
Mean Sensitivity Proofs

March 22, 2020

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

$$f(X) = \frac{1}{n} \sum_{i=1}^n x_i.$$

1 NEIGHBORING DEFINITION: CHANGE ONE

1.1 ℓ_1 -sensitivity

Theorem 1. *Say the space of datapoints \mathcal{X} is bounded above by M and bounded below by m . Then $f(\cdot)$ has ℓ_1 -sensitivity in the change-one model bounded above by*

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

Proof. Say X and X' are neighboring databases which differ at data-point x_j , and let Δf indicate the ℓ_1 -sensitivity of $f(\cdot)$. Then

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

□

1.2 ℓ_2 -sensitivity

Theorem 2. *Say the space of datapoints \mathcal{X} is bounded above by M and bounded below by m . Then $f(\cdot)$ has ℓ_2 -sensitivity in the change-one model bounded above by*

$$\left(\frac{M - m}{n} \right)^2.$$

Proof. Say X and X' are neighboring databases which differ only at index j , and let Δf indicate the ℓ_2 -sensitivity of $f(\cdot)$. Then

$$\begin{aligned}
\Delta f &= \max_{X, X'} (f(X) - f(X'))^2 \\
&= \max_{X, X'} \frac{1}{n^2} \left(\left(\sum_{i \in [n] | i \neq j} x_i \right) + x_j - \left(\sum_{i \in [n] | i \neq j} x'_i \right) - x'_j \right)^2 \\
&= \max_{X, X'} \frac{1}{n^2} (x_j - x'_j)^2 \\
&\leq \frac{(M - m)^2}{n^2} \\
&= \left(\frac{M - m}{n} \right)^2
\end{aligned}$$

□

2 NEIGHBORING DEFINITION: ADD/DROP ONE

2.1 ℓ_1 -sensitivity

Theorem 3. *Say the space of datapoints \mathcal{X} is bounded above by M and bounded below by m . Then $f(\cdot)$ has ℓ_1 -sensitivity in the add/drop-one model bounded above by*

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

Proof. For notational ease, let n always refer to the size of database x . We must consider both adding and removing an element from x . First, consider adding a point:

Let $X' = X \cup \{x\}$. Without loss of generality, assume the point added is the $(n + 1)^{\text{th}}$ element of database X' . Note that

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

Second, consider removing a point:

Let $X' = X \setminus \{x\}$. Without loss of generality assume that the point subtracted is the n^{th}

element of database X .

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

Then, since $\forall n > 0$,

$$\frac{1}{n+1} < \frac{1}{n},$$

the sensitivity of the mean in general is bound from above by

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

□

2.2 ℓ_2 -sensitivity

Theorem 4. *Say the space of datapoints \mathcal{X} is bounded above by M and bounded below by m . Then f has ℓ_2 -sensitivity in the add/drop-one model bounded above by*

$$\left(\frac{M - m}{n} \right)^2.$$

Proof. For notational ease, let n always refer to the size of database X . We must consider both adding and removing an element from X . First, consider adding a point:

Let $X' = X \cup x$. Without loss of generality assume the point added is the $(n+1)^{\text{th}}$ element of database X' . Then,

$$\begin{aligned}
\Delta f &= \max_{X, X'} (f(X) - f(X'))^2 \\
&= \max_{X, X'} \left(\frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n+1} \sum_{i=1}^{n+1} x'_i \right)^2 \\
&= \max_{X, X'} \left(\left(\frac{1}{n} \sum_{i=1}^n x_i \right) - \left(\frac{1}{n+1} \sum_{i=1}^n x'_i \right) - \frac{x}{n+1} \right)^2 \\
&= \max_{X, X'} \left(\frac{(\sum_{i=1}^n x_i) - nx}{n(n+1)} \right)^2 \\
&= \left(\frac{nM - nm}{n(n+1)} \right)^2 \\
&= \left(\frac{M - m}{n+1} \right)^2.
\end{aligned}$$

Second, consider removing an element:

Let $X' = X \setminus \{x\}$. Without loss of generality assume that the point subtracted is the n^{th} element of database X . Then,

$$\begin{aligned}
\Delta f &= \max_{X, X'} (f(X) - f(X'))^2 \\
&= \max_{X, X'} \left(\frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n-1} \sum_{i=1}^{n-1} x'_i \right)^2 \\
&= \max_{X, X'} \left(\left(\frac{1}{n} \sum_{i=1}^{n-1} x_i \right) + \frac{x}{n} - \left(\frac{1}{n-1} \sum_{i=1}^{n-1} x'_i \right) \right)^2 \\
&= \max_{X, X'} \left(\frac{(n-1)x - \sum_{i=1}^{n-1} x_i}{n(n-1)} \right)^2 \\
&= \left(\frac{(n-1)M - (n-1)m}{n(n-1)} \right)^2 \\
&= \left(\frac{M-m}{n} \right)^2.
\end{aligned}$$

Then, since $\forall n > 0$,

$$\frac{1}{n+1} < \frac{1}{n},$$

the sensitivity of the mean in general is bound from above by

$$\frac{(M-m)^2}{n}.$$

□