

# Differentially Private Bayesian Inference

March 12, 2020

## 1 Preliminary

- $\boldsymbol{\theta}$  : The parameter vector of multinomial distribution,  $\boldsymbol{\theta} \in [0, 1]^k$ .
- $\mathbf{x}$  : Observed dataset.  $\mathbf{x} \in \mathcal{X}^n, |\mathcal{X}| = k$
- $\text{Dir}(\boldsymbol{\alpha})$  : Dirichlet distribution. The prior or posterior distribution over  $\boldsymbol{\theta}$ .
- $\text{DirP}(\boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{x})$  : Posterior distribution over  $\boldsymbol{\theta}$  from Bayesian inference given prior distribution  $\text{Dir}(\boldsymbol{\alpha})$  and observed data set  $\mathbf{x}$ .
- $\mathcal{H}(\cdot, \cdot)$  : Hellinger Distance between two distributions.  $\mathcal{H}(\text{Dir}(\boldsymbol{\alpha}_1), \text{Dir}(\boldsymbol{\alpha}_2)) = \sqrt{1 - \frac{B(\frac{\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2}{2})}{\sqrt{B(\boldsymbol{\alpha}_1)B(\boldsymbol{\alpha}_2)}}}$
- $u(\mathbf{x}, r)$  : Scoring function given  $\boldsymbol{\alpha}$  and  $\boldsymbol{\theta}$  for candidate  $r$ .  $u(\mathbf{x}, r) = -\mathcal{H}(\text{DirP}(\boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{x}), r)$
- $GS$  : Global sensitivity of Hellinger distance.  $GS = \sqrt{1 - \pi/4}$
- $LS(\mathbf{x})$  : Local sensitivity of Hellinger distance for  $\mathbf{x}$ .  

$$LS(\mathbf{x}) = \max_{\mathbf{x}' \in \mathcal{X}^n, \text{adj}(\mathbf{x}, \mathbf{x}'), r \in \mathcal{R}_{\boldsymbol{\alpha}}} |\mathcal{H}(\text{DirP}(\mathbf{x}, \boldsymbol{\alpha}), r) - \mathcal{H}(\text{DirP}(\mathbf{x}', \boldsymbol{\alpha}), r)|$$
- $S(\mathbf{x})$  :  $\gamma$ -smooth sensitivity.  $S(\mathbf{x}) = \max_{\mathbf{x}' \in \mathcal{X}^n} \left\{ \frac{1}{\frac{1}{LS(\mathbf{x}')} + \gamma \cdot \text{Himming}(\mathbf{x}, \mathbf{x}')} \right\}$

## 2 Private Mechanisms

---

**Algorithm 1** LSDim - Calibrating noise w.r.t.  $\ell_1$  norm

---

```

 $\mathbf{x} \in \mathcal{X}^n, \text{Dir}(\boldsymbol{\alpha})$ 
let  $\boldsymbol{\alpha}' = \text{DirP}(\boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{x})$ 
Initialize a vector  $\tilde{\boldsymbol{\alpha}} = (0, \dots, 0) \in \mathbb{N}^{|\mathcal{X}|}$ 
For  $i = 1 \dots |\mathcal{X}| - 1$ :
    let  $\eta \sim \text{Lap}(0, \frac{|\mathcal{X}|}{\epsilon})$ 
     $\tilde{\alpha}_i = \alpha_i + \lfloor (\alpha'_i - \alpha_i) + \eta \rfloor_0^n$ 
 $\tilde{\alpha}_{|\mathcal{X}|} = \alpha_{|\mathcal{X}|} + \lfloor n - \sum_{i=1}^{|\mathcal{X}|-1} \lfloor (\alpha'_i - \alpha_i) + \eta_i \rfloor_0^n \rfloor_0^n$ 
return  $\tilde{\boldsymbol{\alpha}}$ 

```

---

---

**Algorithm 2** LSHist - Calibrating noise w.r.t. histogram sensitivity

---

**input**  $\mathbf{x} \in \mathcal{X}^n$ ,  $\text{Dir}(\boldsymbol{\alpha})$   
  **let**  $\boldsymbol{\alpha}' = \text{DirP}(\boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{x})$   
  **let**  $k = \begin{cases} 1 & \text{if } |\mathcal{X}| = 2 \\ 2 & \text{otherwise} \end{cases}$   
  **Initialize** a vector  $\tilde{\boldsymbol{\alpha}} = (0, \dots, 0) \in \mathbb{N}^{|\mathcal{X}|}$   
  **For**  $i = 1 \dots |\mathcal{X}| - 1$ :  
    **let**  $\eta \sim \text{Lap}(0, \frac{k}{\epsilon})$   
     $\tilde{\alpha}_i = \alpha_i + \lfloor (\alpha'_i - \alpha_i) + \eta \rfloor_0^n$   
   $\tilde{\alpha}_{|\mathcal{X}|} = \alpha_{|\mathcal{X}|} + \lfloor n - \sum_{i=1}^{|\mathcal{X}|-1} \lfloor (\alpha'_i - \alpha_i) + \eta_i \rfloor_0^n \rfloor_0^n$   
**return**  $\tilde{\boldsymbol{\alpha}}$

---

---

**Algorithm 3** EHD - Instantiation of the exponential mechanism

---

observed data set  $\mathbf{x} \in \mathcal{X}^n$ , prior:  $\text{Dir}(\boldsymbol{\alpha})$ ,  $\epsilon$   
  **let**  $\text{Dir}(\boldsymbol{\alpha}') = \text{DirP}(\mathbf{x}, \boldsymbol{\alpha})$ .  
  **let**  $GS$  be the global sensitivity for  $\mathbf{x}$ .  
  **set**  $z = r$  with probability  $\frac{\exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \boldsymbol{\alpha}), r)}{2 \cdot GS})}{\sum_{r' \in \mathcal{R}_{\boldsymbol{\alpha}}} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \boldsymbol{\alpha}), r')}{2 \cdot GS})}$   
**return**  $z$

---

---

**Algorithm 4** EHDL - Instantiation of the exponential mechanism with local sensitivity

---

**input** observed data set  $\mathbf{x} \in \mathcal{X}^n$ , prior:  $\text{Dir}(\boldsymbol{\alpha})$ ,  $\epsilon$   
  **let**  $\text{Dir}(\boldsymbol{\alpha}') = \text{DirP}(\mathbf{x}, \boldsymbol{\alpha})$ .  
  **let**  $LS(\mathbf{x})$  be the local sensitivity for  $\mathbf{x}$ .  
  **set**  $z = r$  with probability  $\frac{\exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \boldsymbol{\alpha}), r)}{2 \cdot LS(\mathbf{x})})}{\sum_{r' \in \mathcal{R}_{\boldsymbol{\alpha}}} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \boldsymbol{\alpha}), r')}{2 \cdot LS(\mathbf{x})})}$   
**return**  $z$

---

---

**Algorithm 5** EHDS - Instantiation of the exponential mechanism with  $\gamma$ -smooth sensitivity

---

observed data set  $\mathbf{x} \in \mathcal{X}^n$ , prior:  $\text{Dir}(\boldsymbol{\alpha})$ ,  $\epsilon$   
  **let**  $\text{Dir}(\boldsymbol{\alpha}') = \text{DirP}(\mathbf{x}, \boldsymbol{\alpha})$ .  
  **let**  $S(\mathbf{x})$  be the smooth sensitivity for  $\mathbf{x}$ .  
  **set**  $z = r$  with probability  $\frac{\exp(\frac{\epsilon \cdot u(\mathbf{x}, r)}{4 \cdot S(\mathbf{x})})}{\sum_{r' \in \mathcal{R}_{\boldsymbol{\alpha}}} \exp(\frac{\epsilon \cdot u(\mathbf{x}, r')}{4 \cdot S(\mathbf{x})})}$   
**return**  $z$

---

### 3 Privacy Analysis

**Theorem 3.1.** *The LSDim, LSHist, EHD and EHDS are  $\epsilon$ -differentially private.*

The proofs are in the Arxiv version.

## 4 Accuracy Analysis

**Theorem 4.1.** *To prove the optimality of Laplace mechanism, we are showing*

$$\frac{ELap(\mathbf{x})}{(\epsilon \times LS(\mathbf{x}))}$$

*is  $O(1)$ , considering  $n = |\mathbf{x}| \geq 2$  being the parameter.*

*Where  $LS(\cdot)$  is the local sensitivity, and where  $ELap(\cdot)$  is the measure of the error of the Laplace mechanism, defined in this way:*

$$ELap(\mathbf{x}) = \arg \left( \min_t \{Pr[H(\text{DirP}(\mathbf{x}), \text{LSHist}(\mathbf{x})) < t] \geq 1 - \gamma\} \right).$$

[[Jiawen:

**Theorem 4.2.** *For  $\gamma = e^{O(\epsilon)}$ ,  $\frac{ELap(\mathbf{x})}{(\epsilon \times LS(\mathbf{x}))}$  is  $O(\epsilon)$*

]]

*Proof.* Let  $t = LS(\mathbf{x})$ , we have following by p.d.f. of Laplace distribution:

$$Pr[H(\text{DirP}(\mathbf{x}), \text{LSHist}(\mathbf{x})) < t] \geq 1 - \frac{1}{2}(e^{-\epsilon} + e^{-2\epsilon}) > 1 - e^{-\epsilon}$$

Then we can get when  $\gamma = e^{-\epsilon}$ ,

$$\frac{ELap(\mathbf{x})}{(\epsilon \times LS(\mathbf{x}))} = \frac{1}{\epsilon}$$

□

**Theorem 4.3.** *In order to prove the optimality of Laplace mechanism, instead of prove  $\frac{ELap(\mathbf{x})}{(\epsilon \times LS(\mathbf{x}))}$  is  $O(1)$ , we prove a constant upper bound on following equations:*

$$\begin{aligned} & \frac{\arg \min_t \left\{ Pr[H(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] \geq 1 - \gamma \right\}}{LS(\mathbf{x})} \\ & \leq \frac{\max_{|k| \leq \frac{\lg(\frac{1}{\gamma})}{\epsilon}} H(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha + k, n - \lfloor \alpha + k \rfloor))}{LS(\mathbf{x})} \\ & \leq O\left(\frac{\lg \frac{1}{\gamma}}{\epsilon}\right) \end{aligned}$$

[[Jiawen:

**Theorem 4.4.** *For  $\gamma = e^{O(k)\epsilon}$ , it is proved that  $O(\frac{\lg \frac{1}{\gamma}}{\epsilon})$  is bounded by  $O(k)$ .*

]]

*Proof.* By Laplace distribution, we have:

$$\begin{aligned} \Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] &= \Pr[\{|\text{Lap}(0, \frac{1}{\epsilon})| < O(k) | \mathbf{H}(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha + k, n - \lfloor \alpha + k \rfloor)) < t\}] \\ &\leq 1 - e^{-O(k)\epsilon} \end{aligned}$$

Then we have:

$$\gamma = e^{-O(k)\epsilon}$$

So we can get:

$$O\left(\frac{\lg \frac{1}{\gamma}}{\epsilon}\right) = O\left(\frac{\lg \frac{1}{e^{-O(k)\epsilon}}}{\epsilon}\right) = O(k)$$

□

[[Jiawen:

**Corollary 4.4.1.** For  $-1 \leq k < 2$ , it is proved that  $O(\frac{\lg \frac{1}{\gamma}}{\epsilon})$  is bounded by  $O(1)$ .

]]

*Proof.* Given  $-1 \leq k < 2$ , we have:

$$\mathbf{H}(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha + k, n - \lfloor \alpha + k \rfloor)) \leq LS(\mathbf{x}) \quad (1)$$

For any  $\epsilon$ ,  $k \sim \text{Lap}(0, \frac{1}{\epsilon})$  from Laplace mechanism, we have:

$$\Pr[|k| \leq \frac{b}{\epsilon}] = 1 - \exp(-b)$$

Then we can get:

$$\Pr[-1 \leq k < 2] = 1 - \frac{\exp(-\epsilon) + \exp(-2\epsilon)}{2} \quad (2)$$

By Equation (1) and (2), we can get:

$$\Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha + k, n - \lfloor \alpha + k \rfloor)) \leq LS(\mathbf{x})] \geq 1 - \frac{\exp(-\epsilon) + \exp(-2\epsilon)}{2}$$

i.e.,

$$\begin{aligned} &\frac{\arg \min_t \left\{ \Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] \geq 1 - \frac{\exp(-\epsilon) + \exp(-2\epsilon)}{2} \right\}}{LS(\mathbf{x})} \\ &\leq O\left(\frac{\lg(\frac{2}{\exp(-\epsilon) + \exp(-2\epsilon)})}{\epsilon}\right) \\ &< O\left(\frac{\lg(\frac{2}{2\exp(-2\epsilon)})}{\epsilon}\right) = 2 \end{aligned}$$

□

[[Jiawen:

**Theorem 4.5.** Let  $k = \lfloor k' \rfloor$  be the largest integer that satisfying  $H(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha + k', n - \lfloor \alpha + k' \rfloor)) < t$ , we have:

$$\Pr[H(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] = 1 - \frac{1}{2}(e^{-k\epsilon} + e^{-(k+1)\epsilon}) \geq 1 - e^{-k\epsilon}$$

$$\frac{2ke^{-\epsilon} + 1}{n} < \Pr[H(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t] < \frac{2k+1}{ne^{-\epsilon}}.$$

$$\frac{2k \exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})}) + 1}{n} < \Pr[H(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t] < \frac{2k+1}{n \exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})})}.$$

]]

*Proof.* Given  $k = \lfloor k' \rfloor$  be the largest integer that satisfying  $H(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha + k', n - \lfloor \alpha + k' \rfloor)) < t$ , by the post-processing of Laplace distribution and p.d.f. of Laplace distribution, we have:

$$\Pr[H(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] = \Pr[-k < \text{LSHist}(\mathbf{x}) \leq k+1] = 1 - \frac{1}{2}(e^{-k\epsilon} + e^{-(k+1)\epsilon}) \geq 1 - e^{-k\epsilon}.$$

By definition of EHD and  $GS = \sqrt{1 - \pi/4}$ , we have:

$$\begin{aligned} \Pr[H(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t] &= \sum_{c \geq -t} \frac{\exp(\frac{\epsilon \cdot c}{2 \cdot GS})}{\sum_{r' \in \mathcal{R}_\alpha} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \alpha), r')}{2 \cdot GS})} \\ &\leq \frac{2k \exp(\frac{-0\epsilon}{2 \cdot GS}) + 1}{n \exp(\frac{-\epsilon}{2 \cdot GS})} \\ &= \frac{2k+1}{n \exp(\frac{-\epsilon}{2 \cdot GS})} \\ &< \frac{2k+1}{n \exp(\frac{-\epsilon}{2 \sqrt{1-\pi/4}})} \\ &< \frac{2k+1}{n \exp(-\epsilon)} \end{aligned}$$

$$\begin{aligned} \Pr[H(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t] &= \sum_{c \geq -t} \frac{\exp(\frac{\epsilon \cdot c}{2 \cdot GS})}{\sum_{r' \in \mathcal{R}_\alpha} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \alpha), r')}{2 \cdot GS})} \\ &\geq \frac{2k \exp(\frac{-0\epsilon}{2 \cdot GS}) + 1}{n \exp(\frac{-\epsilon}{2 \cdot GS})} \\ &> \frac{2k \exp(\frac{-\epsilon}{2 \sqrt{1-\pi/4}}) + 1}{n} \\ &> \frac{2ke^{-\epsilon} + 1}{n} \end{aligned}$$

By definition of EHD, we have:

$$\begin{aligned} \Pr[H(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t] &= \sum_{c \geq -t} \frac{\exp(\frac{\epsilon \cdot c}{2 \cdot LS(\mathbf{x})})}{\sum_{r' \in \mathcal{R}_\alpha} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \alpha), r')}{2 \cdot LS(\mathbf{x})})} \\ &\leq \frac{2k \exp(\frac{-0\epsilon}{2 \cdot LS(\mathbf{x})}) + 1}{n \exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})})} \\ &= \frac{2k+1}{n \exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})})} \end{aligned}$$

$$\begin{aligned} \Pr[H(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t] &= \sum_{c \geq -t} \frac{\exp(\frac{\epsilon \cdot c}{2 \cdot LS(\mathbf{x})})}{\sum_{r' \in \mathcal{R}_\alpha} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \alpha), r')}{2 \cdot LS(\mathbf{x})})} \\ &\geq \frac{2k \exp(\frac{-0\epsilon}{2 \cdot LS(\mathbf{x})}) + 1}{n \exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})})} \\ &> \frac{2k \exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})}) + 1}{n} \end{aligned}$$

Since  $\lim_{n \rightarrow \infty} LS(\mathbf{x}) \rightarrow 0$ , we have  $\lim_{n \rightarrow \infty} \frac{-1}{LS(\mathbf{x})} \rightarrow -\infty$ . So  $\exp(\frac{-\epsilon}{2 \cdot LS(\mathbf{x})})$  can only be bounded by 0. We cannot found a tighter lower bound.  $\square$

[[Jiawen:

**Corollary 4.5.1.** *For a reasonable small  $t$ , we have when data size  $n = |\mathbf{x}| > O(\frac{(2k+1)e^\epsilon}{1-e^{-\epsilon}})$ , the accuracy of LSHist is higher than EHD.*

]]

*Proof.* Based on Theorem 2.5, let:

$$\frac{2k+1}{n \exp(-\epsilon)} \leq 1 - e^{-k\epsilon},$$

we can have:

$$\Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] > \Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t].$$

By simplification, we have  $n > \frac{2k+1}{e^{-\epsilon}(1-e^{-k\epsilon})} \sim O(\frac{(2k+1)e^\epsilon}{1-e^{-\epsilon}})$ .  $\square$

[[Jiawen:

**Corollary 4.5.2.** *For a reasonable small  $t$ , we have when data size  $n = |\mathbf{x}| < O(\frac{(2ke^{-\epsilon}+1)}{1-\frac{1}{2}(e^{-k\epsilon}+e^{-(k+1)\epsilon})})$ , the accuracy of EHD is better than LSHist.*

]]

*Proof.* Applying the Theorem 2.5, let:

$$\frac{2ke^{-\epsilon}+1}{n} > 1 - \frac{1}{2}(e^{-k\epsilon} + e^{-(k+1)\epsilon}),$$

we can have:

$$\Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{LSHist}(\mathbf{x})) < t] < \Pr[\mathbf{H}(\text{Beta}(\alpha, \beta), \text{EHD}(\mathbf{x})) < t].$$

By simplification, we have  $n < \frac{(2ke^{-\epsilon}+1)}{1-\frac{1}{2}(e^{-k\epsilon}+e^{-(k+1)\epsilon})}$ .  $\square$

[[Jiawen:

**Corollary 4.5.3.** *Let  $R_g$  be the good output set where  $\forall r \in R, \mathbf{H}(\text{DirP}(\mathbf{x}), r) \leq LS(\mathbf{x})$ , we have:*

$$Pr[\text{LSHist}(\mathbf{x}, \epsilon) \in R_g] > Pr[\text{EHD}(\mathbf{x}, \epsilon) \in R_g]$$

*for data size  $n = |\mathbf{x}| > O(\frac{e^\epsilon}{1-e^{-\epsilon}})$*

]]

*Proof.* simply apply the Theorem 2.5 and corollary 2.5.1, we can get this conclusion.

Let  $R_g$  be the good output set where  $\forall r \in R, \mathbf{H}(\text{DirP}(\mathbf{x}), r) \leq LS(\mathbf{x})$ , we have:

$$Pr[\text{LSHist}(\mathbf{x}) \in R_g] \geq 1 - \frac{1}{2}(e^{-\epsilon} + e^{-2\epsilon}) > 1 - e^{-\epsilon}$$

By definition of EHD and  $GS = \sqrt{1 - \pi/4}$ , we have:

$$\begin{aligned}
Pr[\text{EHD}(\mathbf{x}) \in R_g] &= \sum_{c \geq -LS(\mathbf{x})} \frac{\exp(\frac{\epsilon \cdot c}{2 \cdot GS})}{\sum_{r' \in \mathcal{R}_\alpha} \exp(\frac{-\epsilon \cdot \mathcal{H}(\text{DirP}(\mathbf{x}, \alpha), r')}{2 \cdot GS})} \\
&\leq \frac{2 \exp(-\frac{\epsilon LS(\mathbf{x})}{2 \cdot GS}) + 1}{n \exp(\frac{-\epsilon}{2 \cdot GS})} \\
&\leq \frac{3}{n \exp(\frac{-\epsilon}{2\sqrt{1-\pi/4}})} \\
&\leq \frac{3}{n \exp(-\epsilon)}
\end{aligned}$$

Let  $c = 2\sqrt{1 - \pi/4}$ , we have when  $n > \frac{3}{e^{-\epsilon/c}(1-e^{-\epsilon})} \sim O(\frac{e^\epsilon}{1-e^{-\epsilon}})$  LSHist performs better than EHD.  $\square$