1 Review

- Reduction from learning to testing,
- Lower bound of $O\left(\frac{k}{\varepsilon^2}\right)$ for learning distributions over $[k],$
- Coding theoretic: $p_1, \ldots, p_{2^k}, d_{\text{TV}}(p_i, p_j) > 3\varepsilon$ for any $i, j,$
- KL divergence adds up over product distribution,
- Proving lower bound on learning Gaussian

2 Le Cam’s two point method

Given $\mathbb{P}$ and $\mathbb{Q}$ as two distribution collections, the goal is to prove lower bounds on testing $\mathbb{P}$ and $\mathbb{Q}.$ Uniformity testing is a special case when $\mathbb{P} = \{U\}, \mathbb{Q} = \{q \mid d_{\text{TV}}(q, U) \geq \varepsilon\}$ where $U$ is uniform distribution over support $[k].$

An easy attempt is to take one distribution $p \in \mathbb{P}$ and $q \in \mathbb{Q}.$ Intuitively, lower bound on testing $p$ and $q$ should be a lower bound for testing $\mathbb{P}$ and $\mathbb{Q}.$ With $p, q$ and $d_{\text{TV}}(p, q) = \varepsilon,$ we have

$$\Omega\left(\frac{1}{\varepsilon}\right) < n_{\text{test}} < O\left(\frac{1}{\varepsilon^2}\right)$$

The easy attempt above may run into trouble in complexity when the dimension of the problem increases.

A more sophisticated attempt:

Consider any distribution (priors) $\mu$ over $\mathbb{P}$ and $\nu$ over $\mathbb{Q},$ with probability $\frac{1}{2},$ pick $\mathbb{P}(u)$ or $\mathbb{Q}(u).$ For $\mathbb{P},$ we will pick $p \in \mathbb{P}$ according to $\mu.$ Then, we generate $n$ samples from $p.$ Similarly, for $\mathbb{Q},$ we will pick $q \in \mathbb{Q}$ according to $\nu.$ Then, we generate $n$ samples from $q.$

Given $X_1, \ldots, X_n$ i.i.d, we have

$$\mathbb{P}(X_1, \ldots, X_n) = \sum_{p \in \mathbb{P}} p(X_1, \ldots, X_n)\mu(p) = \mu^{\otimes n}$$

Similarly,

$$\mathbb{Q}(X_1, \ldots, X_n) = \sum_{q \in \mathbb{Q}} q(X_1, \ldots, X_n)\nu(q) = \nu^{\otimes n}$$

Now, we formally define Le Cam’s two points method:

**Definition 1.** For any $u, v,$ the number of samples needed to test between $\mu^{\otimes n}$ and $\nu^{\otimes n}$ is a lower bound on the number of samples to test $\mathbb{P}, \mathbb{Q}.$
3 Example in uniform testing

Given $\mathbb{P} = \{U\}$, $\mathbb{Q} = \{q : d_{TV}(q, U) \geq \varepsilon\}$. We can set $\mu(U) = 1$. However for $q(i)$, we have following two constructions:

We can set a distribution with $q(i)$ for $i = 1, \cdots, k$ such that

$$q(i) = \left\{ \frac{1 + 2\varepsilon}{k} \text{ or } \frac{1 - 2\varepsilon}{k} \right\}$$

We can also set $\bar{Z} = Z_1, Z_2, \cdots, Z_k \in \{\pm1\}^k$ such that

$$q_{\bar{Z}}(2i - 1) = \frac{1 + 2\varepsilon Z_i}{k}, \quad q_{\bar{Z}}(2i) = \frac{1 - 2\varepsilon Z_i}{k}$$

Based on the second construction of $q$, we have $\mu \rightarrow \{U\}$, $\nu \rightarrow \{2^k\}$ distributions. Unless $n > \sqrt{k/\varepsilon^2}$ with ($\varepsilon < 0.2$), $d_{TV}(U \otimes n, \nu q_{\bar{Z}} \otimes n)$ is small.

Now, we will introduce $\chi^2$-distance since $d_{TV}$ does not have nice property with product distribution. Given $p, q$ over $\mathcal{X}$, $\chi^2$-distance is

$$\chi^2(p, q) := \sum_{x \in \mathcal{X}} \frac{(p(x) - q(x))^2}{q(x)} = \left( \sum_{x} \frac{p(x)^2}{q(x)} \right) - 1$$