optimal privacy mechanism Q_θ^* up to arbitrary precision.
- We propose a sequentially interactive private estimation procedure that achieves the asymptotically minimal variance.

Open:

- Numerically efficient algorithms.
- For $p > 1$, consider $\inf_Q \ell(I_\theta(Q)^{-1})$ for an $\ell : \mathbb{R}^{p\times p} \to \mathbb{R}$.
- Nuisance parameters (finite- and infinite-dimensional)

Thank you!

Duchi, J. and Rogers, R. (2019). Lower bounds for locally private estimation via communication complexity. *PMLR*, 99:1161–1191.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In Halevi, S. and Rabin, T., editors, *Theory of Cryptography*, Lecture Notes in Computer Science, pages 265–284. Springer.

Giaconi, G., Gunduz, D., and Poor, H. V. (2018). Privacy-aware smart metering: Progress and challenges. *IEEE Signal Processing Magazine*, 35(6):59–78.

Hájek, J. (1970). A characterization of limiting distributions of regular estimates. *Z. Wahrsch. verw. Gebiete*, 14(4):323–330.

Hucke, U. (2019). Local differential privacy and estimation in the binomial model. Master's thesis, University of Freiburg.

Kairouz, P., Oh, S., and Viswanath, P. (2016). Extremal mechanisms for local differential privacy. *J. Mach. Learn. Res.*, 17(1):492–542.

Kalinin, N. and Steinberger, L. (2024). Efficient estimation of a gaussian mean with local differential privacy. *arXiv:2402.04840*.

Le Cam, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.*, 3:37–98.

Steinberger, L. (2024). Efficiency in local differential privacy. *Ann. Statist.*, 52(5):2139–2166.

Regularity conditions

- **Consistent quantizers** $T_{k,\theta} : \mathcal{X} \to \{1, \dots, k\}$ **exist** if $\mathcal{P} = (P_\theta)_{\theta \in \Theta}$ is DQM with jointly measurable $p_\theta(x)$ and $s_\theta(x)$, $\mathcal{X} \subseteq \mathbb{R}^d$ and the dominating measure μ is finite on compact sets.
- For **uniform continuity of Fisher-Information** we need DQM of the model with jointly measurable $p_\theta(x)$ and $s_\theta(x)$ and continuity of $\theta \mapsto s_\theta \sqrt{p_\theta} : \Theta \to L_2(\mu, \| \cdot \|_2)$.