Duff: A *D*ataset-Based *U*tility *F*unction *F*amily for the Exponential Mechanism

Andrés Muñoz Medina [*1] **Jennifer Gillenwater** [*1]

Abstract

We propose and analyze an instantiation of the exponential mechanism for the release of private statistics. We establish the first ever connection between the exponential mechanism and smooth sensitivity. We also carry out extensive empirical evaluations of the exponential mechanism against other methods, showing improvements of up to 10x in quality.

1. Introduction

Differential privacy (Dwork, 2006) is a very well studied framework for protecting information about individuals in a database. In this paper we are concerned with designing differentially private methods for releasing aggregated statistics over a dataset. Differential privacy can be achieved by perturbing the outcome of a computation. This creates a tension between privacy and utility. On the one hand, a large perturbation will provide more privacy for individuals, on the other hand, perturbing an statistic could degrade information. We propose an instantiation of the well known exponential mechanism and show a previously unknown connection with the smooth sensitivity of a dataset. Our bounds provide the first exponentilally decaying error with variance scaled by the smooth sensitivity. The results of our empirical comparisons also demonstrate that the exponential mechanism is a much better algorithm for releasing private medians than the previous state-of-the-art.

2. Main results

In this section we present some common concepts of differential privacy. We will denote by $\mathbb{S}$ a universe of datasets. A dataset $\mathcal{S} \in \mathbb{S}$ is a collection of information about individuals. We assume that each individual has contributed one value to $\mathcal{S}$.

Definition 1. *Datasets $\mathcal{S}, \mathcal{S}' \in \mathbb{S}$ are **neighbors** if $\mathcal{S}$ can be obtained from $\mathcal{S}'$ by adding or removing a single element.*

We denote the neighbors of a dataset $\mathcal{S}$ by $\mathcal{N}(\mathcal{S})$.

Definition 2. *Let $\mathcal{M}$ denote a (possibly randomized) function mapping a dataset $\mathcal{S}$ to an output in $[a, b]$. We say $\mathcal{M}$ is an (ϵ, δ)-**differentially-private mechanism** if, for any two neighboring datasets $\mathcal{S}$ and $\mathcal{S}'$, and any set of outcomes $A \subseteq [a, b]$, it holds that:*

$$\Pr(\mathcal{M}(\mathcal{S}) \in A) \leq e^{\epsilon} \Pr(\mathcal{M}(\mathcal{S}') \in A) + \delta. \quad (1)$$

Given some target statistic of a dataset, $T : \mathbb{S} \to [a, b]$ (e.g., median, mode), one common mechanism $\mathcal{M}$ that is often used to release a differentially-private version of T is $\mathcal{M}(\mathcal{S}) = T(\mathcal{S}) +$ noise. The exact distribution is scaled to match the *sensitivity* of T.

Definition 3. *Given a function $T : \mathbb{S} \to [a, b]$ and a dataset $\mathcal{S} \in \mathbb{S}$, the **local sensitivity** of T at $\mathcal{S}$ is:*

$$\mathsf{LS}(T, \mathcal{S}) = \max_{\mathcal{S}' \in \mathcal{N}(\mathcal{S})} |T(\mathcal{S}) - T(\mathcal{S}')|.$$

This quantity captures how much T can change if $\mathcal{S}$ is replaced with one if its neighbors $\mathcal{S}'$. (Throughout the paper we will drop the dependency on T and $\mathcal{S}$ when it can be understood from context.) The *global sensitivity* of T builds on the definition of local sensitivity.

Definition 4. *Given a function $T : \mathbb{S} \to [a, b]$, the **(global) sensitivity** of T is:*

$$\mathsf{GS}(T) = \max_{\mathcal{S} \in \mathbb{S}} \mathsf{LS}(T, \mathcal{S}).$$

Adding noise to T with variance proportional to $\mathsf{GS}(T)$ is one way of constructing a mechanism $\mathcal{M}$ that is differentially private. For example, the following mechanism is well-known in differential privacy.

Proposition 1 (Laplace mechanism). *Let T be a function with sensitivity $\mathsf{GS}(T)$. Then the mechanism M that releases $\mathcal{M}(S) = T(S) + \frac{\mathsf{GS}(T)}{\epsilon} Z$, where $Z \sim Lap(0, 1)$, is $(\epsilon, 0)$-differentially private.*

[*]Equal contribution [1]Google Research, New York, USA. Correspondence to: Andrés Muñoz Medina <ammedina@google.com>, Jennifer Gillenwater <jengi@google.com>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 108, 2020.

While the Laplace mechanism provides us with a simple way of releasing statistics, in general the amount of noise added might be more than is strictly necessary. This is because $\mathsf{GS}(T)$ is calculated using the worst possible scenario; it is a max over all possible datasets. It might seem like an easy fix is to simply add noise proportional to $\mathsf{LS}(\mathcal{S})$. However, this is not differentially-private, as $\mathsf{LS}(\mathcal{S})$ is itself a sensitive quantity; see Section 2.1 of Nissim et al. (2007) for a detailed example. To bridge the gap between LS and GS, Nissim et al. (2007) introduced the notion of *smooth sensitivity*, which depends on the distance between datasets.

Definition 5. *Datasets $\mathcal{S}, \mathcal{S}'$ are at **distance** k, denoted $d(\mathcal{S}, \mathcal{S}') = k$, if there exists a sequence $\mathcal{S} = \mathcal{S}_0, \ldots, \mathcal{S}_k = \mathcal{S}'$ such that $\mathcal{S}_i \in \mathcal{N}(\mathcal{S}_{i-1})$ and no sequence satisfying these properties has length less than k.*

Definition 6. *Given a function $T : \mathbb{S} \to [a, b]$, $\beta > 0$ and a dataset $\mathcal{S} \in \mathbb{S}$, the β-**smooth sensitivity** of T at $\mathcal{S}$ is:*

$$\mathsf{SS}_\beta(T, \mathcal{S}) = \max_{k \geq 0} \max_{\mathcal{S}' :\, d(\mathcal{S}, \mathcal{S}') = k} e^{-\beta k} \mathsf{LS}(T, \mathcal{S}').$$

One important property of smooth sensitivity is that for $\beta > 0$:

$$\mathsf{LS}(T, \mathcal{S}) \leq \mathsf{SS}_\beta(T, \mathcal{S}) \leq \mathsf{GS}(T). \tag{2}$$

The other crucial property of SS is:

Proposition 2 (Nissim et al. (2007))**.** *Fix some $\epsilon, \delta > 0$. Let $\alpha = \frac{\epsilon}{2}$ and $\beta = \frac{\epsilon}{2 \log \frac{2}{\delta'}}$, where $\delta' = \frac{2\delta}{(e^{\epsilon/2}+1)}$. Then the mechanism that returns $T(\mathcal{S}) + \frac{\mathsf{SS}_\beta(T,\mathcal{S})}{\alpha} Z$, where $Z \sim Lap(0, 1)$, is (ϵ, δ)-differentially private.*

The above mechanism or small variants of it (Bun & Steinke, 2019) is the state of the art alogirhtm for guaranteeing ϵ, δ-DP in tasks such as median calculations.

We now turn our attention to another popular mechanism for releasing differentially private information and the main focus of this paper: the exponential mechanism. This mechanism defines a distribution using a *utility function*.

Definition 7. *(McSherry & Talwar, 2007, Definition 2) Given a utility function $u : [a, b] \times \mathbb{S} \to \mathbb{R}$, the **exponential mechanism** outputs $x \in [a, b]$ with probability proportional to $\exp\left(\frac{\epsilon u(x,\mathcal{S})}{2\Delta_u}\right)$, where Δ_u is the sensitivity of u:*

$$\Delta_u = \max_{\mathcal{S} \in \mathbb{S}} \max_{\mathcal{S}' \in \mathcal{N}(\mathcal{S})} \max_{x \in [a,b]} |u(x, \mathcal{S}) - u(x, \mathcal{S}')|.$$

We introduce our utility function to be used with the exponential mechanism.

From the example above, it is clear that a good utility function is one that assigns low value to points far away from the true statistic value, $T(\mathcal{S})$, yet has very low sensitivity. We will now define a dataset-based utility function family (*Duff*) that can achieve this for datasets where the smooth sensitivity of the statistic function T (Definition 6) is small.

Definition 8. *For dataset $\mathcal{S}$ and statistic T, Duff is:*

$$u_d(x, \mathcal{S}) = -\min_{\mathcal{S}' :\, T(\mathcal{S}') = x} d(\mathcal{S}, \mathcal{S}').$$

This utility function considers all datasets that have statistic value x, and finds one that is closest to the actual dataset $\mathcal{S}$. Function$d$ then measures how difficult it is to go from $T(\mathcal{S})$ to x, and so its negation is the utility of output x for the dataset $\mathcal{S}$. One of the important features of *Duff* is that the sensitivity of this function is always 1.

Theorem 1. *Let $x \in [a, b]$ denote the output of the exponential mechanism with utility function u_d. Let λ denote the Lebesgue measure and $H_t = \{x \mid u_d(x, \mathcal{S}) \geq -t\}$. Assume $\lambda(H_t) \geq Ct$ for some constant $C > 0$. Let $\beta_{\exp} = \frac{\epsilon}{4W\left(\frac{\epsilon(b-a)}{C\delta}\right)}$, where W is the main branch of the Lambert function[1]. Then with probability at least $1 - \delta$:*

$$|x - T(\mathcal{S})| < 4(e-1)\frac{\mathsf{SS}_{\beta_{exp}}}{\epsilon} W\left(\frac{\epsilon(b-a)}{C\delta}\right). \tag{3}$$

Experiments. In this section we provide an empirical demonstration of the practical advantages of *Duff* for the task of computing medians. For comparison, we test the smooth sensitivity mechanism, SS, of Nissim et al. (2007). As described in Section 2, this mechanism has several variants, and we test three of them:

- SS_ϵ: $(\epsilon, 0)$-differential privacy method of Nissim et al. (2007), with γ set to 2 (yielding the Cauchy distribution), and
- $\mathrm{SS}_{\epsilon,\delta}$: (ϵ, δ)-differential privacy method of Proposition 2, with the standard "reasonable" value of $\delta = 1/|\mathcal{S}|$, as well as the larger, non-private value $\delta = 0.9$.

The results of our experiments can be found in Table 1 where we clearly demonstrate that our utility function yields the best results. A more extensive evaluation can be found in the supplementary material.

Mechanism	ϵ		
	0.5	1.0	2.0
SS, $\delta = 0$	64.3 ($\pm$17.4)	22.6($\pm$10.8)	7.9 ($\pm$5.3)
SS, $\delta = 0.001$	14.8 ($\pm$2.5)	4.0 ($\pm$0.8)	1.2 ($\pm$0.3)
SS, $\delta = 0.9$	2.8 ($\pm$1.0)	1.1 ($\pm$0.7)	0.4 ($\pm$0.2)
Duff	**0.6** ($\pm$0.1)	**0.3** ($\pm$0.1)	**0.2** ($\pm$0.1)

Table 1. Average $|T(\mathcal{S}) - \mathcal{M}(\mathcal{S})| \times 100$ for data from $N(0, 1)$. Standard deviation across datasets is given in parentheses.

[1]The Lambert function is the inverse of the function $x \mapsto xe^x$

References

Bun, M. and Steinke, T. Average-Case Averages: Private Algorithms for Smooth Sensitivity and Mean Estimation. In *Neural Information Processing Systems (NeurIPS)*, 2019.

Dwork, C. Differential Privacy. In *International Colloquium on Automata, Languages, and Programming (ICALP)*, 2006.

McSherry, F. and Talwar, K. Mechanism Design via Differential Privacy. In *Foundations of Computer Science (FOCS)*, 2007.

Nissim, K., Raskhodnikova, S., and Smith, A. Smooth Sensitivity and Sampling in Private Data Analysis. In *Symposium on the Theory of Computing (STOC)*, 2007.