Homework Problem

Disprove: $\forall M$, if M is (ε, δ)-dp wrt $s \in R$ then M is (ε, δ)-dp wrt all $S \subseteq R$

Homework Problem

Disprove: $\forall M$, if M is (ε, δ)-dp wrt $s \in R$

then M is (ε, δ)-dp wrt all $S \subseteq R$

Let $R = \{a, b_1, b_2\}$ be the range of R.

Let x, x' be neighboring databases.

$$M(x) = \begin{cases} a & \text{with prob. } 1-2\delta \\ b_1 & \text{" " } \delta \\ b_2 & \text{" " } \delta \end{cases}$$

Let $\varepsilon = \delta$

Homework Problem

$\varepsilon = \delta$

$$M(x) = \begin{cases} a & \text{wp } 1-2\delta \\ b_1 & \text{wp } \delta \\ b_2 & \text{wp } \delta \end{cases}$$

$$M(x') = \begin{cases} a & \text{wp } 1 \end{cases}$$

Homework Problem

$\varepsilon = \delta \;,\; \delta > 0$

$$M(x) = \begin{cases} a & \text{wp } 1-2\delta \\ b_1 & \text{wp } \delta \\ b_2 & \text{wp } \delta \end{cases}$$

$$M(x') = \begin{cases} a & \text{wp } 1 \end{cases}$$

Claim 1 M is (ε, δ)-dp $\forall s \in R$

$$\Pr[M(x) = a] = 1 - 2\delta$$
$$\Pr[M(x') = a] = 1$$

$$\Pr[M(x) = a] = 1 - 2\delta \le e^{\varepsilon} \cdot 1 + \varepsilon$$

Homework Problem

$\varepsilon = \delta \;, \quad \delta > 0$

$$M(x) = \begin{cases} a & \text{wp } 1-2\delta \\ b_1 & \text{wp } \delta \\ b_2 & \text{wp } \delta \end{cases}$$

$$M(x') = \begin{cases} a & \text{wp } 1 \end{cases}$$

Claim 1 M is (ε, δ)-dp $\forall s \in R$

$$\Pr[M(x) = a] = 1 - 2\delta$$
$$\Pr[M(x') = a] = 1$$

$$\Pr[M(x') = a] = 1 \leq (1+\varepsilon)(1-2\delta) + \delta \leq (1+\delta)(1-2\delta) + \delta \leq 1$$

Homework Problem

$\varepsilon = \delta \,, \quad \delta > 0$

$$M(x) = \begin{cases} a & \text{wp } 1 - 2\delta \\ b_1 & \text{wp } \delta \\ b_2 & \text{wp } \delta \end{cases}$$

$$M(x') = \begin{cases} a & \text{wp } 1 \end{cases}$$

Claim 1 M is (ε, δ)-dp $\forall s \in R$

$$\Pr[M(x) = b_1] = \delta$$

$$\Pr[M(x') = b_1] = 0$$

$$\Pr[M(x) = b_1] = \delta \leq 0 + \delta = e^{\varepsilon} \Pr[M(x') = b_1] + \delta$$

$$\Pr[M(x') = b_1] = 0 \leq e^{\varepsilon} \cdot \delta + \delta$$

Homework Problem

$\varepsilon = \delta, \quad \delta > 0$

$$M(x) = \begin{cases} a & \text{wp } 1-2\delta \\ b_1 & \text{wp } \delta \\ b_2 & \text{wp } \delta \end{cases}$$

$$M(x') = \begin{cases} a & \text{wp } 1 \end{cases}$$

Claim 2 Let $S = \{b_1, b_2\}$

Then $\underbrace{\Pr[M(x) \in S]}_{2\delta} \not\leq \underbrace{e^{\varepsilon} \Pr(M(x') \in S]}_{0} + \underbrace{\delta}_{\delta}$

More DP techniques

1. Composition, Advanced Composition

2. Sparse Vector

3. Blum-Ligett-Roth [BLR]:
   release a sanitized database
   that is DP, and accurate
   for a large family of queries

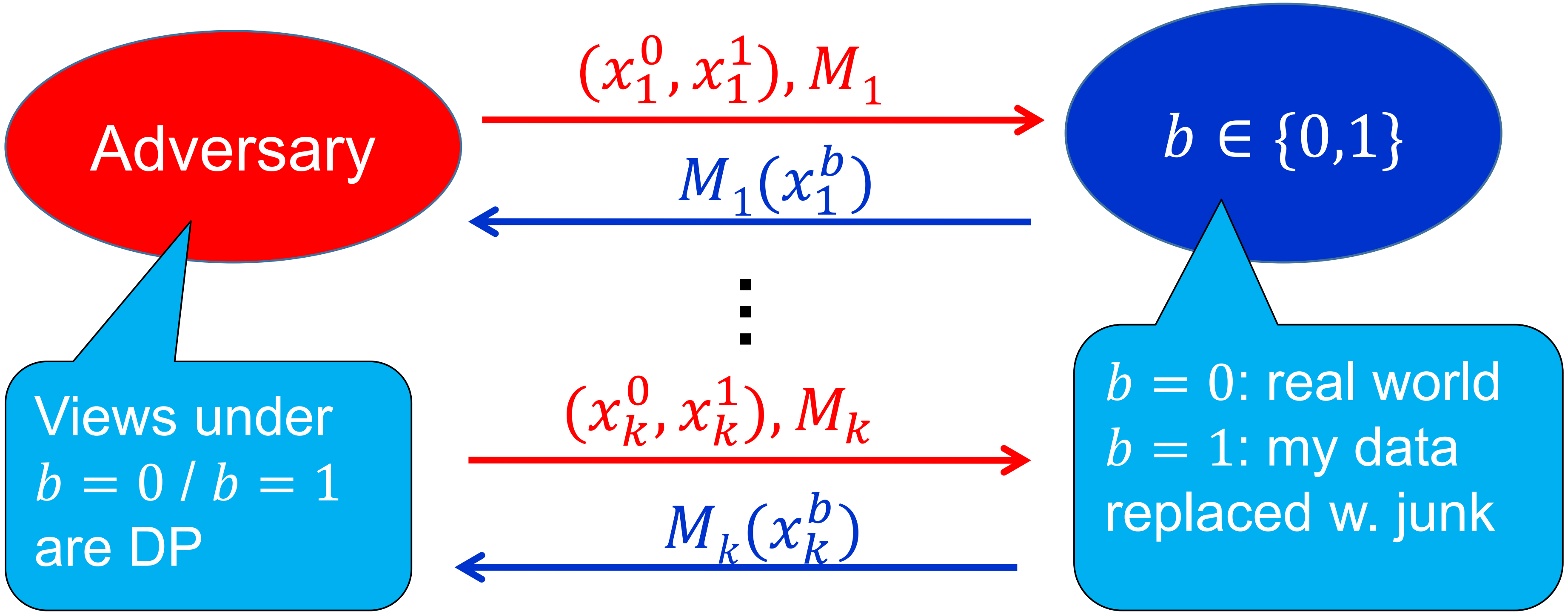
4. DP $\Rightarrow$ generalization

Basic composition

- Setting:
  - M_i be (ϵ_i, δ_i)-differentially private
  - M applies $M_1, \dots, M_t$ on its input (the inner $M_1, \dots, M_t$ use independent randomness).
- Basic composition theorem [DMNS06, DL09]:
  - M is $(\sum_i \epsilon_i, \sum_i \delta_i)$-differentially private
- Basic composition suggests that ϵ (and to a lesser account δ) can be treated as a 'privacy budget':
  - Split 'privacy budget' ϵ into smaller budget $\sum_i \epsilon_i$; allocate portion ϵ_i to mechanism M_i
    - Spend your budget carefully!
- More refined theorems (later):
  - Advanced composition [DRV10]
  - Optimal composition [KOV15, MV15]

Composition in differential privacy

- How do we define it?
  - Both choice of databases and algorithms is adaptive and adversarial [DRV10]



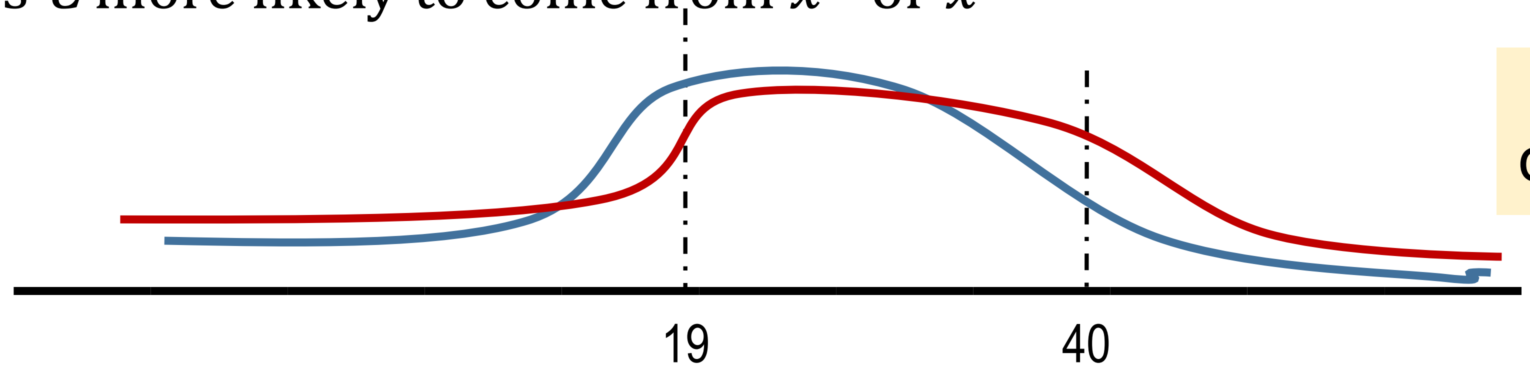
Thx: Guy Rothblum

What is privacy loss?

- Measured by the 'privacy loss' parameter ϵ
- Fix adjacent x^0, x^1, draw $C \leftarrow M(x_0)$
  - Is C more likely to come from x^0 or x^1

"19" more likely as output on x^0 than on x^1

"40" more likely as output on x^1 than on x^0



- Define $Loss(C) = \ln\left[\frac{\Pr[M(x^0)=C]}{\Pr[M(x^1)=C]}\right]$
  - $(\varepsilon, 0) - DP: w.p. 1\ over\ C,\ |Loss(C)| \leq \varepsilon$
  - $(\varepsilon, \delta) - DP^*: w.p. 1 - \delta\ over\ C,\ \ |Loss(C)| \leq \varepsilon$

Log of likelihood ratio

What is privacy loss?

- Fix adjacent x^0, x^1, draw $C \leftarrow M(x_0)$

$$Loss(C) = \ln\left[\frac{\Pr\left[M\left(x^0\right) = C\right]}{\Pr[M(x^1) = C]}\right]$$

- In multiple independent executions *loss* accumulates
  - Worst case: *Loss*= ε for every execution (as in analysis of basic composition)
  - This is pessimistic: *Loss* can be positive, negative → cancellations
  - Random variable, has a mean ([DDN03, DRV10]…)

Privacy Loss in Randomized Response

(general case follows similar argument)

$$RR_\varepsilon(x_i) = \begin{cases} x_i & \text{w.p. } \frac{e^\varepsilon}{e^\varepsilon+1} \\ \neg x_i & \text{w.p. } \frac{1}{e^\varepsilon+1} \end{cases}$$

Privacy loss

$$\ln\left[\frac{\Pr[y_i=0 \mid x_i=0]}{\Pr[y_i=0 \mid x_i=1]}\right] = \ln\left[e^\varepsilon\right] = \varepsilon$$

$$\ln\left[\frac{\Pr[y_i=0 \mid x_i=1]}{\Pr[y_i=0 \mid x_i=0]}\right] = \ln\left[e^{-\varepsilon}\right] = -\varepsilon$$

$c_i \approx$ privacy loss of step i

so $\quad -\varepsilon \le c_i \le \varepsilon$

Privacy Loss in Randomized Response

So $-\varepsilon \leq c_i \leq \varepsilon$

$$E[c_i] = \varepsilon \cdot \frac{e^\varepsilon}{e^\varepsilon+1} - \varepsilon\left[\frac{1}{e^\varepsilon+1}\right] \approx \frac{\varepsilon(1+\varepsilon-1)}{e^\varepsilon+1} \sim \varepsilon^2$$

so $E\left[\sum_{i=1}^{k} c_i\right] = \sum_{i=1}^{k} E[c_i] \sim k \cdot \varepsilon^2$

$\therefore$ Expected cumulative loss $E\left[\sum c_i\right] \sim k\varepsilon^2$

and $\left|\sum_{i=1}^{j+1} c_i - \sum_{i=1}^{j} c_i\right| \leq \varepsilon$

So this is a Martingale

Azuma's Inequality

Let $c_1, c_2, .. c_k$ be real valued r.v.'s satisfying this ε-Lipshitz property: $\forall j$

$$\left| \sum_{i=1}^{j+1} c_i - \sum_{i=1}^{j} c_i \right| \le \varepsilon$$

Then $\forall t \ge 0$

$$\Pr\left[\sum_{i=1}^{k} c_i \ge E\left[\sum_{i=1}^{k} c_i\right] + t \right] \le 2^{-t^2/2k\varepsilon^2}$$

Azuma's Inequality

Let $c_1, c_2, \ldots c_k$ be real valued r.v.'s satisfying this ε-Lypshitz property: $\forall j$

$$\left| \sum_{i=1}^{j+1} c_i - \sum_{i=1}^{j} c_i \right| \leq \varepsilon$$

Then $\forall t \geq 0$

$$\Pr\left[\sum_{i=1}^{k} c_i \geq E\left[\sum_{i=1}^{k} c_i\right] + t \right] \leq 2^{-t^2/2k\varepsilon^2}$$

We have $E\left[\sum_{i=1}^{k} c_i\right] \sim k\varepsilon^2$

choose $t \approx \sqrt{k \log \frac{1}{\delta}}\ \varepsilon$ gives

$$\Pr\left[\sum_{i=1}^{k} \geq \underbrace{k\varepsilon^2 + \sqrt{k\log\frac{1}{\delta}} \cdot \varepsilon}_{\varepsilon'} \right] \leq \delta$$

So we have (ε', δ)-dp

Advanced Composition [DRV10]

Composing k pure-DP algorithms (each ε_0-DP):

$$\varepsilon_g = O\left(\sqrt{k \cdot \ln\frac{1}{\delta_g}} \cdot \varepsilon_0 + k \cdot \varepsilon_0^2\right)$$ with all but δ_g probability.

For all δ_g simultaneously

Dominant if $k \ll \frac{1}{\epsilon_0^2}$

Dominant if $k \gg \frac{1}{\epsilon_0^2}$