K-Randomization

Maxim Zhilyaev David Zeber

April 29, 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: $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{D}^n$.

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 $i = 1, \ldots, n$.

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 6.1.
($\Rightarrow$) Take $S = \{s\}$ in (1.1). □

2 Bit vector reporting

Our goal is to establish differential privacy properties for user data reported in the form of vectors of bits. To protect user privacy, each user record is randomized prior to leaving the client and anonymized on reaching the server. We now describe the randomization procedure, and place ourselves in the setting of Section 1 by representing it as a query applied to a database.

2.1 Bit randomization

For our purposes, a **bit** is an integer $b \in \{0, 1\}$, and a **bit vector** is a vector $x \in \{0, 1\}^L$. Bits and bit vectors are randomized in the following way.

Definition. The **bit randomization** procedure R with **lie probability** $0 < q < 1/2$ flips a bit b with probability q, and leaves it as-is with probability $p := 1 - q$:

$$R(b) = \begin{cases} b & \text{with prob } p \\ 1 - b & \text{with prob } q \end{cases}.$$

This can be expressed concisely as

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

We extend the procedure to **bit vector randomization** by applying the randomization independently to each bit in the vector. Given a bit vector $x = (b_1, \ldots, b_L)$, define

$$R(x) = R(x; \xi) = \big(R(b_1; \xi_1), \ldots, R(b_L; \xi_L)\big) \quad \text{where } \xi = (\xi_1, \ldots, \xi_L) \overset{\text{iid}}{\sim} Ber(p). \tag{2.1}$$

Remark. Note that R reports the original bit value with probability $p = 1 - q > q$, 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 our case, $q = f/2$.

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

The distribution of the randomized bit vectors can be expressed in terms of the Hamming distance between the original and randomized vectors,

$$\delta(x, x') = \sum_{\ell=1}^{L} I(x_\ell \neq x'_\ell) = \sum_{\ell=1}^{L} |x_\ell - x'_\ell|.$$

For a single bit, the randomization has lied when the outcome is different from the original value:

$$\mathsf{P}[R(b) = b'] = p^{I(b=b')} \cdot q^{I(b\neq b')} = p^{1-\delta(b,b')} \cdot q^{\delta(b,b')}.$$

For a bit vector x, this becomes

$$\mathsf{P}[R(x) = y] = p^{\sum I(x_\ell=y_\ell)} \cdot q^{\sum I(x_\ell\neq y_\ell)} = p^{L-\delta(x,y)} \cdot q^{\delta(x,y)}.$$

Note that this probability is maximized when $\delta(x, y) = 0$ (the randomized vector y is identical to the original vector x), and minimized when $\delta(x, y) = L$. In the latter case, we say that y is the **opposite** of x. In other words, the most likely outcome of randomizing a bit vector is obtaining an identical vector.

2.2 Privacy-preserving reporting for bit records

We now adapt the framework of Section 1 to the task of reporting user records encoded as bit-vectors.

Set $\mathcal{D} = \{0, 1\}^L$. We use the term **collection** (of records) interchangeably with "database". We consider a randomized query A that randomizes each record in the collection independently, and aggregates the results by reporting occurrence counts for every possible randomization outcome. We adopt this aggregation step as a model for **anonymization**. After anonymization, we cannot associate a specific randomized record with a specific record out of the original collection, and so the ordering of records in the synthetic collection conveys no information about the ordering of the original records.

(TODO - be more specific about the ordering) In the following, we rely on the fact that $\mathcal{D}$ is finite, and we assume a specific enumeration $\mathcal{D} = (d_1, \ldots, d_{2^L})$. The ordering is unimportant at this point, although it will be convenient to assume that $d_1 = (1, \ldots, 1)$ and $d_{2^L} = (0, \ldots, 0)$.

We consider an aggregation function Φ mapping a collection of elements of $\mathcal{D}$ to a vector $\mathbf{s} = (s_1, \ldots, s_{2^L})$, where s_j counts how many copies of d_j the collection contains. Note that the range of Φ is

$$\mathcal{S}_n := \left\{\mathbf{s} \in \mathbb{Z}^{2^L} : s_j \geq 0,\ s_1 + \cdots + s_{2^L} = n\right\},$$

a subset of the standard $(2^L - 1)$-simplex (whose vertices are the standard basis vectors in $\mathbb{R}^{2^L}$) consisting of the points with integer coordinates.

Definition. The aggregation function $\Phi : \mathcal{D}^n \to \mathcal{S}_n$ counts occurrences of vectors $d_1, \ldots, d_{2^L} \in \mathcal{D}$ in the collection $\mathbf{y} \in \mathcal{D}^n$:

$$\Phi(\mathbf{y}) = \left(\sum_{i=1}^{n} I(y_i = d_1), \ldots, \sum_{i=1}^{n} I(y_i = d_{2^L}) \right)$$

We also extend the bit vector randomization (2.1) to collections $\mathbf{x} \in \mathcal{D}^n$ by applying it independently to each vector in the collection:

$$R(\mathbf{x}) = R(\mathbf{x}; \boldsymbol{\xi}) = \big(R(x_1; \xi_1), \ldots, R(x_n; \xi_n)\big) \quad \text{where } \xi_i = (\xi_{i1}, \ldots, \xi_{iL}) \text{ and } \xi_{i\ell} \overset{\text{iid}}{\sim} Ber(p).$$

We call $R(\mathbf{x})$ the **synthetic** collection obtained from the **original** collection $\mathbf{x}$. Note that $R(\mathbf{x})$ is itself a random element of $\mathcal{D}^n$, with distribution given by

$$\mathsf{P}[R(\mathbf{x}) = \mathbf{y}] = \prod_{i=1}^{n} \mathsf{P}[R(x_i) = y_i] = \prod_{i=1}^{n} p^{L - \delta(x_i, y_i)} q^{\delta(x_i, y_i)} = p^{nL} \left(\frac{q}{p} \right)^{\sum_{i=1}^{n} \delta(x_i, y_i)}.$$

Thus, $\mathsf{P}[R(\mathbf{x}) = \mathbf{y}] > 0$ for any $\mathbf{y} \in \mathcal{D}^n$, i.e., any synthetic collection of size n has a non-zero probability of being generated from any given original collection of the same size.

The randomized query we use for reporting collections of bit vectors may now be defined as follows.

Definition. The randomized query $A : \mathcal{D}^n \to \mathcal{S}_n$ maps collections of bit vectors to occurrence counts according to

$$A = \Phi \circ R.$$

2.3 Next steps

- The distribution of A is invariant under permutations of $\mathbf{x}$, i.e., the distribution remains the same if the original collection is reordered.
- Hence, partition $\mathcal{D}^n$ into equivalence classes of Φ, represented as points $\mathbf{m} = (m_1, \ldots, m_{2^L}) \in \mathcal{S}_n$. Work with original collections represented in terms of $\mathbf{m}$. In the $L = 2$ case, it is convenient to associate the simplex $\mathcal{S}_n \subset \mathbb{R}^4$ with the "corner-of-the-cube" simplex in $\mathbb{R}^3$ defined by the standard basis vectors together with the origin.
- View the randomized query A as a mapping from $\mathcal{S}_n$ to $\mathcal{S}_n$. Hamming distance on $\mathcal{D}^n$ corresponds to L_1 metric on $\mathcal{S}_n$.
- Privacy ratio takes as input two "neighbouring" original collections (defined in terms of a metric on $\mathcal{S}_n$) and a synthetic one. Behaviour of privacy ratio can be studied in terms of "probability ratio", a mapping over a single original collection and two neighbouring synthetic ones.

- Define the **privacy range** as a subset of the most likely synthetic outcomes for a given original collection.
- Goal is to uniformly bound the privacy ratio over all privacy ranges corresponding to all possible original collections.
- Work with privacy ratio for $x = \mathbf{1}$ and $x' = \mathbf{0}$. Argue by symmetry that the bound on the privacy ratio only depends on the choice of x and x', regardless of remaining elements of the collection. Argue that without loss of generality we can choose x' to be the opposite of x.
- Show that the privacy ratio is increasing as the synthetic collection moves towards the one consisting of all $\mathbf{0}$. Direction of maximal increase is along the line connecting $\mu_{\mathbf{m}} = \mathsf{E}[\mathbf{s} \,|\, \mathbf{m}]$ (must be on the line, or just in the direction?)
- Think of the privacy range in terms of a neighbourhood of $\mu_{\mathbf{m}}$.

2.4 Old stuff

(TODO: maybe this is not useful at this point) Writing $m(\ell) := \sum_{i=1}^{n} I(\delta(x_i, y_i) = \ell)$ for $\ell = 0, \ldots, L$, the probability can be expressed as

$$\mathsf{P}[R(\mathbf{x}) = \mathbf{y}] = \prod_{\ell=0}^{L} \left(p^{L-\ell} q^{\ell}\right)^{m(\ell)}.$$

Hence, the distribution is determined by the magnitudes of the pairwise distances, and does not depend on the ordering within the collections.

The distribution of $A(\mathbf{x})$ is given by

$$\begin{aligned}
\mathsf{P}[A(\mathbf{x}) = \mathbf{s}] = \mathsf{P}[\Phi(R(\mathbf{x})) = s] &= \sum_{\mathbf{y}:\Phi(\mathbf{y})=s} \mathsf{P}[R(\mathbf{x}) = \mathbf{y}] \\
&= \sum_{\mathbf{y}:\Phi(\mathbf{y})=s} \prod_{i=1}^{n} \mathsf{P}[R(x_i) = y_i] \qquad (2.2) \\
&= \sum_{\mathbf{y}:\Phi(\mathbf{y})=s} \prod_{i=1}^{n} p^{L-\delta(x_i,y_i)} \cdot q^{\delta(x_i,y_i)} = \sum_{\mathbf{y}:\Phi(\mathbf{y})=s} p^{nL-\delta(\mathbf{x},\mathbf{y})} \cdot q^{\delta(\mathbf{x},\mathbf{y})} \\
&= p^{nL} \sum_{\mathbf{y}:\Phi(\mathbf{y})=s} \left(\frac{q}{p}\right)^{\delta(\mathbf{x},\mathbf{y})} \qquad (2.3)
\end{aligned}$$

where $\delta(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \delta(x_i, y_i)$ can be considered the distance between collections $\mathbf{x}$ and $\mathbf{y}$. Note that the support of A is the support of a multinomial random variable with n trials:

$$\mathcal{S}_n := \left\{ s \in \mathcal{S} : \sum_j s_j = n \right\},$$

and $\{\mathbf{y} : \Phi(\mathbf{y}) = s\}$ is the set of all permutations of the collection

$$\big(\underbrace{d_1, \ldots, d_1}_{s_1}, \underbrace{d_2, \ldots, d_2}_{s_2}, \ldots, \underbrace{d_{2^L}, \ldots, d_{2^L}}_{s_{2^L}}\big).$$

3 The privacy ratio

We will study the differential privacy of the query A in terms of the **privacy ratio**

$$\pi(\mathbf{s}; \mathbf{x}, \mathbf{x}') := \frac{\mathsf{P}[A(\mathbf{x}') = \mathbf{s}]}{\mathsf{P}[A(\mathbf{x}) = \mathbf{s}]}$$

for two collections $\mathbf{x}, \mathbf{x}' \in \mathcal{D}^n$ differing in one row and $\mathbf{s} \in \mathcal{S}_n$. Note that π is well-defined, since any outcome in $\mathcal{S}_n$ occurs with non-zero probability starting from any collection $\mathbf{x}$.

Differential privacy is typically concerned with bounding the privacy ratio over all original and synthetic collections. In particular, by Proposition 1.1, A is ϵ-differentially private if π is bounded everywhere on $\mathcal{S}_n$, with

$$\epsilon = \max_{\substack{\mathbf{s} \in \mathcal{S}_n \\ \mathbf{x}, \mathbf{x}' \in \mathcal{D}^n}} \log \pi(\mathbf{s}; \mathbf{x}, \mathbf{x}').$$

We take this further by studying the behaviour of the privacy ratio and how it varies across the support of A and across collections. This will allow us to understand in which situations it is at or near its bound.

Without loss of generality (in light of (2.2)), assume that the element differing between $\mathbf{x}$ and $\mathbf{x}'$ is the first one, and denote $\tilde{\mathbf{x}} = (x_2, \ldots, x_n)$. Then $\mathbf{x} = (x, \tilde{\mathbf{x}})$ and $\mathbf{x}' = (x', \tilde{\mathbf{x}})$. Also write $\mathbf{1}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$, the vector with a 1 in the j-th position and the rest 0, and let $\mathbf{s}_{-j} := \mathbf{s} - \mathbf{1}_j$.

Conditioning on the value of the modified record, the privacy ratio becomes

$$\pi(\mathbf{s}; \tilde{\mathbf{x}}, x, x') = \frac{\displaystyle\sum_{j=1}^{2^L} \mathsf{P}[A(x') = \mathbf{1}_j]\, \mathsf{P}[A(\tilde{\mathbf{x}}) = \mathbf{s}_{-i}]}{\displaystyle\sum_{j=1}^{2^L} \mathsf{P}[A(x) = \mathbf{1}_j]\, \mathsf{P}[A(\tilde{\mathbf{x}}) = \mathbf{s}_{-i}]} = \frac{\displaystyle\sum_{j=1}^{2^L} \mathsf{P}[R(x') = d_j]\, \mathsf{P}[A(\tilde{\mathbf{x}}) = \mathbf{s}_{-i}]}{\displaystyle\sum_{j=1}^{2^L} \mathsf{P}[R(x) = d_j]\, \mathsf{P}[A(\tilde{\mathbf{x}}) = \mathbf{s}_{-i}]}, \tag{3.1}$$

where we understand $\mathsf{P}[A(\tilde{\mathbf{x}}) = \mathbf{s}_{-i}] = 0$ if $\mathbf{s}_i = 0$. For convenience, we introduce the notation $p_j := \mathsf{P}[R(x) = d_j]$, $p'_j := \mathsf{P}[R(x') = d_j]$, and $P(\mathbf{s}, \mathbf{x}) := \mathsf{P}[A(\mathbf{x}) = \mathbf{s}]$, and express the privacy ratio as

$$\pi(\mathbf{s}; \mathbf{x}, x, x') = \frac{\displaystyle\sum_{j=1}^{2^L} p'_j P(\mathbf{s}_{-i}, \mathbf{x})}{\displaystyle\sum_{j=1}^{2^L} p_j P(\mathbf{s}_{-i}, \mathbf{x})},$$

dropping the tilde so that $\mathbf{x}$ now denotes a collection of size $n - 1$.

Denote $p_{ij} := \mathsf{P}[R(d_i) = d_j]$ and $P_n(\mathbf{s}, \mathbf{m}) := \mathsf{P}[A_n(\mathbf{m}) = \mathbf{s}]$.

3.1 The space $\mathcal{S}_n$

Recall that we consider the randomized query A a mapping from the simplex subset $\mathcal{S}_n$ to itself. Denote the standard basis vectors in $\mathbb{R}^{2^L}$ by $\mathbf{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$.

Viewing a point $\mathbf{s} = (s_1, \ldots, s_{2^L}) \in \mathcal{S}_n$ as a vector of counts, we consider as its "neighbours" those points $\mathbf{s}'$ for which one of the counts differs by 1: $|s_i - s'_i| = 1$ for some j. Because of the constraint $\sum_k s_k = \sum_k s'_k = n$, neighbouring points must differ in exactly two coordinates: $s_i - s'_i = 1 = s'_j - s_j$ for some $i \neq j$. In other words, neighbours of $\mathbf{s}$ are those points belonging to the set $\{\mathbf{s}' \in \mathcal{S}_n : \mathbf{s}' = \mathbf{s} + \mathbf{e}_j - \mathbf{e}_i, i \neq j\}$. We call the vector $\mathbf{e}_{ij} := \mathbf{e}_j - \mathbf{e}_i$ a **step from i to j**; note that adding a step from i to j to the counts vector $\mathbf{s}$ has the effect of shifting one count from bucket i to bucket j. The neighbours of $\mathbf{s}$ are then those points in $\mathcal{S}_n$ that can be reached in a single step. Note also that to step from i to j and remain in $\mathcal{S}_n$, $\mathbf{s}$ must have a non-zero count in bucket i.

In fact, the neighbours of $\mathbf{s}$ are also the closest points in $\mathcal{S}_n$ in the Euclidian sense. The (squared) Euclidian distance between two points $\mathbf{s}$ and $\mathbf{s}'$ is given by

$$\begin{aligned}\|\mathbf{s} - \mathbf{s}'\|_2^2 = \sum_{k=1}^{2^L}(s_k - s'_k)^2 &= \sum_{k=1}^{2^L-1}(s_k - s'_k)^2 + \left[\sum_{k=1}^{2^L-1}(s_k - s'_k)\right]^2 \\ &= \sum_{k=1}^{2^L-1} 2(s_k - s'_k)^2 + \sum_{1\leq k<\ell\leq 2^L-1}(s_k - s'_k)(s_\ell - s'_\ell),\end{aligned}$$

using the fact that $s_{2^L} = n - \sum_{k=1}^{2^L-1} s_k$. On one hand, the distance between neighbouring points is $\|\mathbf{s} - \mathbf{s}'\|_2^2 = 2$. On the other, this is the smallest possible distance between $\mathbf{s} \neq \mathbf{s}'$. Without loss of generality, assume $s_1 \neq s'_1$. Since $s_1, s'_1 \in \{0, \ldots, n\}$, the minimum distance occurs when $s_k = s'_k = 0$ for $k = 2, \ldots, 2^L - 1$, and $|s_k - s'_k| = 1$, i.e., $\mathbf{s}'$ is a neighbour of $\mathbf{s}$.

It is convenient to measure distance between points in $\mathcal{S}_n$ in terms of the number of steps required to reach one from the other. Define the metric δ_S on $\mathcal{S}_n$ as

$$\delta_S(\mathbf{s}, \mathbf{s}') := \frac{\|\mathbf{s} - \mathbf{s}'\|_1}{2} = \frac{1}{2}\left\{\sum_{k=1}^{2^L-1}|s_k - s'_k| + \left|\sum_{k=1}^{2^L-1}(s_k - s'_k)\right|\right\},$$

where $\|\cdot\|_1$ denotes the L_1 metric. Then:

- $\delta_s(\mathbf{s}, \mathbf{s}')$ is equal the minimum number of steps required to reach $\mathbf{s}'$ from $\mathbf{s}$.
- The neighbours of $\mathbf{s}$ are those points $\mathbf{s}' \in \mathcal{S}_n$ such that $\delta_S(\mathbf{s}, \mathbf{s}') = 1$.

3.2 The probability ratio

We observed previously that the privacy ratio can be expressed in terms of the ratio of probabilities of obtaining neighbouring synthetic collections starting from a common original collection. We formalize this notion as follows.

Definition. The **probability ratio** between neighbouring synthetic collections is given by

$$\rho_n(\mathbf{s}, i, j; \mathbf{m}) := \frac{\mathsf{P}[A_n(\mathbf{m}) = \mathbf{s}]}{\mathsf{P}[A_n(\mathbf{m}) = \mathbf{s} + \mathbf{e}_{ij}]} = \frac{P_n(\mathbf{s}, \mathbf{m})}{P_n(\mathbf{s} + \mathbf{e}_{ij}, \mathbf{m})}$$

for $\mathbf{s} \in \mathcal{S}_n$ such that $s_i \geq 1$.

The probability ratio describes how the likelihood of obtaining synthetic collections changes as we step between neighbouring synthetic outcomes in $\mathcal{S}_n$.

Note that, by conditioning on the randomization outcome of one of the vectors in the original collection, the probability ratio for a collection of size n can be expressed in terms of probability ratios for a collection of size $n-1$ over various synthetic collections. Denote $\mathbf{s}_{-i} := \mathbf{s} - \mathbf{e}_i$. Assume without loss of generality that $m_1 \geq 1$ (otherwise reorder the d_j). (TODO: are we assuming a specific ordering anywhere else?)

We have

$$\begin{aligned}
\rho_n(\mathbf{s}, i, j; \mathbf{m}) &= \frac{\displaystyle\sum_{k=1}^{2^L} p_{1k} P_{n-1}(\mathbf{s}_{-k}, \mathbf{m}_{-1})}{\displaystyle\sum_{k=1}^{2^L} p_{1k} P_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} \\
&= \frac{p_{11} \dfrac{P_{n-1}(\mathbf{s}_{-1}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} + \displaystyle\sum_{k=2}^{2^L} p_{1k} \frac{P_{n-1}(\mathbf{s}_{-k}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})}}{p_{11} + \displaystyle\sum_{k=2}^{2^L} p_{1k} \frac{P_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{ij}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})}}.
\end{aligned} \tag{3.2}$$

Observe that

$$\frac{P_{n-1}(\mathbf{s}_{-1}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} = \rho_{n-1}(\mathbf{s}_{-1}, i, j; \mathbf{m}_{-1}),$$

and for $k = 2, \ldots, 2^L$,

$$\begin{aligned}
\frac{P_{n-1}(\mathbf{s}_{-k}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} &= \frac{P_{n-1}(\mathbf{s}_{-k}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{1k}, \mathbf{m}_{-1})} \cdot \frac{P_{n-1}(\mathbf{s}_{-1}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} \\
&= \rho_{n-1}(\mathbf{s}_{-k}, 1, k; \mathbf{m}_{-1}) \cdot \rho_{n-1}(\mathbf{s}_{-1}, i, j; \mathbf{m}_{-1})
\end{aligned}$$

and

$$\frac{P_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{ij}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-1} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} = \frac{P_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{ij}, \mathbf{m}_{-1})}{P_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{1k} + \mathbf{e}_{ij}, \mathbf{m}_{-1})} = \rho_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{ij}, 1, k; \mathbf{m}_{-1}).$$

Applying these identities in (3.2) gives

$$\rho_n(\mathbf{s}, i, j; \mathbf{m}) = \frac{p_{11} \cdot \rho_{n-1}(\mathbf{s}_{-1}, i, j; \mathbf{m}_{-1}) + \displaystyle\sum_{k=2}^{2^L} p_{1k} \cdot \rho_{n-1}(\mathbf{s}_{-k}, 1, k; \mathbf{m}_{-1}) \rho_{n-1}(\mathbf{s}_{-1}, i, j; \mathbf{m}_{-1})}{p_{11} + \displaystyle\sum_{k=2}^{2^L} p_{1k} \cdot \rho_{n-1}(\mathbf{s}_{-k} + \mathbf{e}_{ij}, 1, k; \mathbf{m}_{-1})}.$$

TODO: case when $k = j$ or $k = i$?

TODO: generalize over the index we are conditioning on, and dividing by? Handle the case when $m_1 = 0$, or when any of the probabilities in the first line are zero.

An important property of the probability ratio is monotonicity. In particular, the ratio is nondecreasing when stepping in a direction that decreases the i-th count.

Proposition 3.1.

$$\rho_n(\mathbf{s}, i, j; \mathbf{m}) \leq \rho_n(\mathbf{s} + \mathbf{e}_{i\ell}, i, j; \mathbf{m})$$

for $\mathbf{s} \in \mathcal{S}_n$ *with* $s_i \geq 1$ *and* $\ell \neq i$.

3.3 Further stuff

TODO: Is this true?

Proposition 3.2. *Given an original collection consisting of* $(x, \mathbf{x})$*, the privacy ratio is maximized when the modification* x' *is chosen to be the opposite of* x*:*

$$\pi(\mathbf{s}; \mathbf{x}, x, x^*) \geq \pi(\mathbf{s}; \mathbf{x}, x, x')$$

for any $\mathbf{s} \in \mathcal{S}_n$*,* $x' \in \mathcal{D}$*, where* $\delta(x, x^*) = L$*.*

TODO: consequences of this:
- difference between numerator and denominator is basically a reweighting of the same terms in the sum
- max value of ratio - $(p/q)^L$ but doesn't depend on population size n

3.4 Local differential privacy

As noted above, differential privacy requires the privacy ratio to be uniformly bounded over all original and synthetic collections. This ensures that, regardless of the original collection, the randomization procedure used to generate the synthetic collection carries the same privacy guarantee.

However, applying the differential privacy criterion to our scenario of independently randomized bit vectors has a number of shortcomings. For one thing, the maximum value of the privacy ratio becomes infeasibly large when reporting large vectors, since it grows exponentially in L, the dimensionality of the reports. In order to bound this by what is commonly accepted as a reasonable value (TODO), the lie probability q needs to be set infeasibly high, reducing the utility of the collected data. Also, the maximum value does not depend on the population size n, meaning that no benefit is gained from having a larger population. Another issue is that the maximum value occurs for a synthetic collection which becomes increasing unlikely as the population size grows.

Hence, the high cost in terms of utility that we are incurring is mainly spent on protection for very unlikely outcomes.

We propose to remedy this by relaxing the differential privacy criterion (1.2) to hold for all but the most unlikely synthetic outcomes in $\mathcal{S}$. We say that A is ϵ-**locally differentially private** if

4 The case $L = 1$

It is instructive to first consider the case where each record in the collection consists of a single bit, as the expressions simplify considerably.

When $L = 1$, each original and synthetic record is either 1 or 0, and the transformation R flips each record with probability q. Partition the collection space $\mathcal{D}^n$ according to the number of records that are 1:

$$\mathcal{D}^n = \bigcup_{m=0}^{n} \mathcal{D}_m^n \qquad \text{where} \qquad \mathcal{D}_m^n := \Big\{ \mathbf{x} \in \mathcal{D}^n \,:\, \sum_{i=1}^{n} I(x_i = 1) = m \Big\}.$$

For $\mathbf{x} \in \mathcal{D}_m^n$, we have

$$A(\mathbf{x}) = \Phi \circ R(\mathbf{x}) = \big(A_n(m), n - A_n(m)\big),$$

where

$$\begin{aligned} A_n(m) &:= \sum_{i=1}^{n} I(R(x_i) = 1) = \sum_{i:\, x_i = 1} I(R(1) = 1) + \sum_{i:\, x_i = 0} I(R(0) = 1) \\ &\sim Bin(m, p) + Bin(n - m, q), \end{aligned}$$

a sum of two independent Binomial random variables with support $\{0, \ldots, n\}$. Furthermore, if $\mathbf{x} \in \mathcal{D}_m^n$ and $\mathbf{x}, \mathbf{x}'$ differ in one row, then $\mathbf{x}' \in \mathcal{D}_{m-1}^n \cup \mathcal{D}_{m+1}^n$. Defining

$$\pi_n(s; m) := \frac{\mathsf{P}[A_n(m) = s]}{\mathsf{P}[A_n(m+1) = s]} \qquad \text{for} \quad s \in \{0, \ldots, n\} \text{ and } m \in \{0, \ldots, n-1\},$$

the privacy ratio becomes

$$\pi\big((s, n-s);\, \mathbf{x}, \mathbf{x}'\big) = \begin{cases} \pi_n(s; m-1) & x_1 = 1 \\ \pi_n(s; m)^{-1} & x_1 = 0 \end{cases}.$$

Hence, in the $L = 1$ case, it suffices to study the behaviour of $\pi_n(s; m)$.

4.1 Recursive relationship over n and m

The conditioning argument (3.1) yields a recursive relationship that lets us express the distribution of A_n in terms of that of A_{n-1}.

Recall that $A_n(m)$ is the outcome of applying the bit transformation R to n original bits, m of which are 1 and $n-m$ are 0. For $m \geq 1$, we can condition on the outcome of one of the original 1s:

$$A_n(m) \sim Ber(p) + Bin(m-1,p) + Bin(n-m,q) \sim Ber(p) + A_{n-1}(m-1),$$

and so

$$\mathsf{P}[A_n(m) = s] = p\,\mathsf{P}[A_{n-1}(m-1) = s-1] + q\,\mathsf{P}[A_{n-1}(m-1) = s]. \tag{4.1}$$

If $s = 0$, the first term on the RHS is interpreted as 0, and if $s = n$, the last term is. Similarly, for $m \leq n-1$, conditioning on an original 0,

$$A_n(m) \sim Ber(q) + Bin(m,p) + Bin(n-m-1,q) \sim Ber(q) + A_{n-1}(m),$$

from which

$$\mathsf{P}[A_n(m) = s] = q\,\mathsf{P}[A_{n-1}(m) = s-1] + p\,\mathsf{P}[A_{n-1}(m) = s]. \tag{4.2}$$

The recursive formulas (4.1) and (4.2) give some insight into how the distribution of $A_n(m)$ changes as n and m vary:

- as n increases by 1, the probabilities shift slightly, with $\mathsf{P}[A_n(m) = 0] \leq \mathsf{P}[A_{n-1}(m) = 0]$ and $\mathsf{P}[A_n(m) = s]$ falling between $\mathsf{P}[A_{n-1}(m) = s-1]$ and $\mathsf{P}[A_{n-1}(m) = s]$ for each $s \geq 1$ (i.e., the hump of the pmf shifts to the right);
- the distribution of $A_n(m+1)$ is not so different to that of $A_n(m)$, since $\mathsf{P}[A_n(m) = s]$ and $\mathsf{P}[A_n(m+1) = s]$ both lie between consecutive pmf values of $A_{n-1}(m)$. In particular, this allows us to express the privacy ratio $\pi(s;m)$ in terms of $A_{n-1}(m)$.

Writing $P_{n,m}(s) := \mathsf{P}[A_n(m) = s]$, the formulas (4.1) and (4.2) can be expressed as

$$P_{n,m}(s) = pP_{n-1,m-1}(s-1) + qP_{n-1,m-1}(s) \quad \text{for } \; 0 \leq s \leq n, \;\; 1 \leq m \leq n$$

and

$$P_{n,m}(s) = qP_{n-1,m}(s-1) + pP_{n-1,m}(s) \quad \text{for } \; 0 \leq s \leq n, \;\; 0 \leq m \leq n-1.$$

4.2 The probability ratio

The probabilities in the privacy ratio represent the likelihood of observing the same synthetic collection outcome given two different original collections. In the expression $\pi_n(s;m) = P_{n,m}(s)/P_{n,m+1}(s)$, the probabilities correspond to the distributions of $A_n(m)$ and $A_n(m+1)$, respectively. However, using the decomposition (4.1) and (4.2), we can rewrite π_n in terms of probabilites from the same distribution, which is more convenient to work with.

Applying (4.2) to the numerator and (4.1) to the denominator, we obtain

$$\pi_n(s;m) = \frac{qP_{n-1,m}(s-1) + pP_{n-1,m}(s)}{pP_{n-1,m}(s-1) + qP_{n-1,m}(s)} = \frac{q + p\frac{P_{n-1,m}(s)}{P_{n-1,m}(s-1)}}{p + q\frac{P_{n-1,m}(s)}{P_{n-1,m}(s-1)}}$$

for $s \geq 1$, and $\pi_n(0;m) \equiv p/q$. Define the **probability ratio**

$$\rho_n(s;m) := \frac{P_{n,m}(s)}{P_{n,m}(s-1)} \qquad \text{for } \ 1 \leq s \leq n$$

a ratio of consecutive probabilities from the distribution of $A_n(m)$, and let $g(x) = \frac{q+px}{p+qx}$, so that $\pi_n = g \circ \rho_{n-1}$. The function g is increasing over $x > 0$, since

$$g'(x) = \frac{p-q}{(p+qx)^2} > 0.$$

Therefore, properties of monotonicity and extrema established for ρ_n (for all n) carry over to π_n as well.

The probability ratio can be expressed in a concise way using the following recursive property of the distribution of $A_n(m)$.

Lemma 4.1. *For $n \geq 1$,*

$$(s+1)P_{n,m}(s+1) = \left\{(m-s)\frac{p}{q} + (n-m-s)\frac{q}{p}\right\}P_{n,m}(s) + (n-s+1)P_{n,m}(s-1) \tag{4.3}$$

for $0 \leq m \leq n$ and $0 \leq s \leq n-1$ (with $P_{n,m}(-1) := 0$).

Proof. We proceed by induction on n. Suppose first $n = 1$, $s = 0$. If $m = 1$, then $A_1(1) \sim Ber(p)$, and (4.3) holds since $(mp/q + (1-m)q/p) \cdot P_{1,1}(0) = p = P_{1,1}(1)$. The argument is similar when $m = 0$. Next assume (4.3) holds for $A_{n-1}(m)$, and suppose $m \leq n-1$ and $1 \leq s \leq n-2$. Observe that

$$\begin{aligned}
&\left\{(m-s)\frac{p}{q} + (n-m-s)\frac{q}{p}\right\}P_{n,m}(s) + (n-s+1)P_{n,m}(s-1) \\
&\quad= \left\{(m-s)\frac{p}{q} + (n-1-m-s)\frac{q}{p}\right\}\big[qP_{n-1,m}(s-1) + pP_{n-1,m}(s)\big] \\
&\qquad\qquad + (n-1-s+1)\big[qP_{n-1,m}(s-2) + pP_{n-1,m}(s-1)\big] + \frac{q}{p}P_{n,m}(s) + P_{n,m}(s-1) \\
&\quad= p\left[\left\{(m-s)\frac{p}{q} + (n-1-m-s)\frac{q}{p}\right\}P_{n-1,m}(s) + (n-1-s+1)P_{n-1,m}(s-1)\right] \\
&\qquad + q\left[\left\{(m-(s-1))\frac{p}{q} + (n-1-m-(s-1))\frac{q}{p}\right\}P_{n-1,m}(s-1)\right. \\
&\qquad\qquad\qquad\qquad \left. + (n-1-(s-1)+1)P_{n-1,m}(s-2)\right] \\
&\qquad - \left(p + \frac{q^2}{p}\right)P_{n-1,m}(s-1) - qP_{n-1,m}(s-2) + \frac{q^2}{p}P_{n-1,m}(s-1) + qP_{n-1,m}(s) \\
&\qquad + qP_{n-1,m}(s-2) + pP_{n-1,m}(s-1) \\
&\quad= p(s+1)P_{n-1,m}(s+1) + qsP_{n-1,m}(s) + qP_{n-1,m}(s) \\
&\quad= (s+1)\big[qP_{n-1,m}(s) + pP_{n-1,m}(s+1)\big] = (s+1)P_{n,m}(s+1),
\end{aligned}$$

applying the induction hypothesis for s and for $s-1$ together with (4.2). If $s=0$, the argument is similar:

$$\begin{aligned}\left\{m\frac{p}{q}+(n-m)\frac{q}{p}\right\}P_{n,m}(0) &= p\left\{m\frac{p}{q}+(n-1-m)\frac{q}{p}\right\}P_{n-1,m}(0)+qP_{n-1,m}(0)\\ &= pP_{n-1,m}(1)+qP_{n-1,m}(0)=P_{n,m}(1).\end{aligned}$$

□

Given m, the probability ratio can be expressed using (4.3):

$$\begin{aligned}\rho(s+1;m) &= \frac{m-s}{s+1}\frac{p}{q}+\frac{n-m-s}{s+1}\frac{q}{p}+\frac{n-s+1}{s+1}\frac{1}{\rho(s;m)}\\ \rho(1;m) &= m\frac{p}{q}+(n-m)\frac{q}{p}\end{aligned}$$

Write

$$\eta(s) := \frac{n-s+1}{s+1} \qquad \text{and} \qquad \gamma_m(s) := \frac{1}{s+1}\left[(m-s)\frac{p}{q}+(n-m-s)\frac{q}{p}\right],$$

to get

$$\rho(s+1;m) = \eta(s)\rho(s;m)^{-1}+\gamma_m(s)\,; \quad \rho(1;m)=\gamma_m(0). \tag{4.4}$$

Note also that $\gamma_m(s)$ can be expressed in terms of $\mathsf{E}\,A_n(m)=\mu_m=nq+m(p-q)$:

$$(s+1)\gamma_m(s) = \frac{\mu_m-s}{pq}-n+2s.$$

The probability ratio has the following properties (TODO):

- decreasing in s for fixed m
- increasing in m for fixed s.

4.3 Bounding the probability ratio

For A to satisfy local differential privacy, the privacy ratio $\pi_n(s;m)$ must be bounded for all s except for a set of small probability with respect to the distribution $\mathsf{P}[A_n(m)=\cdot\,]$. Furthermore, this bound must hold regardless of the original collection described through m.

Fix $\delta>0$. Given m, we show that the probability ratio for $s\in[\mu_m-\delta,n]$ is bounded by a value $\rho(s^*;0)$, where s^* is expressed in terms of $\mu_0-\delta$. Together with the fact that $P_{n,m}(\mu_m-\delta)\le P_{n,0}(\mu_0-\delta)$ (TODO - is this necessary?), this implies that the bound for local differential privacy, required to hold for all m, can be computed in terms of $A_n(0)$ alone. Note that, since ρ is decreasing in s for fixed m, it is sufficient to consider the probability ratio at the smallest integer value belonging to the interval $[\mu_m-\delta,n]$.

TODO: how to handle the left endpoint. What is the min value of δ?

For $\delta > 0$ let $s_m(\delta) := \lceil \mu_m - \delta \rceil \vee 0$, and define $R_m(\delta) := \rho(s_m(\delta); m)$. Note that $R_m(\delta) \le R_m(\delta')$ for $\delta \le \delta'$, and $s_m(\delta + 1) = (s_m(\delta) - 1) \vee 0$.

Proposition 4.1.

$$R_m(\delta) \le R_0(\delta + 2) \qquad \textit{for} \;\; m = 0, \ldots, n$$

provided $\delta > \sigma_0 + 1$, *where* $\sigma_0^2 = \mathsf{Var}\, A_n(0) = npq$.

Proof. Fix $\delta > \sigma_0 + 1$. (TODO) If $s_0(\delta) < 2$

Assume $s_0(\delta) \ge 2$, and suppose $R_m(\delta) > R_0(\delta + 2)$ for some m. Then, we have

$$R_0(\delta) \le R_0(\delta + 1) \le R_0(\delta + 2) < R_m(\delta) \le R_m(\delta + 1),$$

implying that

$$R_0(\delta + 2)^{-1} = \frac{R_0(\delta + 1) - \gamma_0(s_0(\delta + 2))}{\eta(s_0(\delta + 2))} > \frac{R_m(\delta) - \gamma_m(s_m(\delta + 1))}{\eta(s_m(\delta + 1))} = R_m(\delta + 1)^{-1}$$

via (4.4). Write $s_m := s_m(\delta + 1)$, $s_0 := s_0(\delta + 2)$. Since $R_m(\delta) > R_0(\delta + 1)$ by assumption, we obtain:

$$\big\{\eta(s_m) - \eta(s_0)\big\} R_0(\delta + 1) + \big\{\eta(s_0)\gamma_m(s_m) - \eta(s_m)\gamma_0(s_0)\big\} > 0. \tag{4.5}$$

Furthermore,

$$\eta(s_m) - \eta(s_0) = \frac{n - s_m + 1}{s_m + 1} - \frac{n - s_0 + 1}{s_0 + 1} = -\frac{(n + 2)(s_m - s_0)}{(s_0 + 1)(s_m + 1)},$$

and

$$\begin{aligned}
&\eta(s_0)\gamma_m(s_m) - \eta(s_m)\gamma_0(s_0) \\
&\quad = \frac{n - s_0 + 1}{s_0 + 1} \cdot \frac{(\mu_m - s_m)/pq - n + 2s_m}{s_m + 1} - \frac{n - s_m + 1}{s_m + 1} \cdot \frac{(\mu_0 - s_0)/pq - n + 2s_0}{s_0 + 1} \\
&\quad = \frac{(n + 2)(s_m - s_0) + (\mu_0 s_m - \mu_m s_0)/pq + (n + 1)[\mu_m - \mu_0 - (s_m - s_0)]/pq}{(s_0 + 1)(s_m + 1)},
\end{aligned}$$

so (4.5) implies

$$\begin{aligned}
&-(n + 2)(s_m - s_0)(R_0(\delta + 1) - 1) + \\
&\qquad \frac{\mu_0(s_m - s_0) - m(p - q)s_0}{pq} + \frac{(n + 1)[m(p - q) - (s_m - s_0)]}{pq} > 0.
\end{aligned} \tag{4.6}$$

Now, let $\delta_0 := \delta - \{\lceil \mu_0 - \delta \rceil - (\mu_0 - \delta)\} = \mu_0 - \lceil \mu_0 - \delta \rceil$, i.e., $\delta_0 = \inf\{\lambda : s_0(\lambda) = s_0(\delta)\}$. Then $s_0(\delta) = s_0(\delta_0) = \mu_0 - \delta_0$, an integer, and $s_m(\delta_0) - s_m(\delta) \in \{0, 1\}$, since $0 \le \delta - \delta_0 < 1$. Consequently, since

$$s_m(\delta_0) - s_0(\delta_0) = \lceil \mu_0 + m(p - q) - \delta_0 \rceil - (\mu_0 - \delta_0) = \lceil m(p - q) \rceil,$$

$$\begin{aligned} s_m - s_0 &= (s_m(\delta) - 1) - (s_0(\delta) - 2) = s_m(\delta) - s_m(\delta_0) + \lceil m(p-q) \rceil + 1 \\ &\in \big\{ \lceil m(p-q) \rceil, \lceil m(p-q) \rceil + 1 \big\}, \end{aligned}$$

and

$$\begin{aligned} \mu_0(s_m - s_0) - m(p-q)s_0 &= \mu_0(s_m - s_0) - m(p-q)(s_0(\delta_0) - 2) \\ &= \mu_0[(s_m - s_0) - m(p-q)] + m(p-q)(\delta_0 + 2). \end{aligned}$$

Applying these identities in (4.6) gives

$$-(n+2)(s_m - s_0)(R_0(\delta+1) - 1) + \frac{m(p-q)(\delta_0+2)}{pq} + \frac{n+1-\mu_0}{pq}\big(m(p-q) - (s_m - s_0)\big) > 0.$$

Since $s_m - s_0 \geq m(p-q)$,

$$(n+2)(R_0(\delta+1) - 1) < \frac{\delta_0 + 2}{pq} \frac{m(p-q)}{s_m - s_0} < \frac{\delta_0 + 2}{pq}. \tag{4.7}$$

Next, recall that $R_0(\delta+1) = P_{n,0}(s_0(\delta+1))/P_{n,0}(s_0(\delta+1) - 1)$. Since $P_{n,0}(\cdot) = \mathsf{P}[Bin(n,q) = \cdot\,]$,

$$R_0(\delta+1) - 1 = \frac{n - s_0(\delta+1) + 1}{s_0(\delta+1)} \cdot \frac{q}{p} - 1 = \frac{\mu_0 - (\mu_0 - \delta_0 - 1) + q}{p(\mu_0 - \delta_0 - 1)} = \frac{\delta_0 + q + 1}{p(\mu_0 - \delta_0 - 1)}.$$

Hence, substituting this expression in (4.7) yields

$$\begin{aligned} & (\delta_0+2)(\mu_0 - \delta_0 - 1) > (n+2)q(\delta_0 + q + 1) > \mu_0(\delta_0 + q + 1) \\ \iff & -\delta_0^2 - 3\delta_0 + 2\mu_0 - 2 > (1+q)\mu_0 \\ \iff & -\delta_0^2 - 3\delta_0 + npq > 0, \end{aligned}$$

which requires that δ_0 lie between the roots of the quadratic equation. In particular,

$$\delta_0 \leq -\frac{3}{2} + \frac{1}{2}\sqrt{9 + 4npq} \leq -\frac{3}{2} + \frac{3}{2} + \sqrt{npq} = \sqrt{npq}.$$

Finally, since $0 \leq \delta - \delta_0 < 1$, we conclude that

$$\delta = \delta_0 + \delta - \delta_0 < \sigma_0 + 1,$$

contradicting our initial choice of δ. □

4.4 Recursive relationship for CDFs

Summing both sides of (4.1) and (4.2) shows that the recursive relationship extends to cdfs as well. Writing $F_{n,m}(x) := \mathsf{P}[A_n(m) \leq x]$, we have

$$F_{n,m}(x) = p \cdot F_{n-1,m-1}(x-1) + q \cdot F_{n-1,m-1}(x) \qquad\qquad m = 1, \ldots, n$$

and

$$F_{n,m}(x) = q \cdot F_{n-1,m}(x-1) + p \cdot F_{n-1,m}(x) \qquad m = 0, \ldots, n-1.$$

Furthermore, we can express the difference on incrementing m as follows:

$$\begin{aligned} &F_{n,m}(x) - F_{n,m-1}(x) \\ &\quad = pF_{n-1,m-1}(x-1) + qF_{n-1,m-1}(x) - qF_{n-1,m-1}(x-1) - pF_{n-1,m-1}(x) \\ &\quad = (p-q)\big[F_{n-1,m-1}(x-1) - F_{n-1,m-1}(x)\big] \\ &\quad = -(p-q)P_{n-1,m-1}(\lfloor x \rfloor) \end{aligned}$$

and

$$\begin{aligned} &F_{n,m}(x+1) - F_{n,m-1}(x) \\ &\quad = pF_{n-1,m-1}(x) + qF_{n-1,m-1}(x+1) - qF_{n-1,m-1}(x-1) - pF_{n-1,m-1}(x) \\ &\quad = q\big[F_{n-1,m-1}(x+1) - F_{n-1,m-1}(x-1)\big] \\ &\quad = q\big\{P_{n-1,m-1}(\lfloor x \rfloor) + P_{n-1,m-1}(\lfloor x+1 \rfloor)\big\} \end{aligned}$$

Therefore, $F_{n,m}(x) \le F_{n,m-1}(x) \le F_{n,m}(x+1)$ and $F_{n,m-1}(x-1) \le F_{n,m}(x) \le F_{n,m-1}(x)$.

For probabilities, the relationship becomes

$$P_{n,m}(s) - P_{n,m-1}(s) = -(p-q)\big\{P_{n-1,m-1}(s) - P_{n-1,m-1}(s-1)\big\}$$

and

$$P_{n,m}(s+1) - P_{n,m-1}(s) = q\big\{P_{n-1,m-1}(s+1) - P_{n-1,m-1}(s-1)\big\}.$$

4.5 Privacy Ratio

Using the recursive relationships (4.1) and (4.2), we have

$$\pi(s;m) = \frac{\mathsf{P}[A_n(m) = s]}{\mathsf{P}[A_n(m+1) = s]} = \frac{qP_{n-1,m}(s-1) + pP_{n-1,m}(s)}{pP_{n-1,m}(s-1) + qP_{n-1,m}(s)}.$$

Lemma 6.2 implies that monotonicity and extrema of the privacy ratio π are determined by those of the *probability ratio* $P_{n-1,m}(s)/P_{n-1,m}(s-1)$:

$$\text{if} \quad \frac{P_{n-1,m}(s)}{P_{n-1,m}(s-1)} \ge \frac{P_{n-1,m'}(s')}{P_{n-1,m'}(s'-1)}, \quad \text{then} \quad \pi_n(s;m) \ge \pi_n(s';m').$$

Hence, studying the behaviour of the privacy ratio reduces to studying ratios of consecutive pmf values. Using this property, we can derive the following facts:

- π is decreasing in s: $\pi(s;m) \ge \pi(s+1;m)$
- π is increasing in m: $\pi(s;m) \le \pi(s;m+1)$

- π is net decreasing when both s and m increase: $\pi(s;m) \geq \pi(s+1;m+1)$.

A different approach, using the difference identities above, gives

$$\begin{aligned}\pi(s;m) = \frac{P_{n,m}(s)}{P_{n,m+1}(s)} &= \frac{P_{n,m-1}(s) + (p-q)\{P_{n-1,m-1}(s-1) - P_{n-1,m-1}(s)\}}{P_{n,m}(s) + (p-q)\{P_{n-1,m}(s-1) - P_{n-1,m}(s)\}} \\ &= \frac{\frac{P_{n,m-1}(s)}{P_{n,m}(s)} + (p-q)\frac{P_{n-1,m-1}(s-1) - P_{n-1,m-1}(s)}{P_{n,m}(s)}}{1 + (p-q)\frac{P_{n-1,m}(s-1) - P_{n-1,m}(s)}{P_{n,m}(s)}}.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\pi(s;m)}{\pi(s;m-1)} &= \frac{P_{n,m}(s)}{P_{n,m-1}(s)}\pi(s;m) \\ &= \frac{1 + (p-q)\frac{P_{n-1,m-1}(s-1) - P_{n-1,m-1}(s)}{P_{n,m-1}(s)}}{1 + (p-q)\frac{P_{n-1,m}(s-1) - P_{n-1,m}(s)}{P_{n,m}(s)}}\end{aligned}$$

4.6 Quantiles

Denote by $\tau_n(m)$ the α-th quantile of $F_{n,m}$:

$$\tau_n(m) = \inf\{y : F_{n,m}(y) \geq \alpha\}.$$

Note that $\tau_n(m)$ is an integer satisfying $F_{n,m}(\tau_n(m)) \geq \alpha$ and $F_{n,m}(\tau_n(m)-1) < \alpha$.

Claim. *For $m = 1, \ldots, n$,*

$$\tau_n(m) \in \{\tau_n(m-1), \tau_n(m-1)+1\}.$$

Proof. Since $\alpha \leq F_{n,m}(\tau_n(m)) \leq F_{n,m-1}(\tau_n(m))$, we have $\tau_n(m) \geq \tau_n(m-1)$. On the other hand, $F_{n,m}(\tau_n(m-1)+1) \geq F_{n,m-1}(\tau_n(m-1)) \geq \alpha$, from which $\tau_n(m) \leq \tau_n(m-1)+1$. □

Furthermore, $\tau_n(m) = \tau_n(m-1)$ if and only if $F_{n,m}(\tau_n(m-1)) \geq \alpha$, i.e.,

$$F_{n,m-1}(\tau_n(m-1)) - \alpha \geq (p-q)\,\mathsf{P}[A_{n-1}(m-1) = \tau_n(m-1)].$$

4.7 Asymptotic approach

4.8 Approximate location relationship over m

The family $\{A_n(m),\ m = 0, \ldots, n\}$ are in fact approximately location-shifted versions of each other. Indeed, note that

$$\mathsf{E}\, A_n(m) = mp + (n-m)q = nq + m(p-q)$$

and

$$\mathsf{Var}\, A_n(m) = mpq + (n-m)qp = npq.$$

In other words, as m varies between 0 and n, the mean varies linearly in m and the variance remains constant. The central portion of the distribution remains approximately the same shape, although transitioning from right-skewed when $m = 0$ to left-skewed when $m = n$. Thus,

$$F_{n,m}(x) \approx F_{n,0}(x - m(p-q)).$$

We quantify this approximation as follows.

5 Maximal collections

The first question we address is for which pair of original and modified collections $\mathbf{x}$ and $\mathbf{x}'$ does π obtain its maximum.

5.1 The case $L = 1$

It is instructive to first consider the case where each record in the collection consists of a single bit, as the expressions simplify considerably. In this case, the outcome of A essentially reduces to the number of 1s obtained in the synthetic collection $R(\mathbf{x})$, since $\mathbf{s} \in \mathcal{S}_n$ can be written as (s_1, s_2) with $s_1 \in \{0, \ldots, n\}$ and $s_2 = n - s_1$. Using this fact, we interpret $A(\mathbf{x})$ as $\sum_{i=1}^{n} I(R(x_i) = 1)$, and express the privacy ratio as

$$\pi(s, \mathbf{x}, x_1') = \frac{\mathsf{P}[R(x_1') = 1]\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s-1] + \mathsf{P}[R(x_1') = 0]\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s]}{\mathsf{P}[R(x_1) = 1]\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s-1] + \mathsf{P}[R(x_1) = 0]\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s]}$$

for $s \in \{0, \ldots, n\}$. Furthermore, the requirement that $x_1 \neq x_1'$ in the single-bit case implies that $x_1' = 1 - x_1$, so

$$\pi(s, \mathbf{x}, x_1') = \frac{(1 - \mathsf{P}[R(x_1) = 1])\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s-1] + (1 - \mathsf{P}[R(x_1) = 0])\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s]}{\mathsf{P}[R(x_1) = 1]\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s-1] + \mathsf{P}[R(x_1) = 0]\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s]}.$$

We fix $s \in \{1, \ldots, n\}$ and investigate which choice of collection $\mathbf{x} = (x_1, \tilde{\mathbf{x}})$ maximizes $\pi(s, \mathbf{x}, x_1')$. Assume first $x_1 = 1$. (TODO: what about when $x_1 = 0$?) Then

$$\pi(s, \tilde{\mathbf{x}}, 0) = \frac{q\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s-1] + p\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s]}{p\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s-1] + q\,\mathsf{P}[A(\tilde{\mathbf{x}}) = s]}.$$

By Lemma 6.2, this ratio is maximized at $\tilde{\mathbf{x}}^*$ satisfying

$$\frac{\mathsf{P}[A(\tilde{\mathbf{x}}^*) = s]}{\mathsf{P}[A(\tilde{\mathbf{x}}^*) = s-1]} \geq \frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s]}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-1]}$$

for any $\tilde{\mathbf{y}} \in \mathcal{D}^{n-1}$. We claim that this is the case when $\tilde{\mathbf{x}}^*$ consists of all 1s:

Claim. *Write* $\mathbf{1} = (1, ..., 1)$ *as an element of* $\mathcal{D}^n$*. Then, given* $s \in \{1, \ldots, n\}$*,*

$$\frac{\mathsf{P}[A(\mathbf{1}) = s]}{\mathsf{P}[A(\mathbf{1}) = s-1]} \geq \frac{\mathsf{P}[A(\mathbf{y}) = s]}{\mathsf{P}[A(\mathbf{y}) = s-1]} \tag{5.1}$$

for any $\mathbf{y} \in \mathcal{D}^n$*.*

Proof. We proceed by induction on n. If $n = 1$, then $A(x) = R(x)$, we need only confirm (5.1) for $s = 1$. Taking $\mathbf{y} = 0$ (the only possibility aside from $\mathbf{1}$), we have

$$\frac{\mathsf{P}[R(0) = 1]}{\mathsf{P}[R(0) = 0]} = \frac{q}{p} \leq \frac{p}{q} = \frac{\mathsf{P}[R(1) = 1]}{\mathsf{P}[R(1) = 0]} \tag{5.2}$$

verifying (5.1) in this case. Next, suppose that (5.1) holds for $s \in \{1, \ldots, n-1\}$ and $\mathbf{1}, \tilde{\mathbf{y}} \in \mathcal{D}^{n-1}$, and consider $\mathbf{y} = (\tilde{\mathbf{y}}, y_n) \in \mathcal{D}^n$. For convenience, write $p(y_n) = \mathsf{P}[R(y_n) = 1]$. Observe that

$$\mathsf{P}[A(\mathbf{y}) = s] = \mathsf{P}[A(\tilde{\mathbf{y}}) = s] \cdot (1 - p(y_n)) + \mathsf{P}[A(\tilde{\mathbf{y}}) = s-1] \cdot p(y_n),$$

conditioning on the value of y_n. Thus,

$$\frac{\mathsf{P}[A(\mathbf{y}) = s]}{\mathsf{P}[A(\mathbf{y}) = s-1]} = \frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s] \cdot (1 - p(y_n)) + \mathsf{P}[A(\tilde{\mathbf{y}}) = s-1] \cdot p(y_n)}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-1] \cdot (1 - p(y_n)) + \mathsf{P}[A(\tilde{\mathbf{y}}) = s-2] \cdot p(y_n)}. \tag{5.3}$$

For $s \in \{2, \ldots, n-1\}$, our induction hypothesis implies that

$$\frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s]}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-1]} \leq \frac{\mathsf{P}[A(\mathbf{1}) = s]}{\mathsf{P}[A(\mathbf{1}) = s-1]} \quad \text{and} \quad \frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-1]}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-2]} \leq \frac{\mathsf{P}[A(\mathbf{1}) = s-1]}{\mathsf{P}[A(\mathbf{1}) = s-2]},$$

and furthermore,

$$\frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s]}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-2]} = \frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s]}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-1]} \cdot \frac{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-1]}{\mathsf{P}[A(\tilde{\mathbf{y}}) = s-2]} \leq \frac{\mathsf{P}[A(\mathbf{1}) = s]}{\mathsf{P}[A(\mathbf{1}) = s-2]}.$$

Therefore, we can apply Lemma 6.3 to (5.3) to obtain

$$\frac{\mathsf{P}[A(\mathbf{y}) = s]}{\mathsf{P}[A(\mathbf{y}) = s-1]} \leq \frac{\mathsf{P}[A(\mathbf{1}) = s] \cdot (1 - p(y_n)) + \mathsf{P}[A(\mathbf{1}) = s-1] \cdot p(y_n)}{\mathsf{P}[A(\mathbf{1}) = s-1] \cdot (1 - p(y_n)) + \mathsf{P}[A(\mathbf{1}) = s-2] \cdot p(y_n)} \tag{5.4}$$

for $s \in \{2, \ldots, n-1\}$. If $s = n$, we consider $\mathsf{P}[A(\tilde{\mathbf{y}}) = s] = 0$ since $\mathsf{P}[A(\tilde{\mathbf{y}}) \in \{0, \ldots, n-1\}] = 1$, and similarly for $\mathsf{P}[A(\mathbf{1}) = s]$. In this case, Lemma 6.3 still applies with $b = b' = 0$. A similar argument establishes (5.4) when $s = 1$. Therefore,

$$\frac{\mathsf{P}[A(\mathbf{y}) = s]}{\mathsf{P}[A(\mathbf{y}) = s-1]} \leq \frac{\mathsf{P}[A((\mathbf{1}, y_n)) = s]}{\mathsf{P}[A((\mathbf{1}, y_n)) = s-1]}$$

for $s \in \{1, \ldots, n\}$. Since $y_n \in \{0, 1\}$, the proof will be complete if we show that

$$\begin{aligned}\frac{\mathsf{P}[A((\mathbf{1}, 1)) = s]}{\mathsf{P}[A((\mathbf{1}, 1)) = s-1]} &= \frac{\mathsf{P}[A(\mathbf{1}) = s] \cdot q + \mathsf{P}[A(\mathbf{1}) = s-1] \cdot p}{\mathsf{P}[A(\mathbf{1}) = s-1] \cdot q + \mathsf{P}[A(\mathbf{1}) = s-2] \cdot p} \\ &\geq \frac{\mathsf{P}[A(\mathbf{1}) = s] \cdot p + \mathsf{P}[A(\mathbf{1}) = s-1] \cdot q}{\mathsf{P}[A(\mathbf{1}) = s-1] \cdot p + \mathsf{P}[A(\mathbf{1}) = s-2] \cdot q} = \frac{\mathsf{P}[A((\mathbf{1}, 0)) = s]}{\mathsf{P}[A((\mathbf{1}, 0)) = s-1]}. \end{aligned} \tag{5.5}$$

Using the fact that $A(\mathbf{1}) \sim Bin(n-1, p)$, observe that

$$\frac{\mathsf{P}[A(\mathbf{1}) = s]}{\mathsf{P}[A(\mathbf{1}) = s-1]} = \frac{\binom{n-1}{s} p^s q^{n-1-s}}{\binom{n-1}{s-1} p^{s-1} q^{n-s}} = \frac{n-s}{s}\frac{p}{q},$$

from which

$$\frac{\lambda_2}{\mu_2} = \frac{\mathsf{P}[A(\mathbf{1}) = s]}{\mathsf{P}[A(\mathbf{1}) = s-1]} = \frac{n-s}{s}\frac{p}{q} \leq \frac{n-s+1}{s-1}\frac{p}{q} = \frac{\mathsf{P}[A(\mathbf{1}) = s-1]}{\mathsf{P}[A(\mathbf{1}) = s-2]} = \frac{\lambda_1}{\mu_1},$$

and hence (5.5) follows from Lemma 6.2 (unless $s = 1$, in which case it follows from a simple direct argument). □

TODO: finish argument in the case when $x_1 = 0$.

6 Ratios of sums: properties

Here we establish some results around bounding and comparing ratios of sums, which will be useful in working with the privacy ratio.

Lemma 6.1. *Suppose $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}$ with $b_i > 0$ all i. Then*

$$\frac{a_1 + \cdots + a_m}{b_1 + \cdots + b_m} \leq \max\left(\frac{a_1}{b_1}, \ldots, \frac{a_m}{b_m}\right).$$

Proof. Write

$$\frac{a_1 + \cdots + a_m}{b_1 + \cdots + b_m} = \frac{a_1}{b_1}\frac{b_1}{b_1 + \cdots + b_m} + \cdots + \frac{a_m}{b_m}\frac{b_m}{b_1 + \cdots + b_m} = \sum_{i=1}^{m} \frac{a_i}{b_i}\lambda_i$$

where $\lambda_1 + \cdots + \lambda_m = 1$. The result follows since each a_i/b_i is bounded by $\max_i a_i/b_i$. □

Lemma 6.2. *Suppose $a_i, a_i', \lambda_i, \mu_i > 0$ for $i = 1, \ldots, m$. Then*

$$\frac{a_1\lambda_1 + \cdots + a_m\lambda_m}{a_1\mu_1 + \cdots + a_m\mu_m} \geq \frac{a_1'\lambda_1 + \cdots + a_m'\lambda_m}{a_1'\mu_1 + \cdots + a_m'\mu_m} \tag{6.1}$$

if

$$\lambda_i/\mu_i \geq \lambda_j/\mu_j \tag{6.2}$$

and

$$a_i/a_j \geq a_i'/a_j' \tag{6.3}$$

whenever $1 \leq i < j \leq m$.

Note that numerator and denominator have the same index (6.2), and different indices in (6.3). It is easy to see that (6.2) is satisfied when $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ and $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_m$. It is also clear from the proof that (6.1) still holds if all the inequalities in (6.2) and (6.3) are reversed. An important special case for us is the following:

Corollary 6.1. *Suppose* $0 < q < p < 1$ *and* $a_i, a_i' > 0$ *for* $i = 1, 2$. *Then*

$$\frac{q \cdot a_1 + p \cdot a_2}{p \cdot a_1 + q \cdot a_2} \geq \frac{q \cdot a_1' + p \cdot a_2'}{p \cdot a_1' + q \cdot a_2'}$$

whenever

$$\frac{a_2}{a_1} \geq \frac{a_2'}{a_1'}.$$

Proof. Cross-multiplying, we see that (6.1) is equivalent to

$$\sum_i \sum_j a_i a_j' \lambda_i \mu_j \geq \sum_i \sum_j a_i a_j' \lambda_j \mu_i \iff \sum_i \sum_{j \neq i} a_i a_j' (\lambda_i \mu_j - \lambda_j \mu_i) \geq 0.$$

Furthermore,

$$\begin{aligned}
\sum_i \sum_{j \neq i} a_i a_j' (\lambda_i \mu_j - \lambda_j \mu_i) &= \sum_i \sum_{j > i} a_i a_j' (\lambda_i \mu_j - \lambda_j \mu_i) + \sum_i \sum_{j < i} a_i a_j' (\lambda_i \mu_j - \lambda_j \mu_i) \\
&= \sum_i \sum_{j > i} a_i a_j' (\lambda_i \mu_j - \lambda_j \mu_i) + \sum_j \sum_{i < j} a_j a_i' (\lambda_j \mu_i - \lambda_i \mu_j) \\
&= \sum_i \sum_{j > i} a_i a_j' (\lambda_i \mu_j - \lambda_j \mu_i) + \sum_i \sum_{j > i} a_j a_i' (\lambda_j \mu_i - \lambda_i \mu_j) \\
&= \sum_i \sum_{j > i} (a_i a_j' - a_j a_i') (\lambda_i \mu_j - \lambda_j \mu_i),
\end{aligned}$$

where the second equality follows from relabeling the summation indices, and the third from reversing the sums. It follows that (6.1) will hold if $(a_i a_j' - a_j a_i')(\lambda_i \mu_j - \lambda_j \mu_i) \geq 0$ for all $1 \leq i < j \leq m$, which is implied by (6.2) and (6.3). □

Lemma 6.3. *Suppose* $a, a', \lambda, \mu > 0$, *and* $b, b', c, c' \geq 0$. *Then*

$$\frac{a\lambda + b\mu}{c\lambda + a\mu} \geq \frac{a'\lambda + b'\mu}{c'\lambda + a'\mu} \tag{6.4}$$

if

$$ac' \geq a'c, \quad ab' \leq a'b, \quad \text{and} \quad bc' \geq b'c. \tag{6.5}$$

Proof. (6.4) holds iff

$$\begin{aligned}
ac'\lambda^2 + aa'\lambda\mu + bc'\lambda\mu + a'b\mu^2 &\geq a'c\lambda^2 + b'c\lambda\mu + aa'\lambda\mu + ab'\mu^2 \\
\iff (ac' - a'c)\lambda^2 + (bc' - b'c)\lambda\mu + (a'b - ab')\mu^2 &\geq 0,
\end{aligned}$$

which is implied by (6.5). □

7 Old stuff

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