K-Randomization

Maxim Zhilyaev David Zeber

January 7, 2016

1 Differential Privacy

The typical setting for differential privacy is as follows. We consider a **database** as a collection of records. Each record is an element of some space $\mathcal{D}$, and a database $\mathbf{x}$ is a vector of n records: $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{D}^n$.

We wish to release information retrieved from the database by means of a **query**, a function A mapping the database into another space: $A : \mathcal{D}^n \to \mathcal{S}$. The result of applying a query to a database is termed a **transcript**. The query usually applies some aggregation to the database records, and so the output space $\mathcal{S}$ is generally of lower dimensionality than the original database. If the query is randomized, i.e., $A(\mathbf{x}) = A(\mathbf{x}; \xi)$ for a random element ξ, then the transcript will be a random element of $\mathcal{S}$.

The notion of differential privacy for a database query is that the resulting transcript does not change substantially when a record in the database is modified, i.e., transcripts are not sensitive to particular individual records in the database. Hence, releasing query transcripts publicly will not jeopardize privacy, since information regarding individual records cannot be gained by analyzing query transcripts.

Differential privacy for a randomized query A is formulated by comparing the transcripts generated by applying A to two very similar databases $\mathbf{x}, \mathbf{x}' \in \mathcal{D}^n$. We say the databases **differ in one row** if $\sum_{i=1}^{n} I(x_i \neq x_i') = 1$.

Definition. A randomized query A is ϵ-**differentially private** if, for any two databases $\mathbf{x}, \mathbf{x}' \in \mathcal{D}^n$ differing in one row,

$$\mathsf{P}[A(\mathbf{x}) \in S] \leq \exp(\epsilon) \cdot \mathsf{P}[A(\mathbf{x}') \in S] \tag{1.1}$$

for all $S \subset \mathcal{S}$ (measurable).

In other words, the transcripts from the two databases databases differing in one row are close in distribution. An alternative notion of differing in one row that is sometimes used is that $\mathbf{x}'$ includes an additional record that is not in $\mathbf{x}$: $\mathbf{x} \in \mathcal{D}^n$, $\mathbf{x}' \in \mathcal{D}^{n+1}$, and $x_i = x_i'$ for $i = 1, \dots, n$.

If $\mathcal{S}$ is finite, which is common in cases where the transcript involves integer counts, then the distribution of the transcript $A(\mathbf{x})$ can be represented using its pmf $\mathsf{P}[A(\mathbf{x}) = s]$ for $s \in \mathcal{S}$. In this case, the differential privacy condition can also be expressed in terms of the pmf.

Proposition 1.1. *If $\mathcal{S}$ is finite, then A is ϵ-**differentially private** if and only if*

$$\mathsf{P}[A(\mathbf{x}) = s] \leq \exp(\epsilon) \cdot \mathsf{P}[A(\mathbf{x}') = s] \tag{1.2}$$

for all $s \in \mathcal{S}$, where $\mathbf{x}, \mathbf{x}'$ differ in one row.

Proof. ($\Leftarrow$) Given $S \subset \mathcal{S}$, we can write $\mathsf{P}[A(\mathbf{x}) \in S] = \sum_{s \in S} \mathsf{P}[A(\mathbf{x}) = s]$. If $\mathsf{P}[A(\mathbf{x}') \in S] = 0$, then $\mathsf{P}[A(\mathbf{x}') = s] = 0$ for each $s \in S$. From (1.2) we have that $\mathsf{P}[A(\mathbf{x}) = s] = 0$ as well, and so $P[A(\mathbf{x}) \in S] = 0$, verifying (1.1). Otherwise, if $\mathsf{P}[A(\mathbf{x}') \in S] > 0$,

$$\frac{\mathsf{P}[A(\mathbf{x}) \in S]}{\mathsf{P}[A(\mathbf{x}') \in S]} = \frac{\sum_{s \in S} \mathsf{P}[A(\mathbf{x}) = s]}{\sum_{s \in S} \mathsf{P}[A(\mathbf{x}') = s]} \leq \max_{s \in S} \frac{\mathsf{P}[A(\mathbf{x}) = s]}{\mathsf{P}[A(\mathbf{x}') = s]} \leq \exp(\epsilon),$$

using Lemma (need ref).
($\Rightarrow$) Take $S = \{s\}$ in (1.1). □

2 Theoretical setup

In the following we work with data in the form of bit vectors. A **bit vector** is a vector $v \in \{0, 1\}^L$.

First we define the randomization procedure we will be applying.

Definition. The randomization procedure R with **lie probability** $0 < q < 1/2$ flips a bit with probability q, and leaves it as-is with probability $1 - q$. In other words, for a bit $b \in \{0, 1\}$,

$$R(b) = R(b; X) = (1 - b) \cdot X + b \cdot (1 - X) \quad \text{where } X \sim Ber(q).$$

When applied to a vector, each bit is randomized independently:

$$R(v) = R\big(v; (X_1, \ldots, X_L)\big) = \big(R(v_1; X_1), \ldots, R(v_L; X_L)\big) \quad \text{where } X \overset{\text{iid}}{\sim} Ber(q).$$

Remark. The randomization R reports the original bit value with probability $1 - q > q$, and lies with probability q. This is equivalent to the randomized response procedure where the value is reported as-is with probability $1 - f$, and with probability f the reported value is the outcome of the toss of a fair coin. In this case, $q = f/2$.

Remark. If $q = 1/2$, then $R(0) \overset{d}{=} R(1)$, and the reported value is “completely” randomly generated, i.e., independently of the original value.

Distribution of $R(v)$.

For a bit b, the randomization lies iff $R(b) \neq b$:

$$P[R(b) = s] = q^{\mathbf{1}_{\{b \neq s\}}} (1 - q)^{\mathbf{1}_{\{b = s\}}}$$

Hence, for a bit vector v,

$$P[R(v) = s] = q^{\sum \mathbf{1}_{\{b_i \neq s_i\}}}(1-q)^{\sum \mathbf{1}_{\{b_i = s_i\}}} = q^{L-m(v,s)}(1-q)^{m(v,s)},$$

where $m(v,s) = |\{i : v_i = s_i\}|$. Note that this probability is maximized when $m(v,s) = L$ (the reported vector s is identical to the original vector v), and minimized when $m(v,s) = 0$. In other words, the most likely outcome of randomizing a bit vector is obtaining an identical vector.

For a collection T,

$$P[s \in R(T)] = 1 - P[s \notin R(T)] = 1 - \prod_{v \in T} P[R(v) \neq s] = 1 - \prod_{v \in T} \left[1 - q^{L-m(v,s)}(1-q)^{m(v,s)}\right].$$

3 Old stuff

Furthermore, if A randomizes each record in the database independently, i.e., $A(\boldsymbol{x}) = A(\boldsymbol{x}, \boldsymbol{X}) := \big(A_0(x_1, X_1), \dots, A_0(x_n, X_n)\big)$ where X_i are independent, then $\boldsymbol{S} = \boldsymbol{S}_0^n$ and $s = (s_1, \dots, s_n)$ with $s_i \in \boldsymbol{S}_0$. In this case $P[A(\boldsymbol{x}) = s] = P[A_0(x_1) = s_1, \dots, A_0(x_n) = s_n] = \prod P[A_0(x_i) = s_i]$. If $\boldsymbol{x}$ and $\boldsymbol{x'}$ differ in one row (wlog $x_1 \neq x_1'$ and $x_i = x_i'$ for $i = 2, \dots, n$), then

$$\frac{P[A(\boldsymbol{x}) = s]}{P[A(\boldsymbol{x'}) = s]} = \frac{P[A_0(x_1) = s_1]}{P[A_0(x_1') = s_1]}.$$

Therefore, in this case, the query A will satisfy differential privacy if

$$P[A_0(x) = s] \leq \epsilon \cdot P[A_0(x') = s]$$

for all $x, x' \in D$ and $s \in \boldsymbol{S}_0$. This is the formulation used in the RAPPOR paper that applies to differences between individual records rather than collections differing on a single element.