Pufferfish Differential Privacy

Dr Chien-Hung (Joseph) Chien

ABS Methodology Division

16 July 2021

Disclaimer

Any findings are not to be considered as official statistics and the opinions and conclusions expressed are those of authors, not the ABS.

Acknowledgement

Excellent work by ABS colleagues: Cedric Wong, Edwin Lu, Jacob Ryan and James Bailie. I would like to thank the following ABS colleagues - Rob Walter, Sarah Kiely, Jenny Spencer and Kirrilie Horswill for supporting this research.

Background and context

- *Statistics Determination 2018* describes circumstances that enable information can be disclosed by the ABS.
- The Determination allows for greater utility, but also requires the ABS provides passive confidentiality under certain circumstances.
- Consequential suppression:
  - is the traditional confidentiality method, but
  - causes information loss.
- Proposing a new approach that:
  - uses a Differential Privacy framework,
  - facilitates privacy risk-utility trade off discussion,
  - enables the ABS's ability to meet user needs, and
  - reduces the costs of managing statistical risks.

Consequential suppression causes utility loss

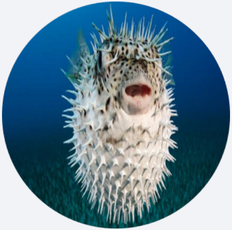
Regions	Contributors	Sales ($000)
A	12	2608
B	3	2562
C	5	2608
D	1	n.p.
E	2	n.p.
Total	23	13427

- D – Primary suppression.
- E – Secondary suppression to protect D.
- Further suppressions on outputs with contributors in E (and D).

Regions	Contributors	Sales ($000)
A	12	2608
B	3	2562
C	5	2608
D	1	2727
E	2	2240
Total	23	12745

- Sensitive record in D is protected by DP noise.
- Perturbed value is used in all outputs.
- D is protected in all outputs that it contributes to.

Why it is called Pufferfish?



$\Longrightarrow$

$\Longrightarrow$

- Pufferfish
  - tetrodotoxin
- Data
  - sensitive info (secrets)

- Certified chef
  - remove toxin
- Formal privacy definition algorithm
  - protect sensitive info (secrets)

- Fugu sashimi
  - no toxin
- Safe data
  - minimum sensitive info (secrets)

Source: adapted from [Machanavajjhala et al., 2017]

Pufferfish DP Framework

Kifer and Machanavajjhala [2014] define given set of potential secrets $\mathbb{S}$, a set of discriminative pairs $\mathbb{S}_{pairs}$, a set of all plausible data distributions $\mathbb{D}$, and a privacy parameter $\epsilon > 0$, a publishing method M satisfies ϵ-Pufferfish($\mathbb{S}$, $\mathbb{S}_{pairs}$, $\mathbb{D}$) privacy if

- for all possible outputs $\omega \in \text{range}(M)$,
- for all pairs $(s_i, s_j) \in \mathbb{S}_{pairs}$ of potential secrets,
- for all distributions $\theta \in \mathbb{D}$ which $P(s_i \mid \theta) \neq 0$ and $P(s_j \mid \theta) \neq 0$

the following holds

$$P(M(\mathfrak{Data}) = \omega \mid s_i, \theta) \leq e^{\epsilon} P(M(\mathfrak{Data}) = \omega \mid s_j, \theta) \quad (1)$$

$$P(M(\mathfrak{Data}) = \omega \mid s_j, \theta) \leq e^{\epsilon} P(M(\mathfrak{Data}) = \omega \mid s_i, \theta) \quad (2)$$

Our q% interval Pufferfish DP

Inspired by the $p\%$ rule. Choose a fixed $q \in (0, 1)$. The set of secrets is defined as

$$\mathbb{S} = \left\{\sigma_{[(1-q)y,(1+q)y)} : y > 0\right\} \cup \left\{\sigma_{((1+q)y,(1-q)y]} : y < 0\right\} \tag{3}$$

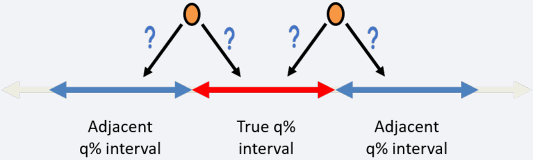
where $\sigma_{[(1-q)y,(1+q)y)}$ is a statement that the record's value is within the interval ${}_{[(1-q)y,(1+q)y)}$. The set of discriminative pairs is defined as

$$\begin{aligned}\mathbb{S}_{pairs} = &\left\{\left(\sigma_{[(1-q)y,(1+q)y)}, \sigma_{[(1+q)y, \frac{(1+q)^2}{(1-q)}y)}\right) : y > 0\right\} \cup \\ &\left\{\left(\sigma_{[\frac{(1+q)^2}{(1-q)}y,(1+q)y)}, \sigma_{[(1+q)y,(1-q)y)}\right) : y < 0\right\}.\end{aligned} \tag{4}$$

The data evolution scenarios $\mathbb{D}$ is defined as the set of probability distributions where $\theta \in \mathbb{D}$ if and only if $P(y > 0 \mid \theta) + P(y < 0 \mid \theta) = 1$, with θ corresponds to a data user's prior knowledge about the data.

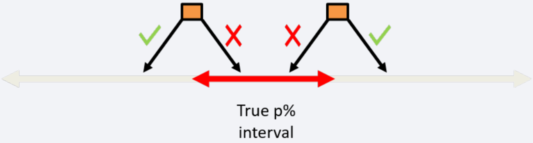
Main ideas

Pufferfish Protection



- Probabilistic (Inferential).
- All sensitive values are protected with noise.

p% Rule



- Deterministic.
- Values that violate the rule must be suppressed.

Log-Laplace multiplicative noise

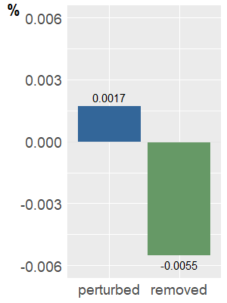
Given the set of potential secrets $\mathbb{S}$, the set of discriminative pairs $\mathbb{S}_{pairs}$, the set of data evolution scenarios $\mathbb{D}$, and privacy parameter $\epsilon > 0$, the log-Laplace multiplicative perturbation mechanism

$$M(\mathfrak{Data} = y) = ce^{X}y$$

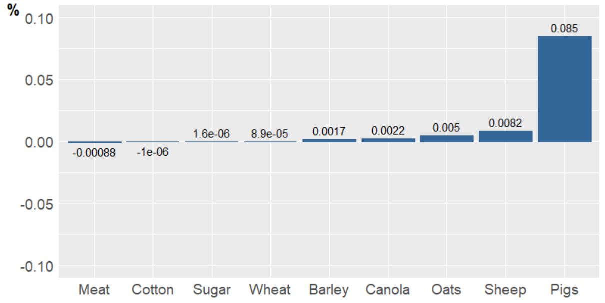
satisfies $(\mathbb{S}, \mathbb{S}_{pairs}, \mathbb{D}, \epsilon)$-Pufferfish, where $X \sim Laplace(0, b)$ with $b = -\frac{4}{\epsilon}\ln(1 - q)$ and $c = 1 - b^2$ (bias correction factor, equal to $1/E(e^X)$. Note that e^X has a log-Laplace distribution if X has a Laplace distribution. Note proof in the upcoming paper.

Case 1 : Impact to Aggregate Estimates

% differences for the total barley production



% difference for the total production for principle commodities



Case 2 : Tuning parameters using sugarcane data

Australian Bureau of Statistics

- Explore utility and risk trade-off under different privacy parameters setting for ϵ and q
- The case study design:
  - Set an arbitrary $p\%$ at 15%
  - 3 contributors in a small area from the sugarcane administrative data

$$Y = y_1 + y_2 + y_3 \tag{5}$$

$$Y^{\ddagger} = y_1 + ce^X y_2 + y_3 \tag{6}$$

    where Y is the true total of sugarcane production, $Y^{\ddagger}$ is the perturbed total of sugarcane production and y_1, y_2 and y_3 are the largest, second and third contributors.
  - The risk scenario is the largest contributor wants to estimate the second largest contributor's true value.

Case study 2 : risk and utility assessment

We perform $M = 1000$ simulations for each combination of arbitrary selected ϵ and q for the following:

- Measure utility loss as

$$RSE = \frac{\sqrt{\sum_{m=1}^{M}(Y_m^{\ddagger} - Y)^2}}{Y}$$

where $Y_m^{\ddagger}$ is the perturbed total of sugarcane production from simulation run m and Y is the true total of sugarcane production.

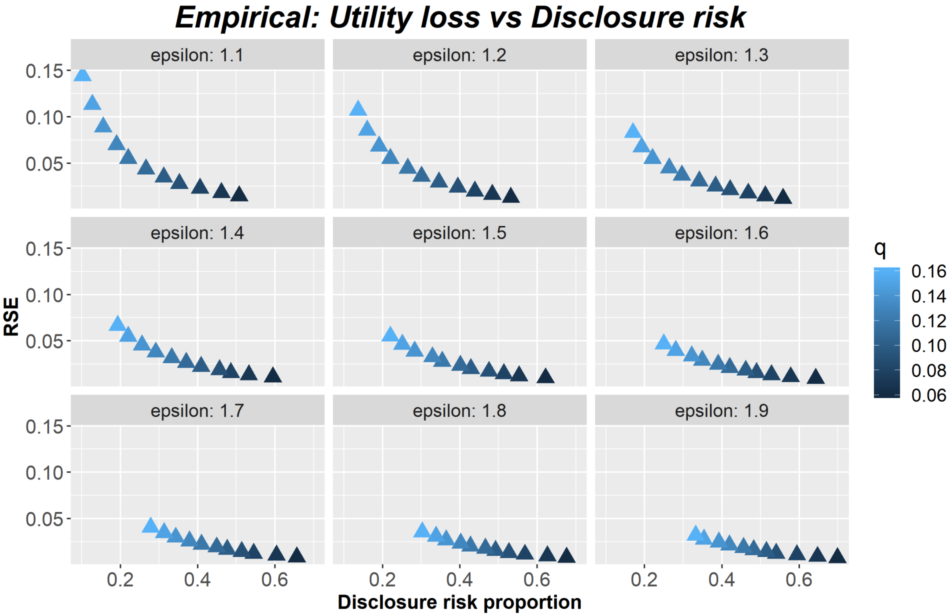
- Assess risk of disclosure as

$$P = \frac{\sum_{M}^{m=1} I_m}{M}$$

where P is the proportion of $p\%$ violation and $I_m =$

$$\begin{cases} 1, & \text{if}(Y_m^{\ddagger} - y_1) \in [y_2(1-p\%), \\ & y_2(1+p\%)], \\ 0, & \text{otherwise.} \end{cases}$$

Case study 2 : empirical results



Conclusion and future directions

- DP Pufferfish offers a nice framework to explore utility and risk trade-off.
- many areas of future research
  - explore a test case for producing experimental statistics.
  - application on two or more passive claimants cases.
  - explore an input based disclosure risk measure.
  - determine an optimal choice of ϵ and q under different settings.
  - explore truncated noise distribution.

References I

Kifer, D. and Machanavajjhala, A. (2014). Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, 39(1):1–36.

Machanavajjhala, A., He, X., and Hay, M. (2017). Differential privacy in the wild: A tutorial on current practices & open challenges. In *Proceedings of the 2017 ACM International Conference on Management of Data*, pages 1727–1730.