The Exponential Mechanism for Medians

May 19, 2020

1 The Exponential Mechanism

Sometimes, the global sensitivity of a function is too great, so the Laplace mechanism will not produce meaningful results. The median is one such function. In many cases, the *Exponential mechanism* is an alternate approach that gives reasonable utility.[1] Introduced in 2007 by McSherry and Talwar, the exponential mechanism posits that for a given database, users prefer some outputs over others. That those preferences may be encapsulated with a utility score, where a high utility score indicates a higher preference for that output. The exponential mechanism releases outputs with probability proportional (in the exponent) to the utility score and the sensitivity of the utility function.

Definition 1. *Let $\mathcal{X}$ be a space of databases and let $[m, M]$ be an arbitrary range. Let $u : \mathcal{X} \times [m, M] \to \mathbb{R}$ be a utility function, which maps pairs of databases and outputs to a utility score. Let Δu be the sensitivity of u with respect to the database argument. The exponential mechanism outputs $r \in [m, M]$ with probability proportional to $\exp\left(\frac{\varepsilon u(x,r)}{2\Delta u}\right)$ [MT07, DR+14].*[2]

Theorem 1. *The exponential mechanism preserves $(\varepsilon, 0)$-differential privacy [MT07, DR+14].*[3]

Note that the exponential mechanism may not be tractable in many cases, as it assumes the existence of a utility function, and even if one exists it may not be tractable to compute it efficiently.

[1] This is not the *only* advantage of the exponential mechanism. It is a way to compute differentially private queries on non-numeric data, unlike the Laplace mechanism it does not assume that the probability of outputting a response ought to be symmetric about the true response, etc.

[2] The original definition is from [MT07], but here we state the version rewritten in [DR+14] as it is slightly clearer.

[3] As written in [MT07], the mechanism actually preserves $(2\varepsilon\Delta u, 0)$-differential privacy; the main difference in the [DR+14] version is that it has the extra factor of $2\Delta u$ to avoid these extra terms.

2 An Exponential Mechanism for a quantile

2.1 Defining a sensible utility function

Note that a user will prefer an output that is closer to the true quantile over one that is further away. Let x be an (ordered) data set, let r be a possible output, and let N be the size of the data set. Let $\#(Z > r)$ refer to the number of points in x above r. Then, the following is a reasonable utility function for a release r for the α-quantile of x.

$$u(x, r) = \max(\alpha, (1-\alpha))N - |(1-\alpha)\#(Z < r) - \alpha\#(Z > r)|. \tag{2.1}$$

2.2 Sensitivity of the utility function

2.2.1 Neighboring Definition: Change One

Lemma 1. *The above utility function u has $\ell_1 1$ sensitivity bounded above by 1 in the change one model.*

Proof. Let $c_1 = \#(Z < r)$ and $c_2 = \#(Z > r)$. In one worst case, c_1 increases by 1 and c_2 decreases by 1. Then,

$$\begin{aligned}
\Delta u &= |(1-\alpha)(c_1+1) - \alpha(c_2-1)| - |(1-\alpha)c_1 - \alpha c_2| \\
&\leq |(1-\alpha)(c_1+1) - \alpha(c_2-1) - (1-\alpha)c_1 + \alpha c_2| \\
&\leq |c_1 + 1 - \alpha c_1 - \alpha - \alpha c_2 + \alpha - c_1 + \alpha c_1 + \alpha c_2| \\
&= 1
\end{aligned}$$

If instead c_2 decreases by 1 and c_1 increases by 1, the same thing will happen except with a negative sign that will not impact the final result due to the absolute values. $\square$

2.2.2 Neighboring Definition: Add/Drop One

Lemma 2. *The above utility function u has $\ell_1 1$ sensitivity bounded above by $\max(1-\alpha, \alpha)$ in the add/drop one model.*

Proof. Let $c_1 = \#(Z < r)$ and $c_2 = \#(Z > r)$. Consider what happens if one point is added. There are two cases that would impact the utility function:

1. c_1 increases by one and nothing happens to c_2.
2. c_2 increases by one and nothing happens to c_1.

Say the first case occurs. Then,

$$\begin{aligned}
\Delta u &= |(1-\alpha)(c_1+1) - \alpha(c_2)| - |(1-\alpha)c_1 - \alpha c_2| \\
&\leq |(1-\alpha)(c_1+1) - \alpha(c_2) - (1-\alpha)c_1 + \alpha c_2| \\
&= 1 - \alpha
\end{aligned}$$

In the second case,

$$\begin{aligned}
\delta u &= |(1-\alpha)(c_1) - \alpha(c_2+1)| - |(1-\alpha)c_1 - \alpha c_2| \\
&\leq |c_1 - \alpha c_1 - \alpha c_2 - \alpha - c_1 + \alpha c_1 + \alpha c_2| \\
&= \alpha
\end{aligned}$$

$\square$

Subtracting a point leads to the same results.

References

[DFM+20] Wenxin Du, Canyon Foot, Monica Moniot, Andrew Bray, and Adam Groce. Differentially private confidence intervals. *arXiv preprint arXiv:2001.02285*, 2020.

[DR+14] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[MT07] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 94–103. IEEE, 2007.

[Smi11] Adam Smith. Privacy-preserving statistical estimation with optimal convergence rates. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 813–822, 2011.