PREM: Privately Answering Statistical Queries with Relative Error

Badih Ghazi BADIHGHAZI@GOOGLE.COM
Google Research

Cristóbal Guzmán CRGUZMAN@GOOGLE.COM
Google Research
Faculty of Mathematics and School of Engineering, Pontificia Universidad Católica de Chile

Pritish Kamath PRITISHK@GOOGLE.COM
Google Research

Alexander Knop ALEXANDERKNOP@GOOGLE.COM
Google Research

Ravi Kumar RAVI.K53@GMAIL.COM
Google Research

Pasin Manurangsi PASIN@GOOGLE.COM
Google Research

Sushant Sachdeva SACHDEVA@CS.TORONTO.EDU
Google Research
University of Toronto

Editors: Nika Haghtalab and Ankur Moitra

Abstract

We introduce PREM (Private Relative Error Multiplicative weight update), a new framework for generating synthetic data that achieves a *relative* error guarantee for statistical queries under (ε, δ)-differential privacy (DP). Namely, for a domain $\mathcal{X}$, a family $\mathcal{F}$ of queries $f : \mathcal{X} \to \{0, 1\}$, and $\zeta > 0$, our framework yields a mechanism that on input dataset $D \in \mathcal{X}^n$ outputs a synthetic dataset $\widehat{D} \in \mathcal{X}^n$ such that all statistical queries in $\mathcal{F}$ on D, namely $\sum_{x \in D} f(x)$ for $f \in \mathcal{F}$, are within a $1 \pm \zeta$ *multiplicative* factor of the corresponding value on $\widehat{D}$ up to an *additive* error that is polynomial in $\log |\mathcal{F}|$, $\log |\mathcal{X}|$, $\log n$, $\log(1/\delta)$, $1/\varepsilon$, and $1/\zeta$. In contrast, any (ε, δ)-DP mechanism is known to require worst-case additive error that is polynomial in at least one of n, $|\mathcal{F}|$, or $|\mathcal{X}|$. We complement our algorithm with nearly matching lower bounds.

Keywords: Differential Privacy, Synthetic Data Generation, Query Answering, Relative Error

Acknowledgments

CG's research partially supported by ANID FONDECYT 1251029 grant and National Center for Artificial Intelligence CENIA FB210017, Basal ANID. SS's research is partially supported by an Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2018-06398 and a Sloan Research Fellowship. Part of this work was done when CG and SS were visiting Google Research.

. Extended abstract. Full version appears as [arXiv 2502.14809,v2]