Harvard Privacy Tools Project

Mean Sensitivity Proofs

March 9, 2020

Definition 1. *The sample mean of database X of size n is*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

1 Neighboring Definition: Change One

1.1 ℓ_1-sensitivity

Theorem 1. *Say database X has size n and is bounded above by M and bounded below by m. Then $\bar{X}$ has ℓ_1-sensitivity bounded above by*

$$\frac{M-m}{n}.$$

Proof. Say X and X' are neighboring databases which differ at data-point x_j. Then

$$\begin{aligned}
\Delta\bar{X} &= \max_{X,X'} \left|\bar{X} - \bar{X}'\right| \\
&= \max_{X,X'} \frac{1}{n}\left|\left(\sum_{\{i\in[n] \mid i\neq j\}} x_i\right) + x_j - \left(\sum_{\{i\in[n] \mid i\neq j\}} x'_i\right) + x'_j\right| \\
&= \max_{X,X'} \frac{1}{n}\left|x_j - x'_j\right| \\
&\leq \frac{M-m}{n}.
\end{aligned}$$

□

1.2 ℓ_2-sensitivity

2 Neighboring Definition: Add/Drop One

2.1 ℓ_1-sensitivity

Theorem 2. *Say database X has size n and is bounded above by M and bounded below by m. Then $\bar{X}$ has ℓ_1-sensitivity bounded above by*

$$\frac{M-m}{n}.$$

Proof. WLOG assume point being added/subtracted is x_n.
Adding a point: $X' = X \cup x$

$$\begin{aligned}
\left|\bar{X} - \bar{X}'\right| &= \left|\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n+1}\sum_{i=1}^{n+1} x_i\right| \\
&= \left|\left(\frac{1}{n} - \frac{1}{n+1}\right)\sum_{i=1}^{n} x_i - \frac{x}{n+1}\right| \\
&= \frac{1}{n+1}\left|\frac{1}{n}\sum_{i=1}^{n} x_i - x\right| \\
&\leq \frac{|M-m|}{n+1}
\end{aligned}$$

Taking a point away: $X' = X \backslash \{x\}$

$$\begin{aligned}
\left|\bar{X} - \bar{X}'\right| &= \left|\frac{1}{n-1}\sum_{i=1}^{n-1} x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right| \\
&= \left|\left(\frac{1}{n-1} - \frac{1}{n}\right)\sum_{i=1}^{n-1} x_i - \frac{x}{n}\right| \\
&= \frac{1}{n}\left|\frac{1}{n-1}\sum_{i=1}^{n-1} x_i - x\right| \\
&\leq \frac{|M-m|}{n}
\end{aligned}$$

□

2.2 ℓ_2-sensitivity