

HW2-1 Report

Laplace Mechanism

1. Task (a) & (b) for $\epsilon = 0.5$

S = sensitivity of the dataset

$$b = \frac{S}{\epsilon}$$

$$\text{error} = 2b^2 = 2 * \left(\frac{S}{\epsilon}\right)^2$$

```
Enter full path for the dataset file
adult.data.txt
For epsilon = 0.5 and query = 1 -> S = 3.8240916E-5 , b = 7.648183E-5 , Error = 1.1698941364112996E-8
For epsilon = 0.5 and query = 2 -> S = 3.0732353E-5 , b = 6.146471E-5 , Error = 7.555820403275854E-9
For epsilon = 0.5 and query = 3 -> S = 3.1451487E-5 , b = 6.2902975E-5 , Error = 7.913568428338363E-9
For epsilon = 0.5 and query = 4 -> S = 3.0899482E-5 , b = 6.1798964E-5 , Error = 7.638224014793851E-9
GS = 4.2961617
Validating that each of the last 3 groups of results and the first group are 0.5-indistinguishable.
0.25 <= 0.4121803
0.25 <= 0.4121803
0.25 <= 0.40146363
```

The global sensitivity for the first 4 groups with $\epsilon = 0.5$ is 4.296

- Comparing the 1st group with 2nd group:
 $Pr[A(D) \in S] \leq e^\epsilon Pr[A(D') \in S] = 0.25 \leq 0.412$
Thus, we can validate that the 2 sets are 0.5-distinguishable
- Comparing the 1st group with 3rd group:
 $Pr[A(D) \in S] \leq e^\epsilon Pr[A(D') \in S] = 0.25 \leq 0.412$
Thus, we can validate that the 2 sets are 0.5-distinguishable
- Comparing the 1st group with 4th group:
 $Pr[A(D) \in S] \leq e^\epsilon Pr[A(D') \in S] = 0.25 \leq 0.401$
Thus, we can validate that the 2 sets are 0.5-distinguishable

1. Task(c) for $\epsilon = 1$

```
For epsilon = 1.0 and query = 1 -> S = 3.8240916E-5 , b = 3.8240916E-5 , Error = 2.924735341028249E-9
For epsilon = 1.0 and query = 2 -> S = 3.0732353E-5 , b = 3.0732353E-5 , Error = 1.8889551008189636E-9
For epsilon = 1.0 and query = 3 -> S = 3.1451487E-5 , b = 3.1451487E-5 , Error = 1.9783921070845907E-9
For epsilon = 1.0 and query = 4 -> S = 3.0899482E-5 , b = 3.0899482E-5 , Error = 1.909556003698463E-9
GS = 4.2961617
Validating that each of the last 3 groups of results and the first group are 1.0-indistinguishable.
0.25 <= 0.67957044
0.25 <= 0.67957044
0.25 <= 0.67821133
```

The global sensitivity for the first 4 groups with $\varepsilon = 1$ is 4.296

- Comparing the 1st group with 2nd group:
 $Pr[A(D) \in S] \leq e^\varepsilon Pr[A(D') \in S] = 0.25 \leq 0.679$
 Thus, we can validate that the 2 sets are 0.5-distinguishable
- Comparing the 1st group with 3rd group:
 $Pr[A(D) \in S] \leq e^\varepsilon Pr[A(D') \in S] = 0.25 \leq 0.679$
 Thus, we can validate that the 2 sets are 0.5-distinguishable
- Comparing the 1st group with 4th group:
 $Pr[A(D) \in S] \leq e^\varepsilon Pr[A(D') \in S] = 0.25 \leq 0.678$
 Thus, we can validate that the 2 sets are 0.5-distinguishable

1. Task (d)

Comparing all the 8 groups we observe

- Group = 1 , $\varepsilon = 0.5$
 - Sensitivity = 3.8240916E-5
 - b = 7.648183E-5
 - Error bound = 1.1698941364112996E-8
- Group = 2 , $\varepsilon = 0.5$
 - Sensitivity = 3.0732353E-5
 - b = 6.146471E-5
 - Error bound = 7.555820403275854E-9
- Group = 3 , $\varepsilon = 0.5$
 - Sensitivity = 3.1451487E-5
 - b = 6.2902975E-5
 - Error bound = 7.913568428338363E-9
- Group = 4 , $\varepsilon = 0.5$
 - Sensitivity = 3.0899482E-5
 - b = 6.1798964E-5
 - Error bound = 7.638224014793851E-9
- Group = 1 , $\varepsilon = 1$
 - Sensitivity = 3.8240916E-5
 - b = 3.8240916E-5
 - Error bound = 2.924735341028249E-9
- Group = 2 , $\varepsilon = 1$
 - Sensitivity = 3.0732353E-5
 - b = 3.0732353E-5
 - Error bound = 1.8889551008189636E-9
- Group = 3 , $\varepsilon = 1$
 - Sensitivity = 3.1451487E-5
 - b = 3.1451487E-5
 - Error bound = 1.9783921070845907E-9
- Group = 4 , $\varepsilon = 1$
 - Sensitivity = 3.0899482E-5
 - b = 3.0899482E-5
 - Error bound = 1.909556003698463E-9

- We can see that when $\varepsilon = 0.5$ the variance is small, the difference between the actual and noisy result (error bound) is also minimum.
- When $\varepsilon = 1$ the variance increases, therefore the difference between the actual and noisy result (error bound) also increases.
- Smaller value of ε gives higher privacy.
- Also, we can see from tasks (a), (b) & (c) that even after removing one tuple from the dataset, the result is still ε – *indistinguishable* when compared to the original dataset.
- Looking at the sensitivity of each dataset, removing a tuple can change the sensitive of the dataset, therefore we consider sensitivity when adding noise to the result.