Local Sensitivity and Smooth Sensitivity

Presenter: Guoming Wang

27 Jan, 2015
Materials

1. S. Raskhodnikova et al.
   1. Smooth Sensitivity and Sampling

2. A. Machanavajjhala
   1. Smooth Sensitivity and Sampling
Outline

- Local Sensitivity
- Smooth Sensitivity
Laplacian Noise

In order for two worst-case neighboring data sets to produce a similar distribution of privatized answers, we need to add noise to span the sensitivity gap.

Adding laplacian Noise is not the only way, but it’s easy.

\[
Prob(R = x \mid D \text{ is the true world}) = \frac{e^{\frac{|x - F(D)|\varepsilon}{2\Delta F}}}{\varepsilon} \]

Global Sensitivity

How to privatize a series of FIVE overlapping counts across a data set? (ie, “How many people in the data set are female?”, “How many like biber?”, “How many are between age 12-16”, etc)

\[ \Delta F = \max_{\{D_1, D_2\}} \left| \left| F(D_1) - F(D_2) \right| \right|_{L_1} \]

Add laplacian noise calibrated to \( \Delta F = 5 \), to each count
Local Sensitivity

Example: median of $x_1, \ldots, x_n [0, 1]$

$X = 0000111 \quad x' = 0001111 \quad \text{Too much noise}$

$\text{Median}(x) = 0, \quad \text{median}(x') = 1$

$G_{s_{\text{median}}} = 1$

Local Sensitivity: $LS_f(x) = \max_{y:d(x, y)=1} || f(x) - f(y) ||$

Global Sensitivity: $GS_f = \max_{x,y:d(x,y)=1} ||f(x) - f(y) ||$

$GS_f = \max_x LS_f(x)$
Local Sensitivity

Add noise proportional to $LS_f(x)$ instead of $GS_f$?

Not a good idea, because it reveals information.

$D_1 = \{0 \ 0 \ 0 \ 0 \ 0 \ \wedge \ \wedge \ \wedge \ \wedge \}$

$Q_{med}(D1) = 0$

$LS_{qmed}(D1) = 0$ \hspace{0.5cm} Noise sampled from Lap(0)

$D_2 = \{0 \ 0 \ 0 \ 0 \ \wedge \ \wedge \ \wedge \ \wedge \}$

$Q_{med}(D2) = 0$

$LS_{qmed}(D2) = \wedge$ \hspace{0.5cm} Noise sampled from Lap($\wedge/\varepsilon$)
Smooth Sensitivity

$S(x)$ is an $\varepsilon$-smooth upper bound on $\text{LS}_f(x)$ if:
- for all $x$: $S(x) \geq \text{LS}_f(x)$
- for all neighbors $x, x'$: $S(x) \leq e^\varepsilon S(x')$
Smooth Sensitivity

\[ S^*_{q}(d) = \max_{d'} \left( LS_q(d') \exp(-m\beta) \right) \]

where \( d \) and \( d' \) differ in \( m \) entries.

\[
\begin{array}{cccccccccc}
& x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} \\
\hline
d & \text{} & \text{} & \text{} & \text{\color{red}x_5} & \text{} & \text{} & \text{} & \text{} & \text{} & \text{} \\
d' & \text{\color{red}x_4} & \text{\color{red}x_5} & \text{} & \text{} & \text{} & \text{0} & \text{0} & \text{0} & \text{} & \text{} \\
\end{array}
\]

- \( x_{5-k} \leq q_{\text{med}}(d') \leq x_{5+k} \)
- \( LS(d') = \max(x_{\text{med}+1} - x_{\text{med}}, x_{\text{med}} - x_{\text{med}-1}) \)

\[ S^*_{q_{\text{med}}}(d) = \max_k \left( \exp(-k\beta) \times \max_{5-k \leq \text{med} \leq 5+k} (x_{\text{med}+1} - x_{\text{med}}, x_{\text{med}} - x_{\text{med}-1}) \right) \]
Smooth Sensitivity

For instance, $\Lambda = 1000$, $\beta = 2$.

$$d \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10$$

$$S^*_{qmed}(d) = \max \left( \max_{0 \leq k \leq 4} (\exp(-\beta \cdot k) \cdot 1), \max_{5 \leq k \leq 10} (\exp(-\beta \cdot k) \cdot \Lambda) \right)$$

$$= 1$$
Thank you – Enjoy the rest of your night