Automatic, Fine-Grained Algorithmic Choice for Differential Privacy

Theorem 1. *The entropy function on disjoint histogram counts $a_1, a_2, \ldots, a_n$ has sensitivity*

Proof. Let $A = \sum_{i=1}^{n} a_i$. Then, the entropy is

$$\sum_{i=1}^{n} \frac{a_i}{A} \log\left(\frac{A}{a_i}\right) = \frac{1}{A}\sum_{i=1}^{n} a_i \log A - \frac{1}{A}\sum_{i=1}^{n} a_i \log(a_i) = \log(A) - \frac{1}{A}\sum_{i=1}^{n} a_i \log(a_i)$$

Suppose bucket a_j is reduced by 1, and the entropy change is

$$\begin{aligned}
&\log(A) - \log(A-1) - \frac{1}{A} a_j \log(a_j) + \frac{1}{A-1}(a_j - 1)\log(a_j - 1) \\
&\leq \frac{1}{\ln(2)(A-1)} - \frac{1}{A}(a_j - 1)\log(a_j - 1) + \frac{1}{A-1}(a_j - 1)\log(a_j - 1) \\
&= \frac{1}{\ln(2)(A-1)} + \frac{1}{A(A-1)}(a_j - 1)\log(a_j - 1) \leq \frac{1}{\ln(2)(A-1)} + \frac{1}{A}\log(A)
\end{aligned}$$

□