HW 8b: DP Wizard and Local Model

CS 208 Applied Privacy for Data Science, Spring 2025

Version 1.1: Due Fri, Apr. 18th, 11:59pm.

Instructions: Submit a PDF file that contains both your written responses as well as your code to the assignment on Gradescope. Read the section "Collaboration & AI Policy" in the syllabus for our guidelines regarding the use of LLMs and other AI assistance on the assignments.

1. **DP-Wizard:** Install DP Wizard on your machine following these instructions.

   Spend some time getting familiar with it. You may want to setup a clean python virtual environment for the installation (see Section 0 notes on Canvas for instructions on setting up virtual environments).

   Note that DP-Wizard is still in-progress and actively being developed. DP Wizard was designed to enable analysts without DP expertise to compute simple differentially private statistics. Assume the following design goals: DP Wizard should help a user: (1) compute simple differentially private statistics without significant involvement from experts in data privacy and (2) understand impacts of lower versus higher values of ϵ on accuracy. In this problem, you will design and run a user study on DP Wizard, bearing in mind its design goals.

   (a) Reflect on the design goals and design a user study to evaluate how well DP Wizard achieves these goals. Write two tasks participants in your study will complete. Each task should correspond to one of the design goals. In addition, write at least two follow-up questions that participants will answer after using DP Wizard. These questions should be open-ended and allow participants to reflect on their experience using the tool. You might consider asking them about challenges they faced, or aspects of the tool they found easy to use. Note: Your response should contain the two tasks, a brief explanation of how each helps test one design goal, two follow-up questions, and a brief explanation for how each follow-up question helps better identify where the tool can be improved.

   (b) Run your user study with one person who does not have DP background but is familiar with quantitative data analysis. Begin by introducing them to DP and walking them through the tool. (Write out a script for the introduction to DP and walk-through of the tool, and include it in your response to this question. Your script should explain DP concepts that your participant will need in order to use DP Wizard to complete your tasks.) As they complete the tasks, ask them to "think aloud" by vocalizing the narrator voice in their head. Describe results from your study, including their responses to the task, how they completed the tasks (e.g., which aspects they struggled with or found straightforward), and their responses to the open-ended interview questions.

   (c) Describe at least two design changes you would recommend based on the results of your study. Provide rationale for each based on the data you collected from your one-person user study. (In practice, you would collect data from more than one participant and look for themes in results to inform design changes. However, for course purposes, it's fine to make some recommendations based on one person's feedback.)

(d) In this exercise, you were not required to include a control condition in your user study. However, in practice, seeing how well potential users use your tool compared to a meaningful baseline can be useful for contextualizing your results. Briefly describe a potential control that you could include in this study, and rationale for why it provides a meaningful baseline.

(e) As previously stated, DP Wizard is an ongoing project. Your responses to this question could help the OpenDP team improve DP Wizard. Would you like to opt-in to having your responses to this question shared with the OpenDP team and others working to improve DP Wizard? You are under no obligation to opt-in, and your decision will not impact your grade in any way.

2. **Statistical Inference in the Local and Shuffle Models:** In class on 3/10, in the context of the Opportunity Atlas application, we discussed the importance of providing uncertainty estimates (like confidence intervals) that take into account the randomness of both the DP mechanism and the data-generating process. This can be quite tricky and consume extra privacy-loss budget for many central DP mechanisms, like DP analogues of OLS regression. This is what motivated Chetty and Friedman to develop their MOS method, which is not strictly DP.

   However, for the local model and shuffle model, providing such uncertainty estimates is much easier, because local randomizers preserve independence. In this problem, you will see how this works for the special case of randomized response on Bernoulli data.

   Suppose we have a dataset of Boolean values $x_1, \ldots, x_n \in \{0, 1\}$ sampled independently from $\text{Ber}(p)$ for a parameter $p \in [0, 1]$. Without privacy, our best estimate for p is the empirical proportion $\hat{p} = (\sum_i x_i)/n$. But we'd like to accompany this point estimate with some measure of uncertainty. Given $x = (x_1, \ldots, x_n)$ and a parameter $\alpha \in [0, 1]$ (e.g. $\alpha = .05$), there are many methods to calculate a *confidence interval* $I = I_\alpha(x) = [\ell, u] \subseteq [0, 1]$ for p. The desired properties of a confidence interval are *coverage*: for every p and n,

$$\Pr_{x_1,\ldots,x_n \sim \text{Ber}(p)}[p \in I_\alpha(x)] \geq 1 - \alpha,$$

   and *small width*: namely that $|I_\alpha(x)| = u - \ell$ is as "small as possible" subject to the coverage constraint.

   The *Wald Interval*, uses the normal approximation and sets

$$I_\alpha(x) = \left[\hat{p} - z_\alpha \sqrt{n\hat{p}(1-\hat{p})}, \hat{p} - z_\alpha \sqrt{n\hat{p}(1-\hat{p})}\right],$$

   where $z_\alpha = \Phi^{-1}(1 - \alpha/2)$. Unfortunately, because the normal approximation is inexact, the Wald Interval does not always satisfy coverage, particularly at small values of p; this is why there are other confidence interval methods.

   (a) Suppose that we have a dataset $x_1, \ldots, x_n \sim \text{Ber}(p)$ (for an unknown value of p, which we wish to learn), and we (or the individual data-holders) release $y_1, \ldots, y_n$ where each $y_i = \text{RR}_\varepsilon(x_i)$ independently (where RR_ε is the ε-DP randomized response). Then, for some value of q, we have $y_1, \ldots, y_n \sim \text{Ber}(q)$. Calculate q in terms of p and ε.

   (b) Given $y_1, \ldots, y_n$, we can calculate a $(1 - \alpha)$-confidence interval $I_\alpha = [\ell, u]$ for q using one of the known confidence-interval methods for Bernoulli data. Show how to convert I_α into a $(1 - \alpha)$-confidence interval J_α for the parameter p.

(c) Suppose instead we use the Shuffle Model protocol where $(y_1, \ldots, y_n) = \text{Shuffle}(\text{RR}_{\varepsilon_0}(x_1), \ldots, \text{RR}_{\varepsilon_0}(x_n))$ where ε_0 is set to give a final (ε, δ) privacy guarantee with $\varepsilon \ll \varepsilon_0$. Explain how we can now get a confidence interval for the parameter p.