Five Building Blocks of Differential Privacy

James Bailie,[*] Ruobin Gong[†] and Xiao-Li Meng[‡]

[*]Chalmers University, [†]Rutgers University, [‡]Harvard University

Adelaide Data Privacy Workshop

26 November 2025

These slides are available online:

jameshbailie.github.io/talks/

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications)

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change:

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.
- These choice are the *building blocks* of a DP specification:

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.
- These choice are the *building blocks* of a DP specification:
  1. The protection *domain* ($\mathcal{X}$)

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.
- These choice are the *building blocks* of a DP specification:
    1. The protection *domain* ($\mathcal{X}$)
    2. The *scope* of protection ($\mathscr{D}$)

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.
- These choice are the *building blocks* of a DP specification:
  1. The protection *domain* ($\mathcal{X}$)
  2. The *scope* of protection ($\mathscr{D}$)
  3. The protection *unit* ($d_{\mathcal{X}}$)

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy ***is*** to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.
- These choice are the *building blocks* of a DP specification:
    1. The protection *domain* ($\mathcal{X}$)
    2. The *scope* of protection ($\mathscr{D}$)
    3. The protection *unit* ($d_{\mathcal{X}}$)
    4. The *standard* of protection (d_{Pr})

Differential Privacy (DP) (Dwork, McSherry, Nissim, and Smith, 2006)

A large family of technical standards (i.e. mathematical specifications) that conceptualizes privacy loss as a rate of change: the *change* in the (distribution of) the *output* statistics per unit *change* in the *input* data.

- Key idea: To protect privacy *is* to limit this rate of change.
- Different DP specifications correspond to different choices of how (and where) to measure *input* and *output changes*, in addition to how much to control this rate.
- These choice are the *building blocks* of a DP specification:
    1. The protection *domain* ($\mathcal{X}$)
    2. The *scope* of protection ($\mathscr{D}$)
    3. The protection *unit* ($d_{\mathcal{X}}$)
    4. The *standard* of protection (d_{Pr})
    5. The *intensity* of protection ($\varepsilon_{\mathcal{D}}$)

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Object of interest: A statistic T – i.e. a function of the data $\boldsymbol{x} \in \mathcal{X}$

For example,

$$T(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Object of interest: A statistic T – i.e. a function of the data $\boldsymbol{x} \in \mathcal{X}$ and some auxiliary random noise Z.

For example,

$$T(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i + Z.$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about T as a function of the dataset $\boldsymbol{x} \in \mathcal{X}$, its derivative is

$$\lim_{\boldsymbol{x}' \to \boldsymbol{x}} \frac{T(\boldsymbol{x}', Z) - T(\boldsymbol{x}, Z)}{\boldsymbol{x}' - \boldsymbol{x}}.$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathrm{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its derivative is

$$\lim_{\boldsymbol{x}' \to \boldsymbol{x}} \frac{\mathrm{P}_{\boldsymbol{x}'}(T) - \mathrm{P}_{\boldsymbol{x}}(T)}{\boldsymbol{x}' - \boldsymbol{x}}.$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \overset{\textit{Data release}}{\rightsquigarrow} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathrm{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its derivative is

$$\lim_{\boldsymbol{x}' \to \boldsymbol{x}} \frac{d_{\mathrm{Pr}}(\mathrm{P}_{\boldsymbol{x}'}, \mathrm{P}_{\boldsymbol{x}})}{\boldsymbol{x}' - \boldsymbol{x}}.$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathsf{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its derivative is

$$\lim_{\boldsymbol{x}' \to \boldsymbol{x}} \frac{d_{\mathsf{Pr}}(\mathsf{P}_{\boldsymbol{x}'}, \mathsf{P}_{\boldsymbol{x}})}{d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x})}.$$

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathsf{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its derivative is

$$\cancel{\lim_{\boldsymbol{x}' \to \boldsymbol{x}}} \frac{d_{\mathsf{Pr}}(\mathsf{P}_{\boldsymbol{x}'}, \mathsf{P}_{\boldsymbol{x}})}{d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x})},$$

for all $\boldsymbol{x}, \boldsymbol{x}'$.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathrm{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest ε such that

$$d_{\mathrm{Pr}}(\mathrm{P}_{\boldsymbol{x}'}, \mathrm{P}_{\boldsymbol{x}}) \le \varepsilon d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}'$.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathsf{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest $\varepsilon_{\mathcal{D}}$ such that

$$d_{\mathsf{Pr}}(\mathsf{P}_{\boldsymbol{x}'}, \mathsf{P}_{\boldsymbol{x}}) \leq \varepsilon_{\mathcal{D}} \, d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all universes $\mathcal{D} \in \mathscr{D}$.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathsf{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest $\varepsilon_{\mathcal{D}}$ such that

$$d_{\mathsf{Pr}}(\mathsf{P}_{\boldsymbol{x}'}, \mathsf{P}_{\boldsymbol{x}}) \leq \varepsilon_{\mathcal{D}}\, d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all universes $\mathcal{D} \in \mathscr{D}$.

Definition: The statistic T is ε-differentially private if its Lipschitz constant is ε.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $P_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest $\varepsilon_{\mathcal{D}}$ such that

$$d_{\mathrm{Pr}}(P_{\boldsymbol{x}'}, P_{\boldsymbol{x}}) \leq \varepsilon_{\mathcal{D}} \, d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all universes $\mathcal{D} \in \mathscr{D}$.

Definition: The statistic T is ε-differentially private if its Lipschitz constant is ε.

- Recall that Lipschitz continuity $\approx$ differentiability.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathsf{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest $\varepsilon_{\mathcal{D}}$ such that

$$d_{\mathsf{Pr}}(\mathsf{P}_{\boldsymbol{x}'}, \mathsf{P}_{\boldsymbol{x}}) \leq \varepsilon_{\mathcal{D}} d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all universes $\mathcal{D} \in \mathscr{D}$.

Definition: The statistic T is ε-differentially private if its Lipschitz constant is ε.

- Recall that Lipschitz continuity $\approx$ differentiability.
- Lipschitz constant is the supremum of the derivative.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathrm{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest $\varepsilon_{\mathcal{D}}$ such that

$$d_{\mathrm{Pr}}(\mathrm{P}_{\boldsymbol{x}'}, \mathrm{P}_{\boldsymbol{x}}) \leq \varepsilon_{\mathcal{D}}\, d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all universes $\mathcal{D} \in \mathscr{D}$.

Definition: The statistic T is ε-differentially private if its Lipschitz constant is ε.

- Recall that Lipschitz continuity $\approx$ differentiability.
- Lipschitz constant is the supremum of the derivative.

Takeaway: Differential privacy is a "bound on the derivative" of T.

The *Derivative* of DP

$$\textit{A population} \xrightarrow{\textit{Data collection}} \textit{Dataset } \boldsymbol{x} \rightsquigarrow^{\textit{Data release}} \textit{Statistic } T(\boldsymbol{x}, Z)$$

Thinking about the distribution $\mathsf{P}_{\boldsymbol{x}}$ of T as a function of $\boldsymbol{x} \in \mathcal{X}$, its ~~derivative~~ Lipschitz constant is the smallest $\varepsilon_{\mathcal{D}}$ such that

$$d_{\mathsf{Pr}}(\mathsf{P}_{\boldsymbol{x}'}, \mathsf{P}_{\boldsymbol{x}}) \leq \varepsilon_{\mathcal{D}}\, d_{\mathcal{X}}(\boldsymbol{x}', \boldsymbol{x}),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all universes $\mathcal{D} \in \mathscr{D}$.

Definition: The statistic T is ε-differentially private if its Lipschitz constant is ε.

- Recall that Lipschitz continuity $\approx$ differentiability.
- Lipschitz constant is the supremum of the derivative.

Takeaway: Differential privacy is a “bound on the derivative” of T.

- The choice of $\mathcal{X}$, $\mathscr{D}$, d_{Pr} and $d_{\mathcal{X}}$ determine the *flavour* of DP.

Example: ε-Indistinguishability (Pure DP)

Definition

(Definition 1 of Dwork, McSherry, Nissim, and Smith, 2006)

Let the dataset $\boldsymbol{x}$ be a vector of n records from some domain $\mathcal{R}$, typically of the form $\{0, 1\}^d$ or $\mathbb{R}^d$. A data release mechanism T is *ε-indistinguishable* if for all neighbors – i.e. pairs of datasets $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}^n$ which differ in exactly one record – and for all outputs $t \in \mathcal{T}$:

$$\left| \ln \frac{\mathsf{P}_{\boldsymbol{x}}\left(T(\boldsymbol{x}, U) = t\right)}{\mathsf{P}_{\boldsymbol{x}'}\left(T(\boldsymbol{x}', U) = t\right)} \right| \leq \varepsilon.$$

Example: ε-Indistinguishability (Pure DP)

Definition

(Definition 1 of Dwork, McSherry, Nissim, and Smith, 2006)

Let the dataset $\boldsymbol{x}$ be a vector of n records from some domain $\mathcal{R}$, typically of the form $\{0,1\}^d$ or $\mathbb{R}^d$. A data release mechanism T is *ε-indistinguishable* if for all neighbors – i.e. pairs of datasets $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}^n$ which differ in exactly one record – ~~and for all outputs $t \in \mathcal{T}$~~:

$$\sup_{t \in \mathcal{T}} \left| \ln \frac{\mathrm{P}_{\boldsymbol{x}}\left(T(\boldsymbol{x}, U) = t\right)}{\mathrm{P}_{\boldsymbol{x}'}\left(T(\boldsymbol{x}', U) = t\right)} \right| \le \varepsilon.$$

Example: ε-Indistinguishability (Pure DP)

Definition

(Definition 1 of Dwork, McSherry, Nissim, and Smith, 2006)

Let the dataset $\boldsymbol{x}$ be a vector of n records from some domain $\mathcal{R}$, typically of the form $\{0, 1\}^d$ or $\mathbb{R}^d$. A data release mechanism T is *ε-indistinguishable* if for all neighbors – i.e. pairs of datasets $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}^n$ which differ in exactly one record:

$$d_{\textsc{Mult}}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) \leq \varepsilon,$$

where

$$d_{\textsc{Mult}}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) = \sup_{t \in \mathcal{T}} \left| \ln \frac{\mathsf{P}_{\boldsymbol{x}}\left(T(\boldsymbol{x}, U) = t\right)}{\mathsf{P}_{\boldsymbol{x}'}\left(T(\boldsymbol{x}', U) = t\right)} \right|.$$

Example: ε-Indistinguishability (Pure DP)

Definition

(Definition 1 of Dwork, McSherry, Nissim, and Smith, 2006)

Let the dataset $\boldsymbol{x}$ be a vector of n records from some domain $\mathcal{R}$, typically of the form $\{0, 1\}^d$ or $\mathbb{R}^d$. A data release mechanism T is *ε-indistinguishable* if for all neighbors – i.e. pairs of datasets $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}^n$ ~~which differ in exactly one record~~ with $d_{\text{Ham}}(\boldsymbol{x}, \boldsymbol{x}') = 1$:

$$d_{\text{MULT}}(\mathrm{P}_{\boldsymbol{x}}, \mathrm{P}_{\boldsymbol{x}'}) \leq \varepsilon,$$

where

$$d_{\text{MULT}}(\mathrm{P}_{\boldsymbol{x}}, \mathrm{P}_{\boldsymbol{x}'}) = \sup_{t \in \mathcal{T}} \left| \ln \frac{\mathrm{P}_{\boldsymbol{x}}\left(T(\boldsymbol{x}, U) = t\right)}{\mathrm{P}_{\boldsymbol{x}'}\left(T(\boldsymbol{x}', U) = t\right)} \right|.$$

Example: ε-Indistinguishability (Pure DP)

Definition

(Definition 1 of Dwork, McSherry, Nissim, and Smith, 2006)

Let the dataset $\boldsymbol{x}$ be a vector of n records from some domain $\mathcal{R}$, typically of the form $\{0,1\}^d$ or $\mathbb{R}^d$. A data release mechanism T is *ε-indistinguishable* if for all neighbors – i.e. pairs of datasets $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}^n$ with $d_{\mathrm{Ham}}(\boldsymbol{x}, \boldsymbol{x}') = 1$:

$$d_{\mathsf{MULT}}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) \leq \varepsilon d_{\mathrm{Ham}}(\boldsymbol{x}, \boldsymbol{x}'),$$

where

$$d_{\mathsf{MULT}}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) = \sup_{t \in \mathcal{T}} \left| \ln \frac{\mathsf{P}_{\boldsymbol{x}}\left(T(\boldsymbol{x}, U) = t\right)}{\mathsf{P}_{\boldsymbol{x}'}\left(T(\boldsymbol{x}', U) = t\right)} \right|.$$

Example: ε-Indistinguishability (Pure DP)

Definition

(Definition 1 of Dwork, McSherry, Nissim, and Smith, 2006)

Let the dataset $\boldsymbol{x}$ be a vector of n records from some domain $\mathcal{R}$, typically of the form $\{0, 1\}^d$ or $\mathbb{R}^d$. A data release mechanism T is *ε-indistinguishable* if ~~for all neighbors – i.e. pairs of datasets~~ $\cancel{\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}^n}$ ~~with~~ $\cancel{d_{\mathrm{Ham}}(\boldsymbol{x}, \boldsymbol{x}') = 1}$~~:~~

$$d_{\text{MULT}}(\mathrm{P}_{\boldsymbol{x}}, \mathrm{P}_{\boldsymbol{x}'}) \leq \varepsilon d_{\mathrm{Ham}}(\boldsymbol{x}, \boldsymbol{x}'),$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X} = \bigcup_{n \in \mathbb{N}} \mathcal{R}^n$, where

$$d_{\text{MULT}}(\mathrm{P}_{\boldsymbol{x}}, \mathrm{P}_{\boldsymbol{x}'}) = \sup_{t \in \mathcal{T}} \left| \ln \frac{\mathrm{P}_{\boldsymbol{x}}\left(T(\boldsymbol{x}, U) = t\right)}{\mathrm{P}_{\boldsymbol{x}'}\left(T(\boldsymbol{x}', U) = t\right)} \right|.$$

A DP Specification $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, d_{\mathsf{Pr}}, \varepsilon_{\mathcal{D}})$

The building blocks of DP:

- The protection domain
- The scope of protection
- The protection unit
- The standard of protection
- The intensity of protection

A DP Specification $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, d_{\mathsf{Pr}}, \varepsilon_{\mathcal{D}})$

The building blocks of DP:

- The protection domain
  - *Who* is eligible for protection?
  - Defined by the set $\mathcal{X}$ of possible input datasets.
- The scope of protection
- The protection unit
- The standard of protection
- The intensity of protection

A DP Specification $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, d_{\mathsf{Pr}}, \varepsilon_{\mathcal{D}})$

The building blocks of DP:

- The protection domain
  - *Who* is eligible for protection?
  - Defined by the set $\mathcal{X}$ of possible input datasets.
- The scope of protection
  - *Where* does the protection extend to?
  - Instantiated by the multiverse $\mathscr{D}$, which is a collection of universes $\mathcal{D} \subset \mathcal{X}$.
- The protection unit
- The standard of protection
- The intensity of protection

A DP Specification $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, d_{\mathsf{Pr}}, \varepsilon_{\mathcal{D}})$

The building blocks of DP:

- The protection domain
  - *Who* is eligible for protection?
  - Defined by the set $\mathcal{X}$ of possible input datasets.
- The scope of protection
  - *Where* does the protection extend to?
  - Instantiated by the multiverse $\mathscr{D}$, which is a collection of universes $\mathcal{D} \subset \mathcal{X}$.
- The protection unit
  - *What* is the granularity of protection?
  - Conceptualized by the input premetric $d_{\mathcal{X}}$ on $\mathcal{X}$.
- The standard of protection
- The intensity of protection

A DP Specification $(\mathscr{X}, \mathscr{D}, d_{\mathcal{X}}, d_{\mathsf{Pr}}, \varepsilon_{\mathcal{D}})$

The building blocks of DP:

- The protection domain
  - *Who* is eligible for protection?
  - Defined by the set $\mathscr{X}$ of possible input datasets.
- The scope of protection
  - *Where* does the protection extend to?
  - Instantiated by the multiverse $\mathscr{D}$, which is a collection of universes $\mathcal{D} \subset \mathcal{X}$.
- The protection unit
  - *What* is the granularity of protection?
  - Conceptualized by the input premetric $d_{\mathcal{X}}$ on $\mathcal{X}$.
- The standard of protection
  - *How* to measure change in the output variations?
  - Captured by the output premetric d_{Pr} on probability distributions.
- The intensity of protection

A DP Specification $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, d_{\mathsf{Pr}}, \varepsilon_{\mathcal{D}})$

The building blocks of DP:

- The protection domain
  - *Who* is eligible for protection?
  - Defined by the set $\mathcal{X}$ of possible input datasets.
- The scope of protection
  - *Where* does the protection extend to?
  - Instantiated by the multiverse $\mathscr{D}$, which is a collection of universes $\mathcal{D} \subset \mathcal{X}$.
- The protection unit
  - *What* is the granularity of protection?
  - Conceptualized by the input premetric $d_{\mathcal{X}}$ on $\mathcal{X}$.
- The standard of protection
  - *How* to measure change in the output variations?
  - Captured by the output premetric d_{Pr} on probability distributions.
- The intensity of protection
  - *How much* protection is afforded?
  - Quantified by the protection loss budget (PLB) $\varepsilon_{\mathcal{D}}$.

Some Examples in the Literature

$\underline{\mathscr{X}}$: DP for network data (Hay, Li, Miklau, and Jensen, 2009) for geospatial data (Andrés, Bordenabe, Chatzikokolakis, and Palamidessi, 2013) Pufferfish DP (Kifer and Machanavajjhala, 2014) noiseless privacy (Bhaskar et al., 2011) privacy under partial knowledge (Seeman, Reimherr, and Slavkovic, 2022) privacy amplification (Balle, Barthe, and Gaboardi, 2020; Beimel, Kasiviswanathan, and Nissim, 2010; Bun et al., 2022)

$\underline{\mathscr{D}}$: privacy under invariants (Ashmead et al., 2019; Dharangutte, Gao, Gong, and Yu, 2023; Gao, Gong, and Yu, 2022; Gong and Meng, 2020) conditioned or empirical DP (Abowd, Schneider, and Vilhuber, 2013; Charest and Hou, 2016) personalized DP (Ebadi, Sands, and Schneider, 2015; Jorgensen, Yu, and Cormode, 2015) individual DP (Feldman and Zrnic, 2022; Soria-Comas, Domingo-Ferrer, Sánchez, and Megías, 2017) bootstrap DP (O'Keefe and Charest, 2019) stratified DP (Bun et al., 2022) per-record DP (Seeman, Sexton, Pujol, and Machanavajjhala, 2024) per-instance DP (Redberg and Wang, 2021; Wang, 2018)

$\underline{d_{\mathcal{X}}}$: $(\mathcal{R}, \varepsilon)$-generic DP (Kifer and Machanavajjhala, 2011a) edge vs node privacy (Hay, Li, Miklau, and Jensen, 2009; McSherry and Mahajan, 2010) d-metric DP (Chatzikokolakis, Andrés, Bordenabe, and Palamidessi, 2013) Blowfish privacy (He, Machanavajjhala, and Ding, 2014) element level DP (Asi, Duchi, and Javidbakht, 2022) distributional privacy (Zhou, Ligett, and Wasserman, 2009) event-level vs user-level DP (Dwork, Naor, Pitassi, and Rothblum, 2010)

$\underline{d_{\mathsf{Pr}}}$: (ε, δ)-approximate DP (Dwork et al., 2006) Rényi DP (Mironov, 2017) concentrated DP (Bun and Steinke, 2016) f-divergence privacy (Barber and Duchi, 2014; Barthe and Olmedo, 2013) f-DP (including Gaussian DP) (Dong, Roth, and Su, 2022)

Examples of $\mathscr{D}$, $d_{\mathcal{X}}$ and d_{Pr}

1. An *invariant-compliant data universe*:

$$\mathscr{D}_{\boldsymbol{c}} = \left\{ \mathcal{D} \subset \mathcal{X} : \boldsymbol{c}(\boldsymbol{x}) = \boldsymbol{c}(\boldsymbol{x}') \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D} \right\},$$

   for some invariants $\boldsymbol{c} : \mathcal{X} \to \mathbb{R}^l$.

Examples of $\mathscr{D}$, $d_{\mathcal{X}}$ and d_{Pr}

1. An *invariant-compliant data universe*:

$$\mathscr{D}_{\boldsymbol{c}} = \left\{ \mathcal{D} \subset \mathcal{X} : \boldsymbol{c}(\boldsymbol{x}) = \boldsymbol{c}(\boldsymbol{x}') \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D} \right\},$$

   for some invariants $\boldsymbol{c} : \mathcal{X} \to \mathbb{R}^l$.

2. *Input premetric* $d_{\mathcal{X}}$ induced by a "neighbour" relation:

$$d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}') = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{x}', \\ 1 & \text{if } \boldsymbol{x} \text{ and } \boldsymbol{x}' \text{ are ``neighbours''}, \\ \infty & \text{otherwise.} \end{cases}$$

Examples of $\mathscr{D}$, $d_{\mathcal{X}}$ and d_{Pr}

3. *Output premetric* d_{Pr} on (the probability distributions over) the output space

Examples of $\mathscr{D}$, $d_{\mathcal{X}}$ and d_{Pr}

3. *Output premetric* d_{Pr} on (the probability distributions over) the output space
   - *Pure DP* (Dwork, McSherry, Nissim, and Smith, 2006): d_{Pr} is the multiplicative distance

$$d_{\mathrm{MULT}}(\mathsf{P}, \mathsf{Q}) = \sup \left\{ \left| \ln \frac{\mathsf{P}(S)}{\mathsf{Q}(S)} \right| : \text{event } S \right\}.$$

Examples of $\mathscr{D}$, $d_{\mathcal{X}}$ and d_{Pr}

3. *Output premetric* d_{Pr} on (the probability distributions over) the output space
   - *Pure DP* (Dwork, McSherry, Nissim, and Smith, 2006): d_{Pr} is the multiplicative distance

$$d_{\text{MULT}}(\mathrm{P}, \mathrm{Q}) = \sup \left\{ \left| \ln \frac{\mathrm{P}(S)}{\mathrm{Q}(S)} \right| : \text{event } S \right\}.$$

   - *Approximate DP* (Dwork et al., 2006):

$$d^{\delta}_{\text{MULT}}(\mathrm{P}, \mathrm{Q}) = \sup_{\text{event } S} \left\{ \ln \frac{[\mathrm{P}(S) - \delta]^{+}}{\mathrm{Q}(S)}, \ln \frac{[\mathrm{Q}(S) - \delta]^{+}}{\mathrm{P}(S)}, 0 \right\},$$

Examples of $\mathscr{D}$, $d_{\mathcal{X}}$ and d_{Pr}

3. *Output premetric* d_{Pr} on (the probability distributions over) the output space
   - *Pure DP* (Dwork, McSherry, Nissim, and Smith, 2006): d_{Pr} is the multiplicative distance

$$d_{\text{MULT}}(\mathsf{P}, \mathsf{Q}) = \sup \left\{ \left| \ln \frac{\mathsf{P}(S)}{\mathsf{Q}(S)} \right| : \text{event } S \right\}.$$

   - *Approximate DP* (Dwork et al., 2006):

$$d^{\delta}_{\text{MULT}}(\mathsf{P}, \mathsf{Q}) = \sup_{\text{event } S} \left\{ \ln \frac{[\mathsf{P}(S) - \delta]^{+}}{\mathsf{Q}(S)}, \ln \frac{[\mathsf{Q}(S) - \delta]^{+}}{\mathsf{P}(S)}, 0 \right\},$$

   - *Zero Concentrated DP* (Bun and Steinke, 2016):

$$D_{\text{nor}}(\mathsf{P}, \mathsf{Q}) = \sup_{\alpha > 1} \frac{1}{\sqrt{\alpha}} \max \left[\sqrt{D_{\alpha}(\mathsf{P}||\mathsf{Q})}, \sqrt{D_{\alpha}(\mathsf{Q}||\mathsf{P})} \right],$$

   where D_{α} is the *Rényi divergence* of order α:

$$D_{\alpha}(\mathsf{P}||\mathsf{Q}) = \frac{1}{\alpha - 1} \ln \int \left[\frac{d\mathsf{P}}{d\mathsf{Q}} \right]^{\alpha} d\mathsf{Q},$$

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.
- Report $T_i = 1$ if $X_i = U_i$, and otherwise report $T_i = 0$.

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.
- Report $T_i = 1$ if $X_i = U_i$, and otherwise report $T_i = 0$.
- At the individual level, $T_i = 1$ can mean i is a cheater or not a cheater.

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.
- Report $T_i = 1$ if $X_i = U_i$, and otherwise report $T_i = 0$.
- At the individual level, $T_i = 1$ can mean i is a cheater or not a cheater.
- But in aggregation:

$$p_T = \Pr(U = X) = p \times p_{\text{cheat}} + (1 - p) \times (1 - p_{\text{cheat}}).$$

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.
- Report $T_i = 1$ if $X_i = U_i$, and otherwise report $T_i = 0$.
- At the individual level, $T_i = 1$ can mean i is a cheater or not a cheater.
- But in aggregation:

$$p_T = \Pr(U = X) = p \times p_{\text{cheat}} + (1 - p) \times (1 - p_{\text{cheat}}).$$

Recovering p_{cheat}:

$$p_{\text{cheat}} = \frac{p_T + p - 1}{2p - 1}$$

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.
- Report $T_i = 1$ if $X_i = U_i$, and otherwise report $T_i = 0$.
- At the individual level, $T_i = 1$ can mean i is a cheater or not a cheater.
- But in aggregation:

$$p_T = \Pr(U = X) = p \times p_{\text{cheat}} + (1 - p) \times (1 - p_{\text{cheat}}).$$

Recovering p_{cheat}:

$$p_{\text{cheat}} = \frac{p_T + p - 1}{2p - 1}$$

Estimate:

$$\hat{p}_{\text{cheat}} = \frac{\bar{T}_n + p - 1}{2p - 1}$$

Example: Randomised Response (Warner, 1965)

- Estimating exam cheating rate p_{cheat}. $X_i = 1$: cheated; $X_i = 0$: not cheated.
- Each student i tosses a biased coin (with $p \neq 0.5$) secretly before answering. $U_i = 1$ if heads, and $U_i = 0$ if tails.
- Report $T_i = 1$ if $X_i = U_i$, and otherwise report $T_i = 0$.
- At the individual level, $T_i = 1$ can mean i is a cheater or not a cheater.
- But in aggregation:

$$p_T = \Pr(U = X) = p \times p_{\text{cheat}} + (1 - p) \times (1 - p_{\text{cheat}}).$$

Recovering p_{cheat}:

$$p_{\text{cheat}} = \frac{p_T + p - 1}{2p - 1}$$

Estimate:

$$\hat{p}_{\text{cheat}} = \frac{\bar{T}_n + p - 1}{2p - 1}$$

Ex: $\bar{T}_n = 0.45$, $p = 0.6$

$$\hat{p}_{\text{cheat}} = \frac{0.45 + 0.6 - 1}{2 \times 0.6 - 1} = 0.25$$

What Is the Loss of Information and What Is the Gain in Protection?

Increased variance:

$$\mathrm{Var}(\hat{p}_{\mathrm{cheat}}) = \frac{1}{n}\frac{p_T(1-p_T)}{(2p-1)^2} \leq \frac{1}{16n}\frac{1}{(p-0.5)^2}.$$

What Is the Loss of Information and What Is the Gain in Protection?

Increased variance:

$$\mathrm{Var}(\hat{p}_{\text{cheat}}) = \frac{1}{n}\frac{p_T(1-p_T)}{(2p-1)^2} \leq \frac{1}{16n}\frac{1}{(p-0.5)^2}.$$

The “first” example of *differential privacy*:

$$\frac{\mathrm{P}\left(T_i = 1 \mid X_i = 1\right)}{\mathrm{P}\left(T_i = 1 \mid X_i = 0\right)} = \frac{p}{1-p} = e^{\varepsilon}, \quad \text{with } \varepsilon = \mathrm{logit}(p),$$

What Is the Loss of Information and What Is the Gain in Protection?

Increased variance:

$$\mathrm{Var}(\hat{p}_{\text{cheat}}) = \frac{1}{n}\frac{p_T(1-p_T)}{(2p-1)^2} \leq \frac{1}{16n}\frac{1}{(p-0.5)^2}.$$

The “first” example of *differential privacy*:

$$\frac{\mathsf{P}\left(T_i = 1 \mid X_i = 1\right)}{\mathsf{P}\left(T_i = 1 \mid X_i = 0\right)} = \frac{p}{1-p} = e^{\varepsilon}, \quad \text{with } \varepsilon = \mathrm{logit}(p),$$

$$\frac{\mathsf{Pr}\left(T_i = 0 \mid X_i = 1\right)}{\mathsf{Pr}\left(T_i = 0 \mid X_i = 0\right)} = \frac{1-p}{p} = e^{-\varepsilon},$$

What Is the Loss of Information and What Is the Gain in Protection?

Increased variance:

$$\mathrm{Var}(\hat{p}_{\mathrm{cheat}}) = \frac{1}{n}\frac{p_T(1-p_T)}{(2p-1)^2} \leq \frac{1}{16n}\frac{1}{(p-0.5)^2}.$$

The "first" example of *differential privacy*:

$$\frac{\mathsf{P}\left(T_i = 1 \mid X_i = 1\right)}{\mathsf{P}\left(T_i = 1 \mid X_i = 0\right)} = \frac{p}{1-p} = e^{\varepsilon}, \quad \text{with } \varepsilon = \mathrm{logit}(p),$$

$$\frac{\mathsf{Pr}\left(T_i = 0 \mid X_i = 1\right)}{\mathsf{Pr}\left(T_i = 0 \mid X_i = 0\right)} = \frac{1-p}{p} = e^{-\varepsilon},$$

$$e^{-\varepsilon} \leq \sup_{t=0,1} \frac{\mathsf{Pr}\left(T_i = t \mid X_i = 1\right)}{\mathsf{Pr}\left(T_i = t \mid X_i = 0\right)} \leq e^{\varepsilon}.$$

What Is the Loss of Information and What Is the Gain in Protection?

Increased variance:

$$\mathrm{Var}(\hat{p}_{\mathrm{cheat}}) = \frac{1}{n}\frac{p_T(1-p_T)}{(2p-1)^2} \leq \frac{1}{16n}\frac{1}{(p-0.5)^2}.$$

The "first" example of *differential privacy*:

$$\frac{\mathsf{P}\left(T_i = 1 \mid X_i = 1\right)}{\mathsf{P}\left(T_i = 1 \mid X_i = 0\right)} = \frac{p}{1-p} = e^{\varepsilon}, \quad \text{with } \varepsilon = \mathrm{logit}(p),$$

$$\frac{\Pr\left(T_i = 0 \mid X_i = 1\right)}{\Pr\left(T_i = 0 \mid X_i = 0\right)} = \frac{1-p}{p} = e^{-\varepsilon},$$

$$e^{-\varepsilon} \leq \sup_{t=0,1} \frac{\Pr\left(T_i = t \mid X_i = 1\right)}{\Pr\left(T_i = t \mid X_i = 0\right)} \leq e^{\varepsilon}.$$

$$d_{\mathrm{MULT}}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) \leq \varepsilon d_{\mathrm{Ham}}(\boldsymbol{x}, \boldsymbol{x}'), \quad \text{for } \boldsymbol{x}, \boldsymbol{x}' \in \{0, 1\}^n.$$

Does Pure DP Control Disclosure?

One can argue (Dalenius, 1977): Control disclosure $\Leftrightarrow$ control the "difference" between $\pi(X_i)$ and $\pi(X_i \mid T = t)$ for any Bayesian attacker π.

Does Pure DP Control Disclosure?

One can argue (Dalenius, 1977): Control disclosure $\Leftrightarrow$ control the "difference" between $\pi(X_i)$ and $\pi(X_i \mid T = t)$ for any Bayesian attacker π.

The "strongest" attacker knows the values of $\boldsymbol{x}_{-i}$:

$$\pi(\boldsymbol{X} = \boldsymbol{x}) = \pi(X_i = x_i)\delta_{\boldsymbol{x}_{-i}=\boldsymbol{x}^*_{-i}}.$$

Does Pure DP Control Disclosure?

One can argue (Dalenius, 1977): Control disclosure $\Leftrightarrow$ control the "difference" between $\pi(X_i)$ and $\pi(X_i \mid T = t)$ for any Bayesian attacker π.

The "strongest" attacker knows the values of $\boldsymbol{x}_{-i}$:

$$\pi(\boldsymbol{X} = \boldsymbol{x}) = \pi(X_i = x_i)\delta_{\boldsymbol{x}_{-i} = \boldsymbol{x}^*_{-i}}.$$

Then

$$\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} = \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')}$$

Does Pure DP Control Disclosure?

One can argue (Dalenius, 1977): Control disclosure $\Leftrightarrow$ control the “difference” between $\pi(X_i)$ and $\pi(X_i \mid T = t)$ for any Bayesian attacker π.

The “strongest” attacker knows the values of $\boldsymbol{x}_{-i}$:

$$\pi(\boldsymbol{X} = \boldsymbol{x}) = \pi(X_i = x_i)\delta_{\boldsymbol{x}_{-i} = \boldsymbol{x}^*_{-i}}.$$

Then

$$\begin{aligned}\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} &= \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\ &= \frac{\int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')}\end{aligned}$$

Does Pure DP Control Disclosure?

One can argue (Dalenius, 1977): Control disclosure $\Leftrightarrow$ control the "difference" between $\pi(X_i)$ and $\pi(X_i \mid T = t)$ for any Bayesian attacker π.

The "strongest" attacker knows the values of $\boldsymbol{x}_{-i}$:

$$\pi(\boldsymbol{X} = \boldsymbol{x}) = \pi(X_i = x_i)\delta_{\boldsymbol{x}_{-i} = \boldsymbol{x}^*_{-i}}.$$

Then

$$\begin{aligned}
\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} &= \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\mathrm{p}(T = t \mid X_i = x_i, \boldsymbol{X}_{-i} = \boldsymbol{x}^*_{-i})}{\int \mathrm{p}(T = t \mid X_i = x'_i, \boldsymbol{X}_{-i} = \boldsymbol{x}^*_{-i}) d\pi(X_i = x'_i)}
\end{aligned}$$

Does Pure DP Control Disclosure?

One can argue (Dalenius, 1977): Control disclosure $\Leftrightarrow$ control the "difference" between $\pi(X_i)$ and $\pi(X_i \mid T = t)$ for any Bayesian attacker π.

The "strongest" attacker knows the values of $\boldsymbol{x}_{-i}$:

$$\pi(\boldsymbol{X} = \boldsymbol{x}) = \pi(X_i = x_i)\delta_{\boldsymbol{x}_{-i} = \boldsymbol{x}^*_{-i}}.$$

Then

$$\begin{aligned}
\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} &= \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\mathrm{p}(T = t \mid X_i = x_i, \boldsymbol{X}_{-i} = \boldsymbol{x}^*_{-i})}{\int \mathrm{p}(T = t \mid X_i = x_i', \boldsymbol{X}_{-i} = \boldsymbol{x}^*_{-i}) d\pi(X_i = x_i')} \\
&\leq e^{\varepsilon}.
\end{aligned}$$

Example: Randomised Response (cont.)

Recall $T_i = 1_{\{X_i = U_i\}}$.

Suppose an adversary's prior for X_1 is $\pi(X_1 = 1) = \theta$. Given $t \in \{0, 1\}$,

$$\begin{aligned} C_\theta(t) &:= \frac{\pi(X_1 = 1 \mid T_1 = t)}{\pi(X_1 = 1)} = \frac{\Pr(T_1 = t \mid X_1 = 1)}{\Pr(T_1 = t)} \\ &= \frac{LR(t)}{LR(t)\theta + (1 - \theta)}, \quad \text{where } LR(t) = \frac{\Pr\left(T_1 = t \mid X_1 = 1\right)}{\Pr\left(T_1 = t \mid X_1 = 0\right)} \end{aligned}$$

Example: Randomised Response (cont.)

Recall $T_i = 1_{\{X_i = U_i\}}$.

Suppose an adversary's prior for X_1 is $\pi(X_1 = 1) = \theta$. Given $t \in \{0, 1\}$,

$$C_\theta(t) := \frac{\pi(X_1 = 1 \mid T_1 = t)}{\pi(X_1 = 1)} = \frac{\Pr(T_1 = t \mid X_1 = 1)}{\Pr(T_1 = t)}$$

$$= \frac{LR(t)}{LR(t)\theta + (1-\theta)}, \quad \text{where } LR(t) = \frac{\Pr\left(T_1 = t \mid X_1 = 1\right)}{\Pr\left(T_1 = t \mid X_1 = 0\right)}$$

$$LR(t) \geq 1 \ \Rightarrow 1 \leq C_\theta(t) \leq LR(t)$$

$$\max_\theta C_\theta(t) = C_0(t) = LR(t)$$

$$\min_\theta C_\theta(t) = C_1(t) = 1$$

Example: Randomised Response (cont.)

Recall $T_i = 1_{\{X_i = U_i\}}$.

Suppose an adversary's prior for X_1 is $\pi(X_1 = 1) = \theta$. Given $t \in \{0, 1\}$,

$$C_\theta(t) := \frac{\pi(X_1 = 1 \mid T_1 = t)}{\pi(X_1 = 1)} = \frac{\Pr(T_1 = t \mid X_1 = 1)}{\Pr(T_1 = t)}$$

$$= \frac{LR(t)}{LR(t)\theta + (1-\theta)}, \quad \text{where } LR(t) = \frac{\Pr\left(T_1 = t \mid X_1 = 1\right)}{\Pr\left(T_1 = t \mid X_1 = 0\right)}$$

$$LR(t) \geq 1 \ \Rightarrow 1 \leq C_\theta(t) \leq LR(t) \qquad LR(t) \leq 1 \ \Rightarrow LR(t) \leq C_\theta(t) \leq 1$$

$$\max_\theta C_\theta(t) = C_0(t) = LR(t) \qquad \max_\theta C_\theta(t) = C_1(t) = 1$$

$$\min_\theta C_\theta(t) = C_1(t) = 1 \qquad \min_\theta C_\theta(t) = C_0(t) = LR(t)$$

Example: Randomised Response (cont.)

Recall $T_i = 1_{\{X_i = U_i\}}$.

Suppose an adversary's prior for X_1 is $\pi(X_1 = 1) = \theta$. Given $t \in \{0, 1\}$,

$$C_\theta(t) := \frac{\pi(X_1 = 1 \mid T_1 = t)}{\pi(X_1 = 1)} = \frac{\Pr(T_1 = t \mid X_1 = 1)}{\Pr(T_1 = t)}$$

$$= \frac{LR(t)}{LR(t)\theta + (1-\theta)}, \quad \text{where } LR(t) = \frac{\Pr(T_1 = t \mid X_1 = 1)}{\Pr(T_1 = t \mid X_1 = 0)}$$

$$LR(t) \geq 1 \Rightarrow 1 \leq C_\theta(t) \leq LR(t) \qquad LR(t) \leq 1 \Rightarrow LR(t) \leq C_\theta(t) \leq 1$$

$$\max_\theta C_\theta(t) = C_0(t) = LR(t) \qquad \max_\theta C_\theta(t) = C_1(t) = 1$$

$$\min_\theta C_\theta(t) = C_1(t) = 1 \qquad \min_\theta C_\theta(t) = C_0(t) = LR(t)$$

The prior-to-posterior semantics for differential privacy:

$$e^{-\varepsilon} \leq C_\theta(t) \leq e^{\varepsilon} \quad \text{for all } \theta \text{ if and only if} \quad e^{-\varepsilon} \leq LR(t) \leq e^{\varepsilon} \quad \text{for all } t$$

However, What if X_1 and X_2 Are *A-Priori* Dependent?

Suppose our prior for (X_1, X_2) is $\pi(X_1 = a, X_2 = b) = \theta_{ab}$. Let

$$C_\pi(t_1, t_2) := \frac{\Pr(X_1 = 1 | T_1 = t_1, T_2 = t_2)}{\Pr(X_1 = 1)} = \frac{\Pr(T_1 = t_1, T_2 = t_2 | X_1 = 1)}{\Pr(T_1 = t_1, T_2 = t_2)}$$

Transferring the bound on likelihood ratio to posterior-to-prior ratio

$$C_\theta(t_1, t_2) = \frac{LR(t_1, t_2)}{LR(t_1, t_2)\theta_{1\cdot} + (1 - \theta_{1\cdot})}, \quad \theta_{1\cdot} = \pi(X_1 = 1) = \theta_{11} + \theta_{10}$$

$$LR(t_1, t_2) = \frac{\Pr(T_1 = t_1, T_2 = t_2 | X_1 = 1)}{\Pr(T_1 = t_1, T_2 = t_2 | X_1 = 0)}.$$

However, What if X_1 and X_2 Are *A-Priori* Dependent?

Suppose our prior for (X_1, X_2) is $\pi(X_1 = a, X_2 = b) = \theta_{ab}$. Let

$$C_\pi(t_1, t_2) := \frac{\Pr(X_1 = 1 | T_1 = t_1, T_2 = t_2)}{\Pr(X_1 = 1)} = \frac{\Pr(T_1 = t_1, T_2 = t_2 | X_1 = 1)}{\Pr(T_1 = t_1, T_2 = t_2)}$$

Transferring the bound on likelihood ratio to posterior-to-prior ratio

$$C_\theta(t_1, t_2) = \frac{LR(t_1, t_2)}{LR(t_1, t_2)\theta_{1\cdot} + (1 - \theta_{1\cdot})}, \quad \theta_{1\cdot} = \pi(X_1 = 1) = \theta_{11} + \theta_{10}$$

$$LR(t_1, t_2) = \frac{\Pr(T_1 = t_1, T_2 = t_2 | X_1 = 1)}{\Pr(T_1 = t_1, T_2 = t_2 | X_1 = 0)}.$$

Consider the case $t_1 = 1, t_2 = 1$, and recall $e^\varepsilon = p/(1 - p)$

$$LR(1, 1) = \frac{e^\varepsilon \frac{\theta_{11}}{\theta_{1\cdot}} + \frac{\theta_{10}}{\theta_{1\cdot}}}{\frac{\theta_{01}}{\theta_{0\cdot}} + e^{-\varepsilon} \frac{\theta_{00}}{\theta_{0\cdot}}}$$

The Dependence Is a Big Trouble Maker

This means that when $\theta_{10} = \theta_{01} = 0$, $LR(1, 1) = e^{2\varepsilon} > e^{\varepsilon}$.

- But $\theta_{10} = \theta_{01} = 0$ means that $X_2 = X_1$, hence X_1 can be learned from the information for X_2. Consequently, the "individual information unit" for X_1 should be the pair $\{X_1, X_2\}$, not merely X_1.

The Dependence Is a Big Trouble Maker

This means that when $\theta_{10} = \theta_{01} = 0$, $LR(1, 1) = e^{2\varepsilon} > e^{\varepsilon}$.

- But $\theta_{10} = \theta_{01} = 0$ means that $X_2 = X_1$, hence X_1 can be learned from the information for X_2. Consequently, the "individual information unit" for X_1 should be the pair $\{X_1, X_2\}$, not merely X_1.
- In fact as soon as $\mathrm{Cov}(X_1, X_2) > 0$, $LR(1, 1) > e^{\varepsilon}$. This is because

$$LR(1,1) > e^{\varepsilon} \iff \pi(X_2 = 1|X_1 = 1) > \pi(X_2 = 1|X_1 = 0)$$

But

$$\begin{aligned}\mathrm{Cov}(X_1, X_2) &= \pi(X_1 = 1, X_2 = 1) - \pi(X_1 = 1)\,\Pr(X_2 = 1) \\ &= [\pi(X_2 = 1|X_1 = 1) - \pi(X_2 = 1|X_1 = 0)]\,\pi(X_1 = 0)\pi(X_1 = 1).\end{aligned}$$

The Dependence Is a Big Trouble Maker

This means that when $\theta_{10} = \theta_{01} = 0$, $LR(1, 1) = e^{2\varepsilon} > e^{\varepsilon}$.

- But $\theta_{10} = \theta_{01} = 0$ means that $X_2 = X_1$, hence X_1 can be learned from the information for X_2. Consequently, the "individual information unit" for X_1 should be the pair $\{X_1, X_2\}$, not merely X_1.
- In fact as soon as $\mathrm{Cov}(X_1, X_2) > 0$, $LR(1, 1) > e^{\varepsilon}$. This is because

$$LR(1,1) > e^{\varepsilon} \quad \Longleftrightarrow \quad \pi(X_2 = 1 | X_1 = 1) > \pi(X_2 = 1 | X_1 = 0)$$

But

$$\begin{aligned} \mathrm{Cov}(X_1, X_2) &= \pi(X_1 = 1, X_2 = 1) - \pi(X_1 = 1)\Pr(X_2 = 1) \\ &= [\pi(X_2 = 1 | X_1 = 1) - \pi(X_2 = 1 | X_1 = 0)]\,\pi(X_1 = 0)\pi(X_1 = 1). \end{aligned}$$

Data are *accidental* representation, not *essential* information

Manipulating data values without considering their interdependence is not a legitimate information operation in general

Does Pure DP Control Disclosure?

For a general prior π,

$$\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} = \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')}$$

Does Pure DP Control Disclosure?

For a general prior π,

$$
\begin{aligned}
\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} &= \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')}
\end{aligned}
$$

Does Pure DP Control Disclosure?

For a general prior π,

$$
\begin{aligned}
\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} &= \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \int \frac{1}{\int \frac{\mathrm{p}_{\boldsymbol{x}'}(T=t)}{\mathrm{p}_{\boldsymbol{x}}(T=t)} d\pi(\boldsymbol{X} = \boldsymbol{x}')} d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)
\end{aligned}
$$

Does Pure DP Control Disclosure?

For a general prior π,

$$
\begin{aligned}
\frac{\pi(X_i = x_i \mid T = t)}{\pi(X_i = x_i)} &= \frac{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\pi(X_i = x_i) \int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \frac{\int \mathrm{p}_{\boldsymbol{x}}(T = t) d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i)}{\int \mathrm{p}_{\boldsymbol{x}'}(T = t) d\pi(\boldsymbol{X} = \boldsymbol{x}')} \\
&= \int \frac{1}{\int \frac{\mathrm{p}_{\boldsymbol{x}'}(T=t)}{\mathrm{p}_{\boldsymbol{x}}(T=t)} d\pi(\boldsymbol{X} = \boldsymbol{x}')} d\pi(\boldsymbol{X}_{-i} = \boldsymbol{x}_{-i} \mid X_i = x_i) \\
&\leq e^{n\varepsilon},
\end{aligned}
$$

with equality as the records of $\boldsymbol{X}$ become totally dependent. (n is the number of records in $\boldsymbol{X}$.) (Dwork, McSherry, Nissim, and Smith, 2006; Kifer and Machanavajjhala, 2011b)

What Does DP Actually Protect?

What Does DP Actually Protect?

- Does ε-DP guarantee the marginal prior-to-posterior ratio

$$e^{-\varepsilon} \leq \frac{\pi(X_i = x | T = t)}{\pi(X_i = x)} \leq e^{\varepsilon}, \quad \forall x, \forall t?$$ **No, not in general**

(Kifer and Machanavajjhala, 2011b, 2012; Tschantz, Sen, and Datta, 2020)

What Does DP Actually Protect?

- Does ε-DP guarantee the marginal prior-to-posterior ratio
$$e^{-\varepsilon} \leq \frac{\pi(X_i = x | T = t)}{\pi(X_i = x)} \leq e^{\varepsilon}, \quad \forall x, \forall t? \quad \textbf{No, not in general}$$

(Kifer and Machanavajjhala, 2011b, 2012; Tschantz, Sen, and Datta, 2020)

- Does ε-DP guarantee the conditional prior-to-posterior ratio
$$e^{-\varepsilon} \leq \frac{\pi(X_i = x_i | T = t, \boldsymbol{X}_{-i})}{\pi(X_i = x | \boldsymbol{X}_{-i})} \leq e^{\varepsilon}, \quad \forall x, \forall t? \quad \textbf{Yes}$$

What Does DP Actually Protect?

- Does ε-DP guarantee the marginal prior-to-posterior ratio

$$e^{-\varepsilon} \leq \frac{\pi(X_i = x|T = t)}{\pi(X_i = x)} \leq e^{\varepsilon}, \quad \forall x, \forall t?$$

No, not in general

(Kifer and Machanavajjhala, 2011b, 2012; Tschantz, Sen, and Datta, 2020)

- Does ε-DP guarantee the conditional prior-to-posterior ratio

$$e^{-\varepsilon} \leq \frac{\pi(X_i = x_i|T = t, \boldsymbol{X}_{-i})}{\pi(X_i = x|\boldsymbol{X}_{-i})} \leq e^{\varepsilon}, \quad \forall x, \forall t?$$

Yes

- Thus the guaranteed limit e^{ε} is only for the **unique individual information**: variations unexplained by anyone else in the database or by knowledge on (and beyond) the database population.

In General, What Does DP Actually Guarantee?

A random statistic $T \in \mathbb{R}^d$ is pure DP with PLB ε and input premetric d_{Ham} if and only if for every prior π on $\boldsymbol{X}$, every sub-σ field $\mathcal{F}$ of the corresponding full σ-field σ_π, every $B \in \mathscr{B}\left(\mathbb{R}^d\right)$, every i, and every $A \in \mathscr{B}(\Theta_i)$ (where Θ_i is the state space of x_i), we have

$$e^{-c_i \varepsilon} \pi(X_i \in A \mid \mathcal{F}) \leq \pi\left(X_i \in A \mid T \in B; \mathcal{F}\right) \leq e^{c_i \varepsilon} \pi(X_i \in A \mid \mathcal{F}),$$

where c_i is the size of the *minimal information chamber* (MIC) for X_i.

In General, What Does DP Actually Guarantee?

A random statistic $T \in \mathbb{R}^d$ is pure DP with PLB ε and input premetric d_{Ham} if and only if for every prior π on $\boldsymbol{X}$, every sub-σ field $\mathcal{F}$ of the corresponding full σ-field σ_π, every $B \in \mathscr{B}\left(\mathbb{R}^d\right)$, every i, and every $A \in \mathscr{B}(\Theta_i)$ (where Θ_i is the state space of x_i), we have

$$e^{-c_i\varepsilon}\pi(X_i \in A \mid \mathcal{F}) \leq \pi\left(X_i \in A \mid T \in B; \mathcal{F}\right) \leq e^{c_i\varepsilon}\pi(X_i \in A \mid \mathcal{F}),$$

where c_i is the size of the *minimal information chamber* (MIC) for X_i.

- $MIC = C_{-i} \cup \{X_i\}$: $C_{-i} \subset \boldsymbol{X}_{-i}$ is the *Markov boundary* for X_i, that is, the smallest subset of $\boldsymbol{X}_{-i}$ such that
$$\pi(X_i|\boldsymbol{X}_{-i}, \mathcal{F}) = \pi(X_i|C_{-i}, \mathcal{F}).$$
- MIC is the X_i's "information family"—knowing any one of them will provide information about X_i, in addition to public knowledge coded into $\mathcal{F}$.

In General, What Does DP Actually Guarantee?

A random statistic $T \in \mathbb{R}^d$ is pure DP with PLB ε and input premetric d_{Ham} if and only if for every prior π on $\boldsymbol{X}$, every sub-σ field $\mathcal{F}$ of the corresponding full σ-field σ_π, every $B \in \mathscr{B}\left(\mathbb{R}^d\right)$, every i, and every $A \in \mathscr{B}(\Theta_i)$ (where Θ_i is the state space of x_i), we have

$$e^{-c_i\varepsilon}\pi(X_i \in A \mid \mathcal{F}) \le \pi\left(X_i \in A \mid T \in B; \mathcal{F}\right) \le e^{c_i\varepsilon}\pi(X_i \in A \mid \mathcal{F}),$$

where c_i is the size of the *minimal information chamber* (MIC) for X_i.

- $MIC = C_{-i} \cup \{X_i\}$: $C_{-i} \subset \boldsymbol{X}_{-i}$ is the *Markov boundary* for X_i, that is, the smallest subset of $\boldsymbol{X}_{-i}$ such that
$$\pi(X_i|\boldsymbol{X}_{-i}, \mathcal{F}) = \pi(X_i|C_{-i}, \mathcal{F}).$$
- MIC is the X_i's “information family”—knowing any one of them will provide information about X_i, in addition to public knowledge coded into $\mathcal{F}$.
- Protecting *relative* risk against “strongest attacker" is the easiest—the more the attacker's prior information, the less left for protection.

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
    - Conceives of data privacy as robustness

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{P_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{\mathrm{P}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
   - Conceives of data privacy as a limit on *probabilistic inference*

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{P_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
   - Conceives of data privacy as a limit on *probabilistic inference*
   - Focus on two aspects of forward-looking harms: the probability and strength of an *inferential, individual-based* (IIB) *disclosure*

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{\mathrm{P}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
   - Conceives of data privacy as a limit on *probabilistic inference*
   - Focus on two aspects of forward-looking harms: the probability and strength of an *inferential, individual-based* (IIB) *disclosure*
   - Assumes a way to quantify IIB disclosures (e.g. via the protection loss random variable, aka the likelihood ratio)

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{\mathrm{P}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
   - Conceives of data privacy as a limit on *probabilistic inference*
   - Focus on two aspects of forward-looking harms: the probability and strength of an *inferential, individual-based* (IIB) *disclosure*
   - Assumes a way to quantify IIB disclosures (e.g. via the protection loss random variable, aka the likelihood ratio)
3. DP is not a holistic framework for assessing privacy

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
    - Conceives of data privacy as robustness
    - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{P_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
    - Conceives of data privacy as a limit on *probabilistic inference*
    - Focus on two aspects of forward-looking harms: the probability and strength of an *inferential, individual-based* (IIB) *disclosure*
    - Assumes a way to quantify IIB disclosures (e.g. via the protection loss random variable, aka the likelihood ratio)
3. DP is not a holistic framework for assessing privacy
    - The theory of DP brackets other privacy concerns

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{P_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
   - Conceives of data privacy as a limit on *probabilistic inference*
   - Focus on two aspects of forward-looking harms: the probability and strength of an *inferential, individual-based* (IIB) *disclosure*
   - Assumes a way to quantify IIB disclosures (e.g. via the protection loss random variable, aka the likelihood ratio)
3. DP is not a holistic framework for assessing privacy
   - The theory of DP brackets other privacy concerns
   - The practice of DP is often left stranded

DP's Framing of Data Privacy (Seeman and Susser, 2023)

1. DP is a condition on the statistic T:
   - Conceives of data privacy as robustness
   - Focuses on *forward-looking, individual-based harms*
2. (More exactly) DP is a restriction on the *data release model* $\{\mathrm{P}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}\}$
   - Conceives of data privacy as a limit on *probabilistic inference*
   - Focus on two aspects of forward-looking harms: the probability and strength of an *inferential, individual-based* (IIB) *disclosure*
   - Assumes a way to quantify IIB disclosures (e.g. via the protection loss random variable, aka the likelihood ratio)
3. DP is not a holistic framework for assessing privacy
   - The theory of DP brackets other privacy concerns
   - The practice of DP is often left stranded
   - DP needs to be integrated into broader theories of privacy (Benthall and Cummings, 2024)

References I

Abowd, J. M., Schneider, M. J., & Vilhuber, L. (2013). Differential privacy applications to Bayesian and linear mixed model estimation. *Journal of Privacy and Confidentiality*, *5*(1).

Andrés, M. E., Bordenabe, N. E., Chatzikokolakis, K., & Palamidessi, C. (2013). Geo-indistinguishability: Differential privacy for location-based systems. *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, 901–914. https://doi.org/10.1145/2508859.2516735

Ashmead, R., Kifer, D., Leclerc, P., Machanavajjhala, A., & Sexton, W. (2019). *Effective privacy after adjusting for invariants with applications to the 2020 Census* (tech. rep.).

Asi, H., Duchi, J. C., & Javidbakht, O. (2022). Element level differential privacy: The right granularity of privacy. *AAAI Workshop on Privacy-Preserving Artificial Intelligence.*

References II

Balle, B., Barthe, G., & Gaboardi, M. (2020). Privacy profiles and amplification by subsampling. *Journal of Privacy and Confidentiality*, *10*(1). https://doi.org/10.29012/jpc.726

Barber, R. F., & Duchi, J. C. (2014). Privacy and statistical risk: Formalisms and minimax bounds. https://doi.org/10.48550/arXiv.1412.4451

Barthe, G., & Olmedo, F. (2013). Beyond differential privacy: Composition theorems and relational logic for f-divergences between probabilistic programs. In F. V. Fomin, R. Freivalds, M. Kwiatkowska, & D. Peleg (Eds.), *Automata, languages, and programming* (pp. 49–60). Springer. https://doi.org/10.1007/978-3-642-39212-2_8

References III

Beimel, A., Kasiviswanathan, S. P., & Nissim, K. (2010). Bounds on the sample complexity for private learning and private data release. In D. Micciancio (Ed.), *Proceedings of the 7th theory of cryptography conference, TCC 2010, Zurich, Switzerland* (pp. 437–454). Springer. https://doi.org/10.1007/978-3-642-11799-2_26

Benthall, S., & Cummings, R. (2024). Integrating differential privacy and contextual integrity. https://doi.org/10.48550/arXiv.2401.15774

Bhaskar, R., Bhowmick, A., Goyal, V., Laxman, S., & Thakurta, A. (2011). Noiseless database privacy. In D. H. Lee & X. Wang (Eds.), *Advances in cryptology – ASIACRYPT 2011* (pp. 215–232). Springer. https://doi.org/10.1007/978-3-642-25385-0_12

Bun, M., Drechsler, J., Gaboardi, M., McMillan, A., & Sarathy, J. (2022). Controlling privacy loss in sampling schemes: An analysis of stratified and cluster sampling. *Foundations of Responsible Computing (FORC 2022)*, 24.

References IV

Bun, M., & Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In M. Hirt & A. Smith (Eds.), *Theory of cryptography* (pp. 635–658). Springer. https://doi.org/10.1007/978-3-662-53641-4_24

Charest, A.-S., & Hou, Y. (2016). On the meaning and limits of empirical differential privacy. *Journal of Privacy and Confidentiality*, *7*(3), 53–66.

Chatzikokolakis, K., Andrés, M. E., Bordenabe, N. E., & Palamidessi, C. (2013). Broadening the Scope of Differential Privacy Using Metrics. In E. De Cristofaro & M. Wright (Eds.), *Privacy Enhancing Technologies* (pp. 82–102). Springer. https://doi.org/10.1007/978-3-642-39077-7_5

Dalenius, T. E. (1977). Towards a methodology for statistical disclosure control. *Statistisk tidskrift*, *15*, 429–444.

References V

Dharangutte, P., Gao, J., Gong, R., & Yu, F.-Y. (2023). Integer subspace differential privacy. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-23).*

Dong, J., Roth, A., & Su, W. J. (2022). Gaussian differential privacy. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *84*(1), 3–37. https://doi.org/10.1111/rssb.12454

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., & Naor, M. (2006). Our data, ourselves: Privacy via distributed noise generation. In S. Vaudenay (Ed.), *Advances in cryptology - EUROCRYPT 2006* (pp. 486–503). Springer. https://doi.org/10.1007/11761679_29

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. *Theory of cryptography conference*, 265–284.

References VI

Dwork, C., Naor, M., Pitassi, T., & Rothblum, G. N. (2010). Differential privacy under continual observation [https://dl.acm.org/doi/10.1145/1806689.1806787]. *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, 715–724. https://doi.org/10.1145/1806689.1806787

Ebadi, H., Sands, D., & Schneider, G. (2015). Differential Privacy: Now it's Getting Personal. *ACM SIGPLAN Notices*, *50*(1), 69–81. https://doi.org/10.1145/2775051.2677005

Feldman, V., & Zrnic, T. (2022). Individual privacy accounting via a Rényi filter.

Gao, J., Gong, R., & Yu, F.-Y. (2022). Subspace differential privacy. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(4), 3986–3995. https://doi.org/10.1609/aaai.v36i4.20315

Gong, R., & Meng, X.-L. (2020). Congenial differential privacy under mandated disclosure. *Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference*, 59–70.

References VII

- Hay, M., Li, C., Miklau, G., & Jensen, D. (2009). Accurate estimation of the degree distribution of private networks. *2009 Ninth IEEE International Conference on Data Mining*, 169–178. https://doi.org/10.1109/ICDM.2009.11
- He, X., Machanavajjhala, A., & Ding, B. (2014). Blowfish privacy: Tuning privacy-utility trade-offs using policies. *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, 1447–1458.
- Jorgensen, Z., Yu, T., & Cormode, G. (2015). Conservative or liberal? Personalized differential privacy [https://ieeexplore.ieee.org/document/7113353]. *2015 IEEE 31st International Conference on Data Engineering*, 1023–1034. https://doi.org/10.1109/ICDE.2015.7113353
- Kifer, D., & Machanavajjhala, A. (2011a). No free lunch in data privacy. *Proceedings of the 2011 International Conference on Management of Data - SIGMOD '11*, 193–204. https://doi.org/10.1145/1989323.1989345

References VIII

- Kifer, D., & Machanavajjhala, A. (2011b). No free lunch in data privacy. *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, 193–204.
- Kifer, D., & Machanavajjhala, A. (2012). A rigorous and customizable framework for privacy. *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI symposium on Principles of Database Systems*, 77–88.
- Kifer, D., & Machanavajjhala, A. (2014). Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, *39*(1), 1–36.
- McSherry, F., & Mahajan, R. (2010). Differentially-private network trace analysis. *Proceedings of the ACM SIGCOMM 2010 Conference*, 123–134. https://doi.org/10.1145/1851182.1851199
- Mironov, I. (2017). Rényi differential privacy. *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, 263–275. https://doi.org/10.1109/CSF.2017.11

References IX

- O'Keefe, C. M., & Charest, A.-S. (2019). Bootstrap differential privacy. *Transactions on Data Privacy*, *12*, 1–28.
- Redberg, R., & Wang, Y.-X. (2021). Privately publishable per-instance privacy. *Advances in Neural Information Processing Systems*, *34*, 17335–17346.
- Seeman, J., Reimherr, M., & Slavkovic, A. (2022). Formal privacy for partially private data.
- Seeman, J., Sexton, W., Pujol, D., & Machanavajjhala, A. (2024). Privately answering queries on skewed data via per record differential privacy. *Proceedings of the VLDB Endowment*, *17*(11), 3138–3150. https://doi.org/10.14778/3681954.3681989
- Seeman, J., & Susser, D. (2023). Between privacy and utility: On differential privacy in theory and practice [https://dl.acm.org/doi/10.1145/3626494]. *ACM Journal on Responsible Computing*. https://doi.org/10.1145/3626494

References X

Soria-Comas, J., Domingo-Ferrer, J., Sánchez, D., & Megías, D. (2017). Individual differential privacy: A utility-preserving formulation of differential privacy guarantees. *IEEE Transactions on Information Forensics and Security*, *12*(6), 1418–1429. https://doi.org/10.1109/TIFS.2017.2663337

Tschantz, M. C., Sen, S., & Datta, A. (2020). SoK: Differential privacy as a causal property. *2020 IEEE Symposium on Security and Privacy (SP)*, 354–371.

Wang, Y.-X. (2018). Per-instance Differential Privacy.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, *60*(309), 63–69.

Zhou, S., Ligett, K., & Wasserman, L. (2009). Differential privacy with compression. *Proceedings of the 2009 IEEE International Conference on Symposium on Information Theory - Volume 4*, 2718–2722.