Statistical Estimation of Wasserstein Distances.

- $\Omega = [0,1]^d$, let $\mu, \nu \in \mathcal{P}(\Omega)$.
- Canonical Statistical Setup: Assume μ and ν are "unknown", but we have access to:

$$X_1, \dots, X_n \overset{iid}{\sim} \mu \quad \& \quad Y_1, \dots, Y_n \overset{iid}{\sim} \nu$$

- Goal: Estimate the Wasserstein distance: $W_p(\mu, \nu)$.
- Popular Choice of Estimator:

$$W_p(\mu_n, \nu_n)$$

$$\underbrace{\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}}_{\text{Empirical Measure.}} \qquad \nu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{Y_i}.$$

$$\mu_n \xrightarrow{a.s.} \mu \text{ weakly} \qquad \int f \, d\mu_n \xrightarrow{a.s} \int f \, d\mu \quad \forall f \in C_b(\Omega)$$

$$\int f \, d\mu_n = \frac{1}{n} \sum_i f(X_i)$$

$$\Rightarrow W_p(\mu_n, \mu) \xrightarrow{a.s.} 0$$

$$\Rightarrow |W_p(\mu_n, \nu_n) - W_p(\mu, \nu)| \leq W_p(\mu_n, \mu) + W_p(\nu_n, \nu) \xrightarrow{a.s} 0$$

$$\Rightarrow W_p(\mu_n, \nu_n) \xrightarrow{a.s} W_p(\mu, \nu) \quad n \to \infty.$$

Goals :

1) Convergence rate of $W_p(\mu_n,\mu)$? [Dudley '69, AKT '80s, Boissard-Le Gouic '14]

2) Convergence rate of $W_p(\mu_n,\nu_n)$ to $W_p(\mu,\nu)$?

3) What is the limiting distribution of $W_p(\mu_n,\nu_n)$?

More recent : e.g del Barrio & Loubes '19, Niles-Weed & Rigollet '21, Chizat et al '20, etc.

I. Convergence rate of $W_p(\mu_n,\mu)$.

$W_p \geq W_1 \quad \forall p \geq 1$

Lower Bounds.

1) $$W_1(\mu_n,\mu) = \inf_{\pi} \int \|x-y\| \, d\pi$$

$$\geq \inf_{\pi} \left\| \int (x-y)\, d\pi \right\|$$

$$= \left\| \int x \, d(\mu_n - \mu) \right\| .$$

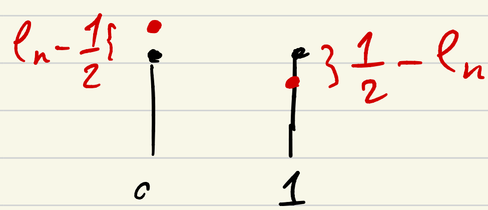
$\Rightarrow \mathbb{E}\, W_1(\mu_n,\mu) \gtrsim n^{-1/2}$. (Apply CLT)

2) $\mu = \mathrm{Ber}(1/2) = \frac{1}{2}(\delta_0 + \delta_1)$

$$\mu_n = \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} I(X_i=0)\right)}_{\ell_n}\delta_0 + \underbrace{\left(\frac{1}{n}\sum I(X_i=1)\right)}_{1-\ell_n}\delta_1 .$$

$$W_p(\mu,\mu_n) = \left(\inf_{\pi} \sum_{i,j=0}^{1} \pi_{ij}\, |i-j|^p\right)^{1/p}$$

$$= \left(\inf_{\pi} (\pi_{01} + \pi_{10})\right)^{1/p} \geq \left|\ell_n - \frac{1}{2}\right|^{1/p}$$



$$\Rightarrow \mathbb{E} W_p(\mu_n,\mu) \geq \mathbb{E}\left[\left|\frac{1}{n}\sum_i I(X_i = a) - \frac{1}{2}\right|^{1/p}\right]$$

$$\gtrsim n^{-1/2p}. \quad \text{(CLT again)}$$

3) Let μ = Lebesgue measure on Ω.

$$W_1(\mu_n,\mu) = \int \|T_n(x) - x\|\, dx$$

T_n is OT map from μ to μ_n.

$$\geq \int \min_{1\leq i\leq n} \|X_i - x\|\, dx.$$

$$\mathbb{P}\left(\min_{1\leq i\leq n} \|X_i - x\| \leq u\right) \qquad u > 0.$$

$$= \mathbb{P}\left(\bigcup_i \{\|X_i - x\| \leq u\}\right)$$

$$\leq \sum_i \mathbb{P}(\|X_i - x\| \leq u) \lesssim \sum_i u^d = n u^d$$

$$\Rightarrow \mathbb{E}\left[\min_i \|X_i - x\|\right] \geq u\, \mathbb{P}\left(\min_i \|X_i - x\| \geq u\right)$$

$$\geq u(1 - n u^d) \simeq n^{-1/d} \quad \text{for } u \simeq n^{-1/d}$$

$$\Rightarrow \mathbb{E} W_1(\mu_n,\mu) \gtrsim n^{-1/d}.$$

$$\Longrightarrow \quad \mathbb{E} W_p(\mu_n, \mu) \gtrsim \begin{cases} n^{-1/2} & \text{always.} \\ n^{-1/2p} & \text{when } \mathrm{supp}(\mu) \text{ is disconnected} \\ n^{-1/d} & \text{when } \mu \text{ has a bounded density w.r.t. Lebesgue.} \end{cases}$$

<u>Theorem</u>:

$$\sup_{\mu \in \mathcal{P}(\Omega)} \mathbb{E} W_p(\mu_n, \mu) \lesssim \begin{cases} n^{-\frac{1}{2p}} & ; d < 2p \\ n^{-\frac{1}{2p}} (\log n)^{1/p} & ; d = 2p \\ n^{-1/d} & ; d > 2p \end{cases} .$$