1. Uniform $[-C/\varepsilon, C/\varepsilon]$

Let there be a value a

then Probability Density function is

$$z(x) = \begin{cases} \dfrac{\varepsilon}{2C} & \text{if } x \in [-C/\varepsilon, C/\varepsilon] \\ 0 & \text{otherwise} \end{cases}$$

$$\Pr[M(x) = a] = \Pr[f(x) + z = a]$$

$$= \Pr[z = a - f(x)]$$

$$= \begin{cases} \dfrac{\varepsilon}{2C} & \text{if } a - f(x) \in [-C/\varepsilon, C/\varepsilon] \\ 0 & \text{otherwise} \end{cases}$$

$$\Pr[M(x') = a] = \Pr[f(x') + z = a]$$

$$= \Pr[z = a - f(x')]$$

$$= \begin{cases} \dfrac{\varepsilon}{2C} & \text{if } a - f(x') \in [-C/\varepsilon, C/\varepsilon] \\ 0 & \text{otherwise} \end{cases}$$

So $\dfrac{\Pr[M(x) = a]}{\Pr[M(x') = a]} = \infty$ if

$a - f(x') \in [-c/\varepsilon, c/\varepsilon]$

but $a - f(x) \notin [-c/\varepsilon, \varepsilon/\varepsilon]$

$= 0$ if

$a - f(x) \in [-c/\varepsilon, c/\varepsilon]$

but $a - f(x') \notin [-c/\varepsilon, \varepsilon/\varepsilon]$

$= 1$ if

$a - f(x) \in [-c/\varepsilon, c/\varepsilon]$

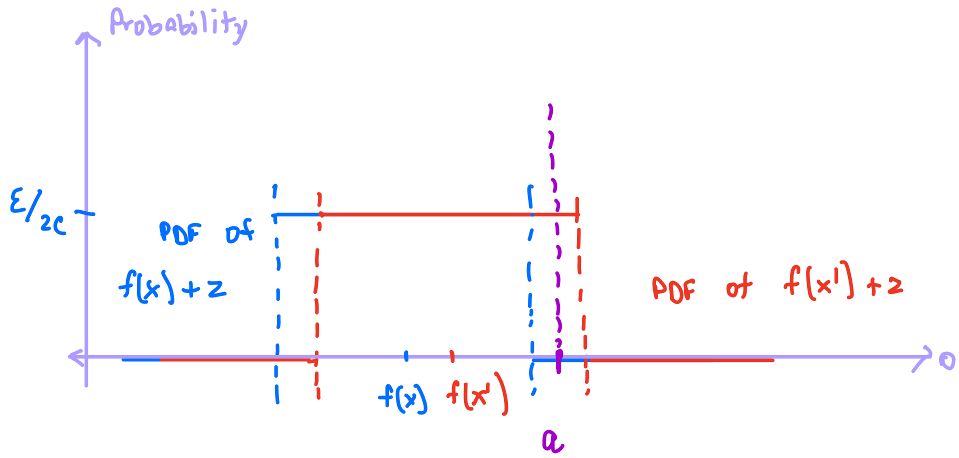
but $a - f(x') \in [-c/\varepsilon, \varepsilon/\varepsilon]$

$=$ undefined otherwise.

In any case, it is not e^{ε}.

Therefore, the given mechanism is not ε- diffentially private//.

eg:



$$Pr[M(x)=a] = 0$$

$$Pr[M(x')=a] = \varepsilon/2c$$

$$\frac{Pr[M(x)=a]}{Pr[M(x')=a]} = \infty$$

So, we cannot set a multicative bound in the change in the probability distribution of the output generated by single input change.

So, the outputs of x and x' can help the find some information regarding the user, a and hence privacy is not maintained.

2. Given R, u

$$u : X^n \times R \to \mathbb{R}$$

$M_E(x, u, R)$ selects and outputs $r \in R$ with probability proportional to $e^{\varepsilon \cdot u(x,r)/\Delta u}$

where $\Delta u = \max\limits_{r \in R} \max\limits_{x,x'} |u(x,r) - u(x',r)|$

To PROVE:

$$\Pr\left[u(M_E(x,u,R)) \leq OPT_u(x) - \frac{\Delta u}{\varepsilon} \cdot (\ln|R| + t)\right] \leq e^{-t}$$

where $OPT_u(x) = \max\limits_{r \in R} u(x,r)$

Let $c = OPT_u(x) - \frac{\Delta u}{\varepsilon} \cdot (\ln|R| + t)$

So $$\Pr[u(M_E(x,u,R)) \leq c] = \sum_{r : u(x,r) \leq c} \frac{e^{\varepsilon \cdot u(x,r)/\Delta u}}{\sum\limits_{s \in R} e^{\varepsilon \cdot u(x,s)/\Delta u}}$$

As these are all the $r \in R$ who have utility $\leq c$

$$\Pr[u(M_E(x,u,R)) \leq c] \leq \sum_{r : u(x,r) \leq c} \frac{e^{\varepsilon \cdot c/\Delta u}}{\sum\limits_{s \in R} e^{\varepsilon \cdot u(x,s)/\Delta u}} \quad -① $$

For the denominator

$$\sum_{s \in R} e^{\varepsilon \cdot u(x,s)/\Delta u} \geq e^{\varepsilon \cdot OPT_u(x)/\Delta u} \quad - \text{(2)}$$

includes the optimal answer

So (2) in (1)

$$\Pr[u(M_\varepsilon(x,u,R)) \leq c] \leq \sum_{r : u(x,r) \leq c} \frac{e^{\varepsilon \cdot c/\Delta u}}{\underbrace{e^{\varepsilon \cdot OPT_u(x)/\Delta u}}_{\text{independent of } r}} \quad - \text{(3)}$$

So

$$\Pr[u(M_\varepsilon(x,u,R)) \leq c] \leq \frac{e^{\varepsilon \cdot c/\Delta u}}{e^{\varepsilon \cdot OPT_u(x)/\Delta u}} \times \text{count of } r \in R \text{ such that } u(x,r) \leq c$$

$$\leq \frac{e^{\varepsilon \cdot c/\Delta u}}{e^{\varepsilon \cdot OPT_u(x)/\Delta u}} \times |R|$$

$$\leq |R| \, e^{\varepsilon/\Delta u \, (c - OPT_u(x))} \quad - \text{(4)}$$

Substitute $c = OPT_u(x) - \frac{\Delta u}{\varepsilon} \cdot (\ln|R| + t)$ in (4)

$$\frac{\varepsilon}{\Delta u}(c - OPT_u(x)) = \frac{-\varepsilon}{\Delta u} \times \frac{\Delta u}{\varepsilon} \cdot (\ln|R| + t)$$

$$= -(\ln|R| + t)$$

$$\Pr\left[u\left(M_E(x, u, R)\right) \leq OPT_u(x) - \frac{\Delta u}{\varepsilon} \cdot (\ln|R| + t)\right] \leq |R| \cdot e^{-(\ln|R| + t)}$$

$$\leq |R| \cdot \frac{1}{|R|} \times e^{-t}$$

$$\leq e^{-t} \;//$$

HENCE PROVED//

3. a. Sensitivity

So, we can change 1 value and this can drastically change median

eg: $x = \{1, 2, N\}$ and $x' = \{1, N-1, N\}$

So $|f(x) - f(x')| = (N-1) - 2 = N-3$

eg: $x = \{0\}$ and $x' = \{N\}$

So, $|f(x) - f(x')| = N$

As no number can be greater than N or less than 0, we can argue that

$|f(x) - f(x')| \leq N$

Therefore, $S_1(\text{median}) = N$

b. LAPLACIAN MECHANISM:

We need to apply Laplacian Mechanism with noise $b = N/\varepsilon$ to achieve ε-DP.

$$\Pr[\text{error} > Nt/\varepsilon] \leq e^{-t}$$

But as N can be large, we need to apply a large noise N/ε and hence, the output results can be erroneous.

WE WILL LOSE ACCURACY //

MECHANISM:

1. Compute $f(x) = \text{Median}(x)$

2. Release $f(x) + z$

   where $z \sim \text{Laplacian}(N/\varepsilon)$

c) $x \in [N]^n$ $\qquad x \in \{0, 1, 2 \dots N\}$

$u(x, r) = -\min |x \oplus y|$ such that median $(y) = r$

where $x \oplus y$ can be seen as number of elements that differ between x and y.

↳ This can also be seen as L_0 norm of $(x - y)$.

So if $x = \{0, 3, 5, 7, 100\}$

$r = 6$

We can choose y as

$\{0, 3, 6, 7, 100\}$

So $u(x, r) = -1$ as only one element is different between x and y.

SENSITIVITY :

$$\Delta u = \max_{r \in R} \max_{x, x'} |u(x, r) - u(x', r)|$$

where x, x' are two neighboring databases.

So $|x \oplus x'| = 1$ $\quad$ - ①

Let us assume $u(x, r)$ is s for any arbitrary r.

As per the utility function definition,

$$u(x, r) = -\min |x \oplus y| \text{ such that } \operatorname{median}(y) = r$$

i.e. there exists some y^*

such that

$$|x \oplus y^*| = -s \text{ and } \operatorname{median}(y^*) = r \quad -(2)$$

So $|x' \oplus y^*| \leq |x' \oplus x| + |x \oplus y^*|$

$$= -s + 1 \quad (\text{From } (1), (2))$$

We know $\operatorname{median}(y^*) = r$

So by the utility function definition,

$$u(x', r) \geq -(-s+1)$$

$$= s - 1$$

So, $|u(x, r) - u(x', r)| \leq |s - (s-1)| = 1$

So $\Delta u = \max\limits_{r \in R} \max\limits_{x, x'} |u(x, r) - u(x', r)|$

$$\boxed{\Delta u \leq 1}$$

This is true for all x, x' such that they are neighboring and for all $r \in R$ //

So, Sensitivity is independent of N, n.

So

$$\Pr\left[u(M_\varepsilon(x,u,R)) \leq OPT_u(x) - \frac{\Delta u}{\varepsilon}\cdot\left(\ln|R| + t\right)\right] \leq e^{-t}$$

i.e. $$\Pr\left[\text{Error} \geq \frac{1}{\varepsilon}\left(\ln|n| + t\right)\right] \leq e^{-t}$$

d. $x \in [N]^n$ Let $f(x) = 90^{th}$ percentile of x.

$$u(x, r) = -\min |x \oplus y| \text{ such that } f(y) = r$$

where $x \oplus y$ can be seen as number of elements that differ between x and y.

↳ This can also be seen as L_0 norm of $(x - y)$.

SENSITIVITY :

$$\Delta u = \max_{r \in R} \max_{x, x'} |u(x, r) - u(x', r)|$$

where x, x' are two neighboring databases.

So $|x \oplus x'| = 1$ — ①

Let us assume $u(x, r)$ is s for any arbitrary r.

As per the utility function definition,

$$u(x, r) = -\min |x \oplus y| \text{ such that } f(y) = r$$

i.e. there exists some y^* such that

$$|x \oplus y^*| = -s \text{ and } f(y^*) = r \quad - ②$$

So $|x' \oplus y^*| \leq |x' \oplus x| + |x \oplus y^*|$

$= -s + 1$ (From ①, ②)

We know $f(y^*) = r$

So by the utility function definition,

$$u(x', r) \geq -(-s+1)$$

$$= s - 1$$

So, $|u(x,r) - u(x',r)| \leq |s - (s-1)| = 1$

So $\Delta u = \max\limits_{r \in R} \max\limits_{x, x'} |u(x,r) - u(x',r)|$

$$\boxed{\Delta u \leq 1}$$

This is true for all x, x' such that they are neighboring and for all $r \in R$ //

So, Sensitivity is independent of N, n.