Count Sensitivity Proofs

May 21, 2020

Definition 1. *Let $\mathcal{X}$ be the universe of possible rows (individuals) and let $I : \mathcal{X} \to \{0,1\}$ be a predicate on rows. Let $x \in \mathcal{X}^n$ be a dataset. Then a count over x is defined as*

$$q(x) = \sum_{i=1}^{n} I(x_i).$$

Definition 2. *Let $q_1, \ldots, q_k$ be a series of counts with predicates $I_1, \ldots, I_k$. These counts are disjoint for every row in the database if only one of them evaluates to 1. In other words, they are disjoint if $\forall x_i \in \mathcal{X}$,*

$$\sum_{j=1}^{k} I_j(x_i) = 1.$$

1 NEIGHBORING DEFINITION: CHANGE ONE

1.1 ℓ_1-sensitivity

Theorem 1. *A single count query has ℓ_1-sensitivity 1 in the change-one model. A series of k disjoint counts has ℓ_1-sensitivity 2 in the change-one model.*

Proof. Let q be a count query with predicate I, and let x' be equal to x with point x_i changed to x_i'. Then,

$$
\begin{aligned}
\left| q(x') - q(x) \right| &= \left| \sum_{j=1}^{n} I(x_j') - \sum_{j=1}^{n} I(x_j) \right| \\
&= \left| \sum_{\{i \in [n] | i \neq j\}} I(x_j) + I(x_i') - \sum_{\{i \in [n] | i \neq j\}} I(x_j) - I(x_i) \right| \\
&= \left| I(x_i') - I(x_i) \right| \\
&\leq 1.
\end{aligned}
$$

Consider a series of k disjoint count queries $\mathbf{q} = \{q_1, \ldots, q_k\}$ on the same databases x and x'. Note that since the counts are disjoint, x_i and x_i' can at most each increment a single one of the k counts by 1. Call the count that x_i impacts q_i, and the count that x_i' impacts q_j. Then,

$$
\begin{aligned}
\left| \mathbf{q}(x) - \mathbf{q}(x') \right| &= \left| \left(q_1(x) - q_1(x') \right) + \ldots + \left(q_k(x) - q_k(x') \right) \right| \\
&= \left| \left(q_i(x) - q_i(x') \right) + \left(q_i(x) - q_j(x') \right) \right| \\
&\leq 2
\end{aligned}
$$

$\square$

1.2 ℓ_2-sensitivity

Theorem 2. *A single count query has ℓ_2-sensitivity 1 in the change-one model. A series of k disjoint counts has ℓ_2-sensitivity $\sqrt{2}$ in the change-one model.*

Proof. From the proof of Theorem 1, the difference between counts on two neighboring databases is at most 1. Squaring this gives the same value. For a series of k disjoint counts,

$$
\begin{aligned}
\left| \mathbf{q}(x) - \mathbf{q}(x') \right|_2 &= \left(\left(q_1(x) - q_1(x') \right)^2 + \ldots + \left(q_k(x) - q_k(x') \right)^2 \right)^{1/2} \\
&\leq \left(1^2 + 1^2 \right)^{1/2} \\
&= \sqrt{2}.
\end{aligned}
$$

$\square$

2 Neighboring Definition: Add/Drop One

2.1 ℓ_1-sensitivity

Theorem 3. *A single count query has ℓ_1-sensitivity 1 in the add/drop-one model. A series of k disjoint counts also has ℓ_1-sensitivity 1 in the add/drop-one model.*

Proof. Let q be a count query with predicate I, and let x' be equal to database x with point x_i removed. Then

$$
\begin{aligned}
\left| q(x) - q(x') \right| &= \left| \sum_{j=1}^{n} I(x_j) - \sum_{\{i \in [n] | i \neq j\}} I(x_j) \right| \\
&= \left| \sum_{\{i \in [n] | i \neq j\}} I(x_j) + I(x_i) - \sum_{\{i \in [n] | i \neq j\}} I(x_j) \right| \\
&\leq 1.
\end{aligned}
$$

A nearly identical argument holds for adding a point.

Consider a series of k disjoint count queries $\mathbf{q} = \{q_1, \ldots, q_k\}$ and consider database x' equal to database x with point x_i removed. Note that only a single one of the k queries will be affected by the change from x to x', so

$$
\begin{aligned}
\left| \mathbf{q}(x) - \mathbf{q}(x') \right| &= \left| \left(q_1(x) - q_1(x') \right) + \ldots + \left(q_k(x) - q_k(x') \right) \right| \\
&\leq 1.
\end{aligned}
$$

The same argument holds for x' equal to x with a single point added. □

2.2 ℓ_2-sensitivity

Theorem 4. *A single count query has ℓ_2-sensitivity 1 in the add/drop-one model. A series of k disjoint counts also has ℓ_2-sensitivity 1 in the add/drop-one model.*