Generalize: "stable histograms".

Laplace histograms: some entry off by $\Theta\left(\frac{\log k}{\varepsilon}\right)$

k could be huge: arbitrarily larger than n: $|\mathcal{U}|$; even infinite!

Want to get $\Theta\left(\frac{1}{\varepsilon}\ln\frac{1}{\delta}\right)$.

Think of $\delta \le \frac{1}{n}$, $n=|\mathcal{U}|$, so looking for $\frac{1}{\varepsilon}\ln n$ error.

Idea: would be done if only released nonzero counts, but would violate privacy

- only release counts above a threshold, round all others down to 0!

SH:

- For each $x \in X$:
  - if $D_x = 0$, set $a_x = 0$ ($D_* =$ count of $*$ in D)
  - if $D_x > 0$:
    - Set $a_x = D_x + \mathrm{Lap}\left(\frac{1}{\varepsilon}\right)$
    - If $a_x < \frac{1}{\varepsilon}\ln\frac{1}{\delta} + 1$, set $a_x = 0$

Thm: For $\delta \le \frac{1}{n}$, w.h.p. $|a_x - D_x| \le O\left(\frac{1}{\varepsilon}\ln\frac{1}{\delta}\right)$ $\forall x \in X$

Pf: Definitely true if $D_x = 0$

else, w.h.p. $D_x + Lap\left(\frac{1}{\varepsilon}\right) \le D_x + O\left(\frac{1}{\varepsilon} \ln n\right)$

$\le D_x + O\left(\frac{1}{\varepsilon} \ln \frac{1}{\delta}\right) \quad \forall x \in X$

$\Rightarrow |a_x - D_x| \le O\left(\frac{1}{\varepsilon} \ln \frac{1}{\delta}\right) + \frac{1}{\varepsilon} \ln \frac{1}{\delta} + 1$

$\le O\left(\frac{1}{\varepsilon} \ln \frac{1}{\delta}\right)$ w.h.p.

Thm: (ε, δ)-DP.

Pf: Let $D \sim D'$ where $D_x = D'_x + 1$ (added x to D)

$\Rightarrow$ only diff. bw $SH(D)$ and $SH(D')$ is in q_x (all other values identically distributed)

Case 1: $D'_x > 0$

$\Rightarrow$ add $Lap\left(\frac{1}{\varepsilon}\right)$-noise to both D'_x and D_x

$\Rightarrow \varepsilon$-DP (truncation is post-processing)

Case 2: $D'_x = 0 \Rightarrow a'_x = 0$

$\Rightarrow D_x = 1$

$\Rightarrow Pr[a_x \neq 0] = Pr\left[Lap\left(\frac{1}{\varepsilon}\right) > \frac{1}{\varepsilon} \ln \frac{1}{\delta}\right] \le \delta$

$\Rightarrow (0, \delta)$-DP

$\Rightarrow (\varepsilon, \delta)$-DP overall.

privately bounding LS:

- Instead of passing β arbitrarily in PTR, add noise to LS!

- Compute $\hat{\beta}(D)$ privately so that w/prob. $\geq 1-\delta$,

$\Delta_{LS}(D,f) \leq \hat{\beta}(D)$

$\Rightarrow$ output $f(D) + Lap\left(\frac{\hat{\beta}}{\varepsilon}\right)$, set (ε,δ)-DP.

Doesn't help w/ median, but what about other problems?

Ex: triangle counting!

Given graph $G=(V,E)$, $G'=(V,E')$, neighboring if

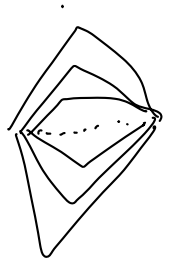
$|E \Delta E'| = 1$

(edge-privacy)

(often want node-privacy, but edge-privacy much easier).

Goal: Let $f(G)$ = # triangles in G.

$\Delta f = \max_{G,G': G \sim G'} |f(G) - f(G')| = n-2$:



So obvious ε-DP mechanism returns $f(G) + Lap\left(\frac{n-2}{\varepsilon}\right)$

What about local sensitivity?

$\Delta_{CS}(f,G) = \max_{u,v \in V} |N(u) \cap N(v)|$: # pairs in common

Need a private estimate which is β-local sensitivity of CS?

Add/remove edge $\{u,v\}$: affects $|N(u)|$ by 1, $|N(v)|$ by 1

$\Rightarrow \Delta(\Delta_{CS}(f,G)) \leq 1$

$M(G)$:

- Compute $\hat{\beta} = \Delta_{CS}(f,G) + \underbrace{Lap(\frac{1}{\varepsilon})}_{Z_\beta} + \frac{1}{\varepsilon}\ln\frac{1}{\delta}$
- Output $f(G) + \underbrace{Lap(\frac{\hat{\beta}}{\varepsilon})}_{Z_f} = \hat{f}(G)$

Thm: $E[|f(G) - \hat{f}(G)|] = O\left(\frac{\Delta_{CS}(f,G) + \frac{1}{\varepsilon}(1+\ln\frac{1}{\delta})}{\varepsilon}\right)$

pf: $E[|f(G) - \hat{f}(G)|] = E[|Z_f|]$

$= O(\frac{1}{\varepsilon} E[\hat{\beta}]) = O(\frac{1}{\varepsilon}[\Delta_{CS}(f,G) + E[Z_\beta] + \frac{1}{\varepsilon}\ln\frac{1}{\delta}])$

$= O(\frac{1}{\varepsilon}(\Delta_{CS}(f,G) + \frac{1}{\varepsilon}(1+\ln\frac{1}{\delta})))$

Thm: $(2\varepsilon, \delta)$-DP

pf: Computing $\hat{\beta}$ is ε-DP by Laplace mechanism

$$\Pr\left[\hat{\beta} \le \Delta_{CS}(f,G)\right] \cdot \Pr\left[Z_\beta \le \frac{1}{\varepsilon}\ln\frac{1}{\delta}\right] = e^{-\ln\frac{1}{\delta}} = \delta$$

If $\hat{\beta} \ge \Delta_{CS}(f,G)$, then second step ε-DP by Laplace mechanism

$\Rightarrow 2\varepsilon$-DP overall

$\Rightarrow 2\varepsilon$-DP w/ prob. $\ge 1-\delta \Rightarrow (2\varepsilon, \delta)$-DP

Smooth Sensitivity:

$$\Delta_{SS}(f,D) = \max_{D'} \left(\Delta_{CS}(f,D')\, e^{-\varepsilon d(D,D')} \right)$$

($d(D,D')$: distance in # changes)

So D' closer to D matters, further downweighted

- Δ_{SS} can be very hard to compute, but can do so for median, trimmed mean

Surprising fact: Laplace noise only gives (ε,δ)-DP!

Facts:

- Adding Laplace or Gaussian noise scaled by $\Delta_{SS}(f,D)$ only (ε,δ)-DP

- To get more flexible variant of Cauchy distribution
  - polynomial tails, no finite expectation!