Count Sensitivity Proofs

March 17, 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, 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 sensitivity 1. A series of k disjoint counts has sensitivity 2.*

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,

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

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,

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

□

1.2 ℓ_2-sensitivity

Theorem 2. *A single count query has sensitivity 1. A series of k disjoint counts has sensitivity 2.*

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,

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

□

2 Neighboring Definition: Add/Drop One

2.1 ℓ_1-sensitivity

Theorem 3. *A single count query has sensitivity 1. A series of k disjoint counts also has sensitivity 1.*

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

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

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

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

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 sensitivity 1. A series of k disjoint counts also has sensitivity 1.*

Proof. Squaring the sensitivity bounds from Theorem 3 gives 1 as an upper bound on the ℓ_2 sensitivity for both a single count and a series of k disjoint counts. $\square$