

Theory of Differential Privacy

EN.601.438/638, Fall 2026

Problem Set 1: DP Fundamentals

100 points • 20% of course grade • Due: October 5, 23:59

Submission. Submit one PDF on Gradescope. Include your name and list, for each problem, any collaborators and outside resources or tools you used. Unless a problem explicitly asks only for discussion, justify every answer with a proof.

Collaboration. You may discuss definitions, concepts, and high-level ideas with classmates and course staff. You may not view another student's written solution, jointly compose a proof, or reuse text. After a discussion, put away your notes and write your solution independently from scratch.

Generative AI. You may use an AI tool as a peer to clarify definitions, request a new example, or quiz yourself. Do not paste the assignment questions into a tool, request a solution or solution outline, keep the assignment open while using the tool, or submit AI-generated reasoning. Close the tool before writing your solution, and disclose how you used it. Some problems or closely related exercises may appear in the online materials for the BU/NEU or Waterloo courses; you may not consult or copy solutions, if any, from those courses.

Conventions. Unless stated otherwise, datasets have fixed size and neighboring datasets differ in one row (replace-one adjacency). All logarithms are natural. Privacy and failure parameters such as δ and β lie in $(0, 1)$ unless stated otherwise. For a real number u , we write $(u)_+ = \max\{u, 0\}$. You may use results proved in lecture, including post-processing, adaptive composition, the Laplace and Gaussian mechanisms, and the privacy and utility theorems for the Exponential Mechanism.

Facts you may use

- If $Z \sim \text{Lap}(b)$, then for every $t \geq 0$,

$$\Pr(|Z| \geq t) = e^{-t/b}, \quad \Pr(Z \geq t) = \Pr(Z \leq -t) = \frac{1}{2}e^{-t/b}.$$

- If $G \sim N(0, 1)$, then $\mathbb{E}[e^{\lambda G}] = e^{\lambda^2/2}$ for every $\lambda \in \mathbb{R}$.
- Basic adaptive composition for approximate DP: composing mechanisms with parameters $(\varepsilon_t, \delta_t)$ gives $(\sum_t \varepsilon_t, \sum_t \delta_t)$ -DP.
- Gaussian mechanism theorem from lecture: if f has ℓ_2 -sensitivity at most Δ , then

$$f(x) + N(0, \sigma^2 I)$$

is (ε, δ) -DP for

$$\sigma \geq \frac{\Delta \sqrt{2 \log(2/\delta)}}{\varepsilon}.$$

[18 points]

Problem 1. Privacy calculus and privacy loss

(a) **Parallel composition.** [4 points] Let $[n] = I_1 \sqcup \dots \sqcup I_m$ be a fixed, public partition of the row indices. For every $j \in [m]$, let

$$M_j : \mathcal{X}^{I_j} \rightarrow \mathcal{Y}_j$$

be ε_j -DP with respect to replacing one row inside block I_j . Assume the mechanisms use independent randomness, and define

$$M(x) = (M_1(x_{I_1}), \dots, M_m(x_{I_m})).$$

Prove that M is $\max_j \varepsilon_j$ -DP.

(b) **Private and nonprivate data-dependent selection.** [5 points] A biologist wants to release correlations among ten diseases selected from a list of d diseases.

- (i) She first uses an ε -DP mechanism A_1 to release approximate disease frequencies. She selects the ten diseases with the largest *released* frequencies, and then applies an ε -DP mechanism A_2 to the original dataset to release correlations among the selected diseases. Assume that, for every fixed set of ten diseases, the corresponding instance of A_2 is ε -DP. Prove that releasing both stages is 2ε -DP.
- (ii) She instead selects the ten diseases with the largest *true* frequencies in the original dataset, and then applies an ε -DP mechanism for the selected correlations. Explain why privacy does not follow from the fact that the second-stage mechanism is private for every fixed set of diseases. Give an explicit counterexample in which every fixed-choice mechanism is 0-DP but the resulting data-dependent procedure reveals the private input exactly.

(c) **Exact characterization by the privacy-loss random variable.** [9 points] Let P and Q be probability distributions on a common measurable space, with densities p and q with respect to a common dominating measure. Define

$$L_{P\|Q}(y) = \log \frac{p(y)}{q(y)},$$

using the usual extended-real conventions when one density is zero, and define the one-sided privacy deficit

$$\delta_\varepsilon(P\|Q) = \sup_S (P(S) - e^\varepsilon Q(S)),$$

where the supremum is over measurable events S .

- (i) Prove the identity

$$\delta_\varepsilon(P\|Q) = \mathbb{E}_{Y \sim P} \left[\left(1 - e^{\varepsilon - L_{P\|Q}(Y)} \right)_+ \right].$$

Identify an event attaining the supremum.

- (ii) Deduce that

$$\Pr_{Y \sim P} [L_{P\|Q}(Y) > \varepsilon] \leq \delta \implies \delta_\varepsilon(P\|Q) \leq \delta.$$

- (iii) **A partial converse, necessarily with slack.** The implication in part (ii) cannot in general be reversed at the same threshold ε . Show instead that if $\delta_\varepsilon(P\|Q) \leq \delta$, then, for every $t > 0$,

$$\Pr_{Y \sim P} [L_{P\|Q}(Y) > \varepsilon + t] \leq \frac{\delta}{1 - e^{-t}}.$$

- (iv) **What is equivalent to approximate DP?** Show that a mechanism A is (ϵ, δ) -DP if and only if

$$\delta_\epsilon(A(x) \| A(x')) \leq \delta$$

for every ordered neighboring pair (x, x') . This privacy-deficit condition is the exact characterization. In contrast, the unshifted PLRV tail condition in part (ii) is sufficient but not necessary. Verify this using the following two-point example: for $0 < \delta < 1$, take

$$P = \left(\frac{1}{2} + \frac{\delta}{2}, \frac{1}{2} - \frac{\delta}{2}\right), \quad Q = \left(\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\right),$$

and compare $\delta_0(P \| Q)$ with $\Pr_{Y \sim P}[L_{P \| Q}(Y) > 0]$.

[18 points]

Problem 2. Differential privacy at the limit of reconstruction

Let

$$X \sim \text{Unif}(\{0, 1\}^n),$$

let $A : \{0, 1\}^n \rightarrow \mathcal{Y}$ be ϵ -DP, and assume for simplicity that $\mathcal{Y}$ is finite. A possibly randomized reconstruction algorithm B observes only $A(X)$ and outputs $\hat{X} \in \{0, 1\}^n$.

(a) **Posterior odds.** [7 points] For every coordinate $i \in [n]$ and every transcript y with positive probability, prove

$$e^{-\epsilon} \leq \frac{\Pr[X_i = 1 \mid A(X) = y]}{\Pr[X_i = 0 \mid A(X) = y]} \leq e^\epsilon.$$

You may find it useful to write, for $b \in \{0, 1\}$,

$$\Pr[A(X) = y \mid X_i = b] = 2^{-(n-1)} \sum_{u \in \{0, 1\}^{n-1}} \Pr[A(b, u) = y].$$

(b) **Reconstruction lower bound.** [7 points] Use part (a) and linearity of expectation to prove

$$\mathbb{E}[d_H(X, \hat{X})] \geq \frac{n}{1 + e^\epsilon},$$

where the expectation is over X , the randomness of A , and the randomness of B .

(c) **Tightness.** [4 points] Define randomized response on one bit by

$$\text{RR}_\epsilon(b) = \begin{cases} b, & \text{with probability } \frac{e^\epsilon}{1 + e^\epsilon}, \\ 1 - b, & \text{with probability } \frac{1}{1 + e^\epsilon}. \end{cases}$$

Apply this mechanism independently to all coordinates and output

$$A_{\text{RR}}(x) = (\text{RR}_\epsilon(x_1), \dots, \text{RR}_\epsilon(x_n)).$$

Prove that A_{RR} is ϵ -DP, rather than $n\epsilon$ -DP, and show that the decoder $\hat{X} = A_{\text{RR}}(X)$ attains equality in part (b).

[22 points]

Problem 3. A first privacy accountant: adaptive Gaussian composition

A mechanism interacts with an analyst for k rounds. At round t , after observing the previous outputs

$$y_{<t} = (y_1, \dots, y_{t-1}),$$

the analyst chooses a function

$$f_t^{y_{<t}} : \mathcal{X}^n \rightarrow \mathbb{R}$$

whose global sensitivity is at most Δ_t . The mechanism releases

$$Y_t = f_t^{Y_{<t}}(x) + Z_t, \quad Z_t \sim N(0, \sigma_t^2),$$

where the Z_t are mutually independent and independent of the past. The query choice may be fully adaptive, but Δ_t and σ_t are public.

Fix neighboring datasets $x \sim x'$. Let P and Q be the distributions of the full transcript $Y = (Y_1, \dots, Y_k)$ under x and x' , respectively.

(a) Chain rule for transcript privacy loss. [3 points] Let $p_x(y_t | y_{<t})$ denote the conditional density of the round- t output under dataset x , and define

$$L_t = \log \frac{p_x(Y_t | Y_{<t})}{p_{x'}(Y_t | Y_{<t})}.$$

Prove

$$L_{P\|Q}(Y) = \sum_{t=1}^k L_t.$$

Explain why, after conditioning on the same history $Y_{<t} = y_{<t}$, the same adaptively chosen query $f_t^{y_{<t}}$ appears in the numerator and denominator.

(b) Conditional privacy loss of one Gaussian release. [4 points] For a fixed history $y_{<t}$, write

$$d_t = f_t^{y_{<t}}(x) - f_t^{y_{<t}}(x'), \quad |d_t| \leq \Delta_t.$$

Under P , prove that, conditional on $Y_{<t}$,

$$L_t = \frac{d_t^2}{2\sigma_t^2} + \frac{d_t}{\sigma_t} G_t,$$

where $G_t \sim N(0, 1)$ is fresh randomness independent of the past. Deduce the conditional mean and variance of L_t .

(c) Composition by a conditional moment bound. [7 points] Define

$$\rho_t = \frac{\Delta_t^2}{2\sigma_t^2}, \quad \rho = \sum_{t=1}^k \rho_t.$$

Prove that for every $\lambda > 0$,

$$\mathbb{E}_P \left[\exp \left(\lambda \sum_{t=1}^k \left(L_t - \frac{d_t^2}{2\sigma_t^2} \right) \right) \right] \leq e^{\lambda^2 \rho}.$$

Hint. Condition on the final history $Y_{<k}$, apply the Gaussian moment-generating-function identity to the k th factor, and then iterate the tower property backward through the rounds.

Use this bound, Markov's inequality, and an optimized choice of λ to prove

$$\Pr_{Y \sim P} \left[L_{P \parallel Q}(Y) > \rho + 2\sqrt{\rho \log(1/\delta)} \right] \leq \delta.$$

Repeat the argument with x and x' interchanged, and use Problem 1(c) to conclude that the full adaptive transcript is

$$\left(\rho + 2\sqrt{\rho \log(1/\delta)}, \delta \right)\text{-DP}.$$

(d) Basic composition versus the privacy-loss accountant. [6 points] Suppose every query has sensitivity 1 and every release uses the same standard deviation σ .

- (i) If each round is separately calibrated to be $(\varepsilon/k, \delta/k)$ -DP and one applies basic composition, show that it is sufficient to take

$$\sigma_{\text{basic}} = \frac{k\sqrt{2 \log(2k/\delta)}}{\varepsilon}.$$

- (ii) Using part (c), show that the composed privacy parameter is

$$\varepsilon_{\text{acct}}(\sigma) = \frac{k}{2\sigma^2} + \frac{\sqrt{2k \log(1/\delta)}}{\sigma}.$$

Assuming $0 < \varepsilon \leq 1$ and $0 < \delta \leq e^{-1}$, prove that

$$\sigma = \frac{2\sqrt{2k \log(1/\delta)}}{\varepsilon}$$

is sufficient for (ε, δ) -DP.

- (iii) Compare the dependence on k in the two calibrations.

(e) Sequential inside groups, parallel across groups. [2 points] Suppose a fixed public partition divides the rows into m disjoint groups. On group j , an adaptive Gaussian protocol has total parameter

$$\rho_j = \frac{1}{2} \sum_t \frac{\Delta_{j,t}^2}{\sigma_{j,t}^2}.$$

All group transcripts are released, and the group protocols use independent randomness. Prove that the overall mechanism is

$$\left(\max_{j \in [m]} \left\{ \rho_j + 2\sqrt{\rho_j \log(1/\delta)} \right\}, \delta \right)\text{-DP}.$$

[22 points]

Problem 4. Private localization and mean estimation

For a dataset $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, write

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

(a) Why a boundedness assumption is necessary. [4 points] Prove that, for every finite $t \geq 0$, there is no 1-DP mechanism $M : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying

$$\Pr[|M(x) - \bar{x}| \leq t] \geq 0.9$$

for every dataset $x \in \mathbb{R}^n$.

(b) Mean estimation in a known interval. [4 points] Let $J = [a, b]$ be a public interval of width $W = b - a$, and define

$$\Pi_J(z) = \min\{\max\{z, a\}, b\}.$$

For a privacy parameter $\varepsilon_{\text{mean}} > 0$, consider

$$A_J(x) = \frac{1}{n} \sum_{i=1}^n \Pi_J(x_i) + Z, \quad Z \sim \text{Lap}\left(\frac{W}{n\varepsilon_{\text{mean}}}\right).$$

Prove that A_J is $\varepsilon_{\text{mean}}$ -DP on every dataset in $\mathbb{R}^n$. If all $x_i \in J$, prove that, with probability at least $1 - \beta$,

$$|A_J(x) - \bar{x}| \leq \frac{W}{n\varepsilon_{\text{mean}}} \log \frac{1}{\beta}.$$

(c) Privately locate an unknown short interval. [7 points] Let $R \geq 1$ be an integer known to the algorithm, and suppose the data lie in $[-R, R]$. Partition this range into the $2R$ unit bins

$$B_j = [j, j+1), \quad j = -R, -R+1, \dots, R-1,$$

and define the count score

$$c_j(x) = \#\{i : x_i \in B_j\}.$$

Assume there exists an unknown interval $I \subseteq [-R, R]$ of width at most 2 containing every data point. Use the Exponential Mechanism with privacy parameter ε_{loc} and score $c_j(x)$ to choose an index $\hat{j}$. Prove all of the following:

- (i) Some unit bin contains at least $n/3$ data points.
- (ii) Each score c_j has sensitivity 1.
- (iii) If

$$n > \frac{6}{\varepsilon_{\text{loc}}} \log \frac{2R}{\beta_{\text{loc}}},$$

then with probability at least $1 - \beta_{\text{loc}}$, the selected bin $B_{\hat{j}}$ contains at least one data point.

- (iv) On this event, the interval

$$J_{\hat{j}} = [\hat{j} - 2, \hat{j} + 3]$$

has width 5 and contains I .

Notice that your algorithm is private on every input, as the DP definition requires, but its accuracy guarantee is conditional on the assumption of the existence of the promised width-2 interval I .

(d) Compose localization and estimation. [7 points] Choose privacy parameters $\varepsilon_{\text{loc}}, \varepsilon_{\text{mean}} > 0$ satisfying

$$\varepsilon_{\text{loc}} + \varepsilon_{\text{mean}} = \varepsilon.$$

First run the localization procedure from part (c), and then, conditional on its chosen interval $J_{\hat{j}}$, run the clipped-mean mechanism from part (b). The final algorithm may output only the estimated mean.

Prove that the resulting algorithm is ε -DP on every dataset. Under the width-2 promise, prove that if

$$n > \frac{6}{\varepsilon_{\text{loc}}} \log \frac{4R}{\beta},$$

then, with probability at least $1 - \beta$,

$$|\hat{\mu} - \bar{x}| \leq \frac{5}{n\varepsilon_{\text{mean}}} \log \frac{2}{\beta}.$$

Finally, state the resulting sample-size condition and error bound for the even split

$$\varepsilon_{\text{loc}} = \varepsilon_{\text{mean}} = \varepsilon/2.$$

[20 points]

Problem 5. Private medians: worst-case guarantees and stable instances

Let $R \geq 2$, let $n = 2r + 1$ be odd, and let $x = (x_1, \dots, x_n) \in [R]^n$. Write

$$x_{(1)} \leq \dots \leq x_{(n)} \quad \text{and} \quad \text{med}(x) = x_{(r+1)}.$$

For $y \in [R]$, define

$$L_x(y) = \#\{i : x_i < y\}, \quad R_x(y) = \#\{i : x_i > y\},$$

and the rank violation

$$v_x(y) = \max\{L_x(y) - r, R_x(y) - r, 0\}.$$

(a) Median via the Exponential Mechanism. [7 points] Let $q(x, y) = -v_x(y)$. Prove that q has global sensitivity 1 under replace-one adjacency, and that $v_x(y) = 0$ if and only if $y = \text{med}(x)$. Consider the Exponential Mechanism that samples $Y \in [R]$ with probability proportional to

$$\exp\left(\frac{\varepsilon q(x, Y)}{2}\right).$$

Conclude that it is ε -DP, and prove that with probability at least $1 - \beta$,

$$v_x(Y) \leq \frac{2}{\varepsilon} \left(\log R + \log \frac{1}{\beta} \right).$$

(b) Distance to instability. [5 points] Define

$$D(x) = \min\{d_H(x, x') : \text{med}(x') \neq \text{med}(x)\}.$$

Prove that D has global sensitivity 1. Then let

$$v = \text{med}(x), \quad \ell = \#\{i : x_i < v\}, \quad g = \#\{i : x_i > v\}.$$

Prove the explicit formula

$$D(x) = \min\{r + 1 - \ell, r + 1 - g\}.$$

(c) Propose-Test-Release for the median. [8 points] Fix $0 < \delta < 1/2$ and define

$$T = 1 + \frac{1}{\varepsilon} \log \frac{1}{2\delta}.$$

The mechanism draws $Z \sim \text{Lap}(1/\varepsilon)$, forms

$$\tilde{D} = D(x) + Z,$$

and outputs

$$M_{\text{PTR}}(x) = \begin{cases} \text{med}(x), & \tilde{D} > T, \\ \perp, & \tilde{D} \leq T. \end{cases}$$

Prove that M_{PTR} is (ε, δ) -DP.

(d) Stability-dependent exact recovery. [0 points]

- (i) Prove that for every $y \neq \text{med}(x)$,

$$v_x(y) \geq D(x).$$

Use this fact to show that the Exponential Mechanism from part (a) satisfies

$$\Pr[Y \neq \text{med}(x)] \leq (R - 1) \exp\left(-\frac{\varepsilon D(x)}{2}\right).$$

Conclude that it returns the exact median with probability at least $1 - \beta$ whenever

$$D(x) \geq \frac{2}{\varepsilon} \log \frac{R - 1}{\beta}.$$

- (ii) Show that the PTR mechanism from part (c) has the following properties:

- (a) If $D(x) = 1$, then

$$\Pr[M_{\text{PTR}}(x) = \text{med}(x)] = \delta, \quad \Pr[M_{\text{PTR}}(x) = \perp] = 1 - \delta.$$

- (b) If

$$D(x) \geq 1 + \frac{1}{\varepsilon} \log \frac{1}{4\delta\beta},$$

then PTR returns the exact median with probability at least $1 - \beta$.

- (iii) Briefly compare the two guarantees. Consider when an algorithm does not return an answer and how often it returns the exact median. For a concrete comparison, set

$$\varepsilon = 4 \log 10, \quad \delta = 2.5 \cdot 10^{-6}, \quad \beta = 0.1, \quad n = 13, \quad R = 10^6, \quad m = 500000.$$

Which algorithm would you choose if you knew the instance were

$$x = (1, 2, 3, 4, \underbrace{m, m, m, m, m}_{5 \text{ copies}}, R - 3, R - 2, R - 1, R)$$

versus

$$x' = (m - 6, m - 5, \dots, m + 5, m + 6)?$$

For each instance, calculate D and bound the exact-recovery probability of each mechanism. For the Exponential Mechanism on x' , use its exact output probabilities rather than only the worst-case bound from part (i). You may use the identity $e^{-\varepsilon j/2} = 10^{-2j}$ for integer $j \geq 0$.

- (e) **Median via Above Threshold.** [0 points] For $y \in [R]$, define the prefix-count query

$$c_y(x) = \#\{i : x_i \leq y\}.$$

Run the Above Threshold algorithm from lecture on $c_1, c_2, \dots, c_R$ with threshold $r + 1$ and total privacy parameter ε , returning the first index declared above threshold. Use the version in which the threshold receives $\text{Lap}(2/\varepsilon)$ noise and each query receives $\text{Lap}(4/\varepsilon)$ noise.

Prove that the mechanism is ε -DP. Assuming

$$r \geq \frac{2}{\varepsilon} \log \frac{2}{\beta} + \frac{4}{\varepsilon} \log \frac{2R}{\beta},$$

prove that, with probability at least $1 - \beta$, it returns some $Y \in [R]$ satisfying

$$v_x(Y) \leq 1 + \frac{2}{\varepsilon} \log \frac{2}{\beta} + \frac{4}{\varepsilon} \log \frac{2R}{\beta}.$$

Compare this rank-error guarantee with the guarantee for the Exponential Mechanism in part (a).