HW6: Variants of Sensitivity and Data-Dependent Bounds

CS 208 Applied Privacy for Data Science, Spring 2022

Version 1.2: Due Fri, Mar. 11, 5:00pm.

Instructions: Submit a single PDF file to Gradescope containing your solutions, code, plots, and analyses. Make sure to list all collaborators and references.

1. **Graph Privacy and Different Types of Sensitivity:** For $n \geq 2$, let $\mathcal{G} =$ the set of undirected graphs (without self-loops) on vertex set $V = \{1, \ldots, n\}$, and for $G, G' \in \mathcal{G}$, define $G \sim G'$ if there is a vertex $v \in V$ such that the only differences between G and G' involve edges incident to the vertex v. (That is, we are considering node-level privacy.) For an integer $d \in [2, n-1]$, let $\mathcal{H} \subseteq \mathcal{G}$ denote the set of graphs of degree at most d. Define $q : \mathcal{G} \to \mathbb{N}$ by taking $q(G)$ to be the number of isolated (i.e., degree 0) nodes in G. Calculate the following measures of sensitivity of q:

   (a) The global sensitivity: GS_q.

   (b) The minimum local sensitivity: $\min_{G \in \mathcal{G}} \mathrm{LS}_q(G)$. (Some of the approaches we mentioned in class, like Privately Bounding Local Sensitivity, Propose-Test-Release, and Smooth Sensitivity aim to add noise that's not too much larger than the local sensitivity, which can sometimes be much smaller than the global sensitivity. It's not always possible to do this while preserving DP, but local sensitivity calculations like here and below help give a sense of how much we can gain from such methods.)

   (c) The maximum local sensitivity on $\mathcal{H}$: $\max_{G \in \mathcal{H}} \mathrm{LS}_q(G)$. [1]

   (d) The restricted sensitivity on $\mathcal{H}$: $\mathrm{RS}_q^{\mathcal{H}} = \max_{G, G' \in \mathcal{H}, G \sim G'} |q(G) - q(G')|$.[2] (The material we surveyed on graph privacy and restricted sensitivity tells us that there is a mechanism that is ϵ-DP on all of $\mathcal{G}$, but only adds noise proportional to $\mathrm{RS}_q^{\mathcal{H}}$ for graphs in $\mathcal{H}$.)

2. **Data-Dependent Clipping Bounds:** In all of the parts below, the dataset is $x \in [0, B]^n$. In all of the implementation parts, you should write code that takes as input $B \geq 0$, $n \in \mathbb{N}$, $x \in [0, B]^n$, and $\varepsilon > 0$.

   (a) Show that the following algorithm for estimating a Trimmed mean is ε-DP:

$$M(x) = \frac{1}{.9n} \cdot \left(\sum_{\lfloor .05n \rceil \leq i \leq \lfloor .95n \rceil} \mathrm{sort}(x)_i \right) + \mathrm{Lap}\left(\frac{B}{0.9\varepsilon n} \right),$$

   where $\mathrm{sort}(x)$ is a sorting of x and $\lfloor z \rceil$ denotes the nearest integer to z (breaking ties by rounding down). That is, we are applying the Laplace mechanism after removing the bottom and top 5% of the dataset. (Hint: Think about how changing one row of x affects the trimmed dataset.)

[1] The answer differs from and motivates why we use restricted sensitivity.

[2] The general definition of restricted sensitivity is a bit more involved, and also considers datasets G and G' that are not neighbors, but this simplified version is equivalent in the special case of $\mathcal{H}$ and $\sim$ considered here.

(b) Show that for large enough n, the analogous algorithm for the *Winsorized* mean is *not* ε-DP:

$$M(x) = \frac{1}{n} \cdot \sum_{i=1}^{n} [x_i]_{P_{.05}}^{P_{.95}} + \text{Lap}\left(\frac{B}{\varepsilon n}\right),$$

where $P_t = \text{sort}(x)_{\lfloor tn \rceil}$ is the t'th percentile of x and $[x]_a^b$ is defined as in HW3. In Winsorization (which you saw in the Opportunity Insights Application), we clamp points rather than drop them. To prove that the mechanism is not ε-DP, you should exhibit two adjacent datasets $x, x' \in [0, B]^n$ for which the distributions of $M(x)$ and $M(x')$ are not within an e^{ϵ} factor of each other.

(c) In HW5, you implemented a continuous version of the exponential mechanism for releasing a median. Describe and implement a continuous version of the exponential mechanism for releasing an estimate of the tth percentile P_t of a dataset $x \in [0, B]^n$ for any desired $t \in [0, 100]$. Your function should take t as an input.

(d) Consider the following algorithm for estimating a Winsorized mean of a dataset: use your algorithm from Part 2c to get $\varepsilon/3$-DP estimates $\hat{P}_{.05}$ and $\hat{P}_{.95}$ of the 5th and 95th percentiles, and output

$$M(x) = \frac{1}{n} \cdot \left(\sum_{i=1}^{n} [x_i]_{\hat{P}_{.05}}^{\hat{P}_{.95}}\right) + \text{Lap}\left(\frac{3(\hat{P}_{.95} - \hat{P}_{.05})}{\varepsilon n}\right).$$

What DP properties does M use that makes it ε-DP, even though the algorithm in Part 2b is not?

(e) The dataset `FultonPUMS5full.csv` provides the 5% PUMS Census file for Fulton County, Georgia. For $\varepsilon = 1$ and each $B \in \{5 \times 10^5, 5 \times 10^6, 5 \times 10^7\}$, estimate the RMSE of DP mean income for each PUMA in Fulton County.[3] Run this analysis to compare (i) the ordinary Laplace mechanism for a mean and (ii) the algorithm from Part 2d. Show box-and-whisker plots of the DP mean incomes for each PUMA and algorithm, noting the true means. (In the GitHub repo, we have given you `hw6_starter.py` for producing such plots comparing the winsorized mean algorithm from Part 2d to the ordinary Laplace mechanism.) Order PUMA by mean income, or perhaps skew of income, or anything you think reveals an interesting pattern. Give an intuitive explanation of the cases (datasets and parameter settings) in which algorithm (i) performs better than algorithm (ii) and vice-versa.

3. **Participation Highlights:** Recall that participation is an important part of CS208, and that there are many ways in which you can participate (Perusall, Ed, in-class discussions, section, collaboration with classmates, etc.) To help us in assessing participation, please share with us brief descriptions of up to 3 highlights of your participation in the class so far (up to Friday March 5), including a reflection on how each contributed positively to the learning environment of the class. Also briefly discuss any adjustments you intend to make to your participation in the second half of the course.

[3] You can assume that the size of each PUMA dataset is public information.