HW2 out. Due 2/7/25

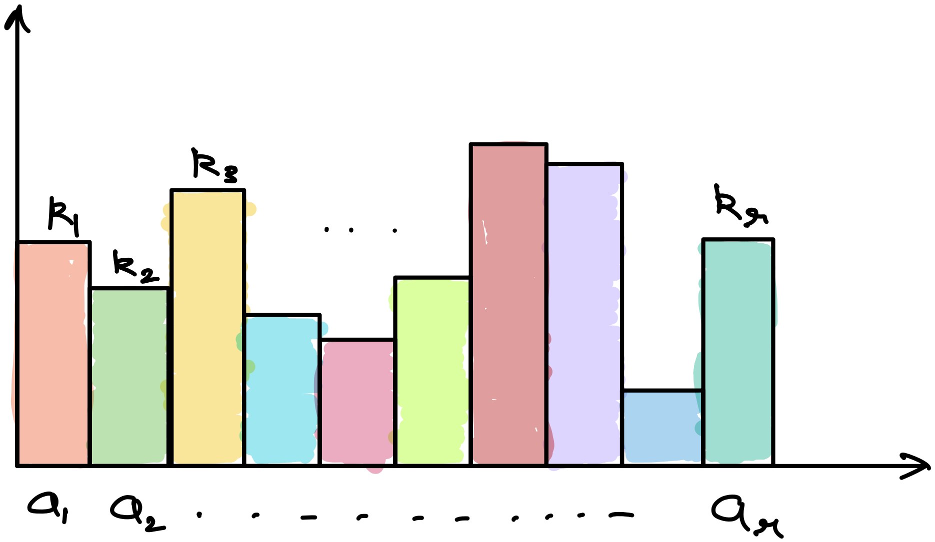
Recap

$\mathcal{X} = \{a_1, \ldots \ldots, a_r\}$

$\bar{x} = (x_1 \ldots x_n) \in \mathcal{X}^n$

$P_{\bar{x}}(a_i) = \frac{k_i}{n}$

Types:



Type class: $\frac{2^{n \cdot H(P)}}{(n+1)^r} \le |\mathcal{C}_P| \le 2^{n \cdot H(P)}$

$\{\bar{x} \mid P_{\bar{x}} = P\}$ (sequences of type P)

Concentration: $\frac{2^{-D(P\|Q)\cdot n}}{(n+1)^r} \le \Pr_{\bar{x} \sim Q^n}\left[P_{\bar{x}} = P\right] \le 2^{-D(P\|Q)\cdot n}$

Sanov's theorem

$T_n = \{ P_{\bar{x}} \mid \bar{x} \in \mathcal{X}^n \}$

$|T_n| \leq (n+1)^{r}$

- Set Π of distributions on $\mathcal{X}$

(Part I)

$$\mathbb{P}_{\bar{x} \sim Q^n} [P_{\bar{x}} \in \Pi] \leq (n+1)^{r} \cdot 2^{-n \cdot \delta}$$

$$\delta = \inf_{P \in \Pi} D(P \| Q)$$

<u>Proof:</u> $\mathbb{P}_{\bar{x} \sim Q^n} [P_{\bar{x}} \in \Pi] = \sum_{P \in \Pi \cap T_n} \mathbb{P}[P_{\bar{x}} = P]$

$$\leq \sum_{P \in \Pi \cap T_n} \underbrace{2^{-n \cdot D(P \| Q)}}_{\leq 2^{-n \cdot \delta}}$$

$$\leq (n+1)^{r} \cdot 2^{-n \cdot \delta}$$

Sanov's theorem

(Part II) If Π is the closure of an open set, then

$$\frac{1}{n}\log\left(\mathbb{P}_{\bar{x}\sim Q^n}[P_{\bar{x}}\in\Pi]\right)\to -\Big(\underbrace{\inf_{P\in\Pi} D(P\|Q)}_{\delta}\Big)$$

Proof: Consider sequence $\{P^{(n)}\}$ s.t. $P^{(n)}\in T_n\cap\Pi$ and $D(P^{(n)}\|Q)\to\delta$

$$\frac{2^{-n\cdot D(P^{(n)}\|Q)}}{(n+1)^{r}}\le \mathbb{P}_{\bar{x}\sim Q^n}[P_{\bar{x}}=P^{(n)}]\le \mathbb{P}_{\bar{x}\sim Q^n}[P_{\bar{x}}\in\Pi]\le (n+1)^{r}\cdot 2^{-n\cdot\delta}$$

$$\underbrace{-D(P^{(n)}\|Q)}_{\to\delta}+\underbrace{\frac{r\cdot\log(n+1)}{n}}_{\to 0}\le \frac{1}{n}\log\left(\mathbb{P}_{\bar{x}\sim Q^n}[P_{\bar{x}}\in\Pi]\right)\le -\delta+\underbrace{\frac{r\cdot\log(n+1)}{n}}_{\to 0}$$

Infinite Sanov

- Also holds when $\mathcal{X}$ is infinite
- Can use definition of $D(P\|Q)$ as supremum over finite partitions

$$\frac{1}{n}\log\left(\mathbb{P}_{\underline{X}\sim Q^n}\left[P_{\underline{X}}\in\Pi\right]\right)\longrightarrow -\inf_{P\in\Pi} D(P\|Q)$$

- What does a "histogram" mean!?

(Binary) Hypothesis Testing

Distinguish $(x_1, \dots, x_n) \sim \underbrace{P_0^n}_{\text{null hypothesis}}$ from $(x_1, \dots, x_n) \sim P_1^n$

Test: $T : \mathcal{X}^n \longrightarrow \{0, 1\}$

Two kinds of errors

$$\alpha(T) = \mathbb{P}_{\bar{x} \sim P_0^n} [T(\bar{x}) = 1]$$] False positive

$$\beta(T) = \mathbb{P}_{\bar{x} \sim P_1^n} [T(\bar{x}) = 0]$$] False negative

Minimizing (total?) error

▶ $\forall T : \mathcal{X}^n \to \{0,1\}$ $\quad \alpha(T) + \beta(T) \geq 1 - \delta_{TV}(P_0^n, P_1^n)$

Proof:

$$\mathbb{E}_{\bar{x}\sim P_0^n} T(\bar{x}) + \mathbb{E}_{\bar{x}\sim P_1^n} (1 - T(\bar{x}))$$

$$= 1 - \underbrace{\left(\mathbb{E}_{\bar{x}\sim P_1^n} T(\bar{x}) - \mathbb{E}_{\bar{x}\sim P_0^n} T(\bar{x}) \right)}_{\leq\, \delta_{TV}(P_0^n,\, P_1^n)}$$

Ex: Find $T : \mathcal{X}^n \to \{0,1\}$ achieving above bound

Being greedy : Optimize both $\alpha(T)$ and $\beta(T)$

- [Neyman-Pearson] : Ratio tests are simultaneously optimal

Let

$$T(\bar{x}) = \begin{cases} 1 & \text{if } \dfrac{P_1^n(\bar{x})}{P_0^n(\bar{x})} \geq \Delta \\ 0 & \text{if } \dfrac{P_1^n(\bar{x})}{P_0^n(\bar{x})} < \Delta \end{cases} \qquad (\Delta > 0)$$

Let $T' : \mathcal{X}^n \longrightarrow \{0,1\}$ be any other test. Then,

$$\alpha(T') \geq \alpha(T) \quad \text{or} \quad \beta(T') \geq \beta(T)$$

Proof:

$$T(\bar{x}) = \begin{cases} 1 & \text{if } \dfrac{p_1^n(\bar{x})}{p_0^n(\bar{x})} \geq \Delta \\[2ex] 0 & \text{if } \dfrac{p_1^n(\bar{x})}{p_0^n(\bar{x})} < \Delta \end{cases} \qquad T' : \mathcal{X}^n \to \{0,1\}$$

$$\forall \bar{x} \quad \big(T(\bar{x}) - T'(\bar{x})\big)\big(p_1^n(\bar{x}) - \Delta\, p_0^n(\bar{x})\big) \geq 0$$

≥ 0 $\quad$ < 0

≤ 0

$\sum_{\bar{x}}$

$$\underbrace{\mathop{\mathbb{E}}_{\bar{x}\sim p_1^n} T(\bar{x})}_{1-\beta(T)} - \underbrace{\mathop{\mathbb{E}}_{\bar{x}\sim p_1^n} T'(\bar{x})}_{1-\beta(T')} - \Delta\cdot \underbrace{\mathop{\mathbb{E}}_{\bar{x}\sim p_0^n} T(\bar{x})}_{\alpha(T)} + \Delta\cdot \underbrace{\mathop{\mathbb{E}}_{\bar{x}\sim p_0^n} T'(\bar{x})}_{\alpha(T')} \geq 0$$

$$\frac{\beta(T') - \beta(T)}{\alpha(T) - \alpha(T')} \geq \Delta > 0$$

Understanding ratio tests

- $$\underbrace{\frac{P_1^n(\bar{x})}{P_0^n(\bar{x})}}_{T(\bar{x})=1} \geq \Delta \iff \underbrace{D(P_{\bar{x}} \| P_0) - D(P_{\bar{x}} \| P_1) \geq \frac{1}{n}\log \Delta}_{P_{\bar{x}} \in \Pi}$$

Proof: If $P_{\bar{x}} = P$ then $\dfrac{Q^n(\bar{x})}{P^n(\bar{x})} = 2^{-nD(P\|Q)}$

$$\frac{P_1^n(\bar{x}) / P_{\bar{x}}^n(\bar{x})}{P_0^n(\bar{x}) / P_{\bar{x}}^n(\bar{x})} \geq \Delta \iff \frac{2^{-n \cdot D(P_{\bar{x}} \| P_1)}}{2^{-n \cdot D(P_{\bar{x}} \| P_0)}} \geq \Delta$$

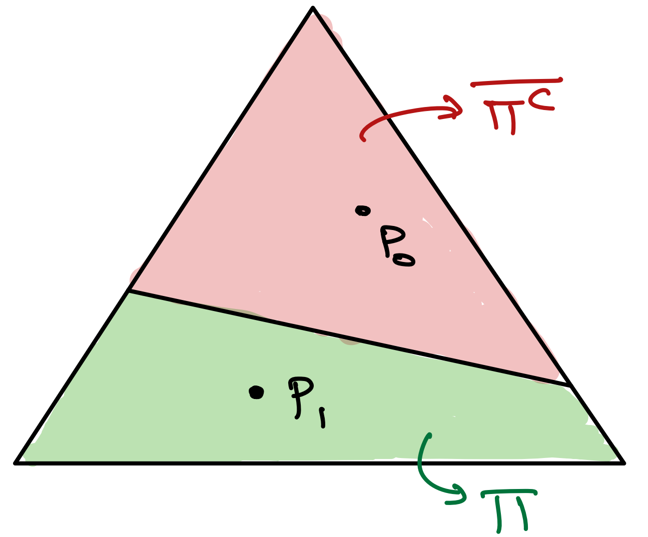
$$\Updownarrow$$

$$D(P_{\bar{x}} \| P_0) - D(P_{\bar{x}} \| P_1) \geq \frac{1}{n}\log \Delta$$

Error exponents via Sanov's theorem

$$\Pi = \left\{ P \mid D(P\|P_0) - D(P\|P_1) \geq \frac{1}{n}\log\Delta \right\}$$

$$\Pi^c = \left\{ P \mid D(P\|P_0) - D(P\|P_1) < \frac{1}{n}\log\Delta \right\}$$

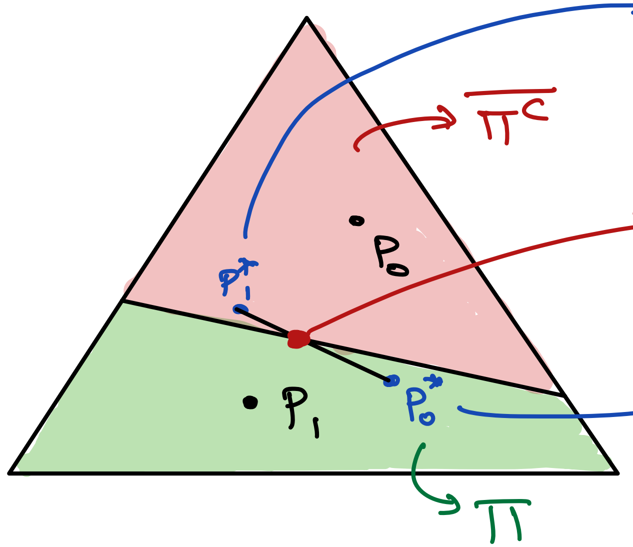


Ex: Both Π and Π^c are polytopes

$$\alpha(T) = \mathbb{P}_{\bar{x}\sim P_0^n}\left[T(\bar{x}) = 1\right] = \mathbb{P}_{\bar{x}\sim P_0^n}\left[P_{\bar{x}} \in \Pi\right] \simeq 2^{-n \cdot \inf_{P\in\Pi} D(P\|P_0)}$$

$$\beta(T) = \mathbb{P}_{\bar{x}\sim P_1^n}\left[T(\bar{x}) = 0\right] = \mathbb{P}_{\bar{x}\sim P_1^n}\left[P_{\bar{x}} \in \Pi^c\right] \simeq 2^{-n \cdot \inf_{P\in\Pi^c} D(P\|P_1)}$$

$$P_0^* = P_1^* = P^*$$



$P_1^* = \underset{P \in \overline{\Pi^c}}{\operatorname{argmin}} \, D(P \,\|\, P_1)$ } governs $\beta(T)$

If $P_0^* \neq P_1^*$, $\alpha \cdot P_0^* + (1-\alpha) \cdot P_1^*$ improves on at least one.

$P_0^* = \underset{P \in \overline{\Pi}}{\operatorname{argmin}} \, D(P \,\|\, P_0)$

governs $\alpha(T)$

<u>Ex:</u> Show that $P_0^* = P_1^*$ (If different, a convex combination on boundary improves $D(P_0^* \| P_0)$ or $D(P_1^* \| P_1)$

$$P_0^*(x) = P_1^*(x) = \frac{(P_0(x))^{\lambda} \cdot (P_1(x))^{1-\lambda}}{\sum_y (P_0(y))^{\lambda} \cdot (P_1(y))^{1-\lambda}}$$ } Can show via convex duality

λ s.t. $D(P^* \| P_0) - D(P^* \| P_1) = \frac{1}{n} \log \Delta$