16 The Laplace Mechanism

This section is based heavily on the course Privacy in Statistics and Machine Learning by Adam Smith and Jonathan Ullman, primarily Lectures 4 and 5 (Spring 2023) [Smi23a, Smi23b].

Randomized response adds noise to each individual's answer. If we trust the curator with the original data, a more accurate approach is often to compute the statistic first and add noise to the result. How much noise do we need? The key is how much the statistic can change when we change one record.

16.1 The Laplace Mechanism in One Dimension

For a fixed function $f : \mathcal{X}^n \to \mathbb{R}$, define its *global sensitivity* by

$$\Delta_1(f) := \sup_{S,S' \text{ neighboring}} |f(S) - f(S')|.$$

For example, a count has sensitivity at most 1, and the empirical mean of values in $[0,1]$ has sensitivity at most $1/n$. We use a bound that holds over all neighboring datasets, rather than only the dataset we happen to observe.

The mean-zero Laplace distribution with scale $b > 0$, denoted $\mathrm{Lap}(b)$, has density

$$p_b(w) = \frac{1}{2b} \exp\left(-\frac{|w|}{b}\right).$$

If $W \sim \mathrm{Lap}(b)$, then $\mathbb{E}[W] = 0$, $\mathbb{E}[|W|] = b$, $\mathrm{Var}(W) = 2b^2$, and $\mathbb{P}[|W| > t] = e^{-t/b}$ for $t \geq 0$.

Algorithm 8: The one-dimensional Laplace mechanism

Input: A dataset $S \in \mathcal{X}^n$, a fixed function $f : \mathcal{X}^n \to \mathbb{R}$, a fixed bound $\Delta \geq \Delta_1(f)$ with $\Delta > 0$, and a parameter $\varepsilon > 0$.
Output: $\mathcal{A}(S) = f(S) + W$, where $W \sim \mathrm{Lap}(\Delta/\varepsilon)$.

Theorem 16.1. *Algorithm 8 is ε-differentially private. Its expected absolute error is Δ/ε and its root mean squared error is $\sqrt{2}\Delta/\varepsilon$. For every $\beta \in (0,1)$, with probability at least $1-\beta$,*

$$|\mathcal{A}(S) - f(S)| \leq \frac{\Delta}{\varepsilon} \log \frac{1}{\beta}.$$

Proof. Fix neighboring S, S' and let $p_S, p_{S'}$ be the densities of their outputs. For every $y \in \mathbb{R}$, the triangle inequality gives

$$\frac{p_S(y)}{p_{S'}(y)} = \exp\left(\frac{\varepsilon}{\Delta}\left(|y - f(S')| - |y - f(S)|\right)\right) \leq \exp\left(\frac{\varepsilon}{\Delta}|f(S) - f(S')|\right) \leq e^{\varepsilon}.$$

Integrating over any event R proves differential privacy. The error guarantees follow from the moments and tail probability of W. □

Swapping S, S' gives the lower bound as well: $e^{-\varepsilon} \le p_S(y)/p_{S'}(y) \le e^{\varepsilon}$ for every output y. Figure 12 illustrates this pointwise closeness under a shift.

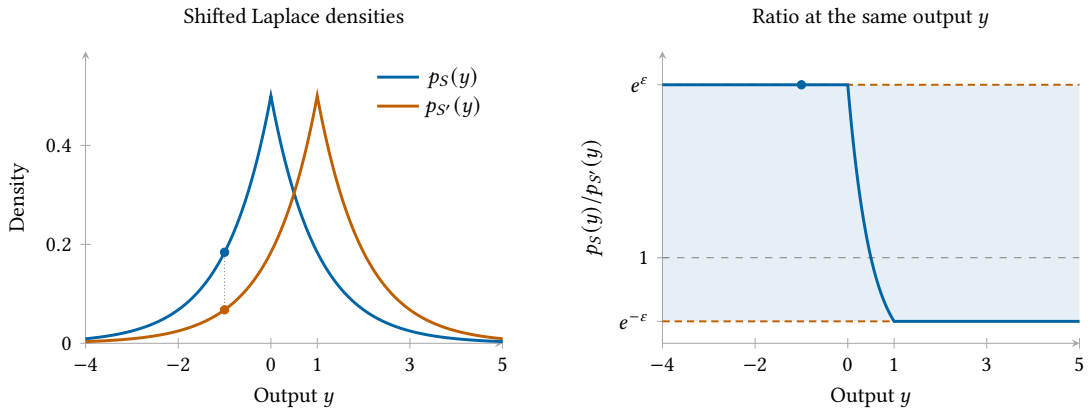


Figure 12: Two Laplace output distributions with $f(S) = 0$, $f(S') = 1$, and scale $b = \Delta/\varepsilon = 1$, where $\Delta = \varepsilon = 1$. Left: the dots compare the two densities at the same output $y = -1$. Right: their ratio stays between $e^{-\varepsilon}$ and e^{ε} everywhere, attaining the bounds in the tails.

For the empirical mean $\mu_S = \frac{1}{n}\sum_{i=1}^{n} \varphi(X_i)$ with $\varphi\colon \mathcal{X} \to [0,1]$, we can take $\Delta = 1/n$. The root mean squared error is then $\sqrt{2}/\varepsilon n$, compared with $\Theta(1/\varepsilon\sqrt{n})$ for randomized response when $0 < \varepsilon \le 1$. To get error at most α with probability at least $1 - \beta$, it suffices to have $n \ge \log(1/\beta)/\varepsilon\alpha$. These are errors relative to the empirical mean; estimating a population mean would also incur sampling error. If $\Delta_1(f) = 0$, the function is constant and we can release it without noise.

Figure 13 compares the sampling distributions when $X_i \sim \mathrm{Ber}(0.4)$, $n = 100$, and $\varepsilon = 0.5$. The Laplace estimate has almost the same spread as the empirical mean, while randomized response is much less concentrated. Here the distributions include both sampling noise and the mechanisms' random choices.

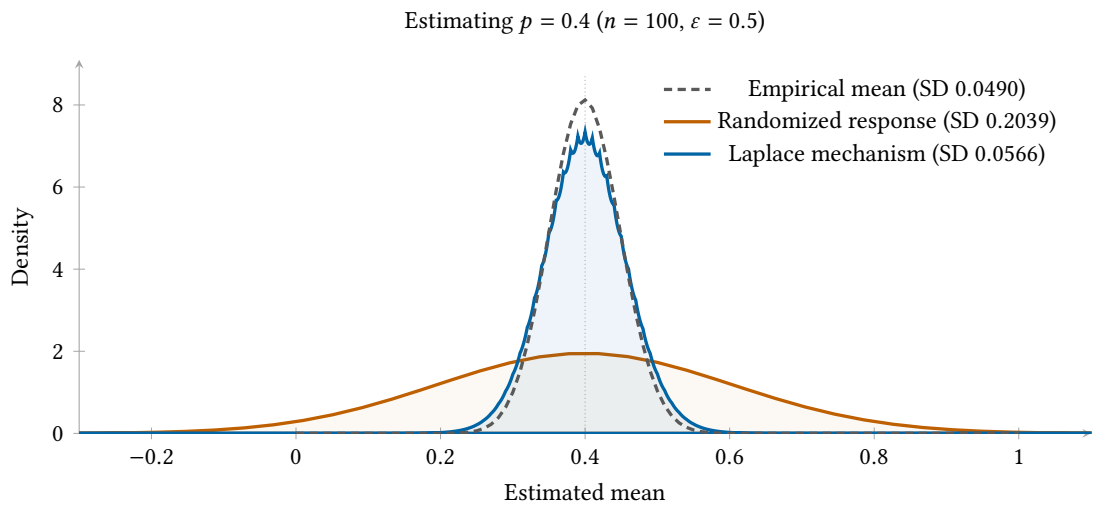


Figure 13: Sampling distributions for $n = 100$ independent Ber(0.4) records and $\varepsilon = 0.5$. Discrete curves join exact probability masses divided by grid spacing; the Laplace curve is the exact mixture density. Estimates are unclipped, and the standard deviations include data and mechanism randomness.

16.2 The Laplace Mechanism in d Dimensions

For a vector-valued function $f\colon \mathcal{X}^n \to \mathbb{R}^d$, we measure sensitivity in the ℓ_1 norm:

$$\Delta_1(f) := \sup_{S,S' \text{ neighboring}} \|f(S) - f(S')\|_1, \qquad \|v\|_1 = \sum_{j=1}^{d} |v_j|.$$

The mechanism adds independent Laplace noise to each coordinate, calibrated to the sensitivity of the *whole vector*.

Algorithm 9: The d-dimensional Laplace mechanism

Input: A dataset $S \in \mathcal{X}^n$, a fixed function $f\colon \mathcal{X}^n \to \mathbb{R}^d$, a fixed bound $\Delta \geq \Delta_1(f)$ with $\Delta > 0$, and a parameter $\varepsilon > 0$.
Output: $\mathcal{A}(S) = f(S) + (W_1, \ldots, W_d)$, where the $W_j \sim \mathrm{Lap}(\Delta/\varepsilon)$ are independent.

Theorem 16.2. *Algorithm 9 is ε-differentially private. Its error satisfies*

$$\mathbb{E}[\|\mathcal{A}(S) - f(S)\|_1] = \frac{d\Delta}{\varepsilon}, \qquad \sqrt{\mathbb{E}[\|\mathcal{A}(S) - f(S)\|_2^2]} = \frac{\sqrt{2d}\Delta}{\varepsilon}.$$

Moreover, for every $\beta \in (0,1)$, with probability at least $1 - \beta$,

$$\|\mathcal{A}(S) - f(S)\|_\infty \leq \frac{\Delta}{\varepsilon} \log \frac{d}{\beta}.$$

Proof. The joint output density is

$$p_S(y) = \left(\frac{\varepsilon}{2\Delta}\right)^d \exp\left(-\frac{\varepsilon}{\Delta}\|y - f(S)\|_1\right).$$

For neighboring S, S', the same triangle inequality as in one dimension gives

$$\frac{p_S(y)}{p_{S'}(y)} \le \exp\left(\frac{\varepsilon}{\Delta}\|f(S) - f(S')\|_1\right) \le e^{\varepsilon}.$$

Integrating proves privacy. The first two error bounds follow by summing the absolute and second moments of the coordinates. For the last bound, apply the Laplace tail bound with failure probability β/d to each coordinate and take a union bound. □

Example: histograms. Fix a partition $B_1, \dots, B_d$ of $\mathcal{X}$ and let $f(S)$ be the vector of counts,

$$f_j(S) = \sum_{i=1}^{n} \mathbb{1}\{X_i \in B_j\}.$$

Replacing one record can decrease one count by 1 and increase another by 1, so $\Delta_1(f) \le 2$, regardless of d. Thus, adding independent $\mathrm{Lap}(2/\varepsilon)$ noise to every count gives ε-differential privacy. Each count has expected absolute error $2/\varepsilon$, and the largest error is at most $(2/\varepsilon)\log(d/\beta)$ with probability at least $1 - \beta$.

This is a case where analyzing the vector directly is much better than using composition. If we instead gave each of the d counts a privacy budget ε/d, their individual sensitivity bound of 1 would require noise of scale d/ε. The joint sensitivity analysis captures the fact that changing one record affects at most two bins.

References

[Smi23a] Adam Smith. *Lecture 4: Differential Privacy I.* Privacy in Statistics and Machine Learning, Spring 2023. 2023. (Link).

[Smi23b] Adam Smith. *Lecture 5: Differential Privacy II.* Privacy in Statistics and Machine Learning, Spring 2023. Based on materials developed with Jonathan Ullman. 2023. (Link).